Università degli Studi di Padova

Dipartimento di Scienze Storiche, Geografiche e dell'Antichità

Corso di Laurea Magistrale in Scienze Storiche

Curriculum: Mobility Studies

Exploring Sentiment Analysis on Twitter:

Investigating Public Opinion on Migration in Brazil from 2015 to 2020

Relatore: Prof. Dr. Daniel ZILIO

Laureanda:

Ana Beatriz BASELICE PELICIONI (2041703)

Academic Year 2022/2023

To my family

Quando vim da minha terra,
se é que vim da minha terra
(não estou morto por lá?),
a correnteza do rio
me sussurrou vagamente
que eu havia de quedar
lá donde me despedia.

Os morros, empalidecidos
no entrecerrar-se da tarde,
pareciam me dizer
que não se pode voltar,
porque tudo é consequência
de um certo nascer ali.

Quando vim, se é que vim
de algum para outro lugar,
o mundo girava, alheio
à minha baça pessoa,
e no seu giro entrevi
que não se vai nem se volta
de sítio algum a nenhum.

Que carregamos as coisas,
moldura da nossa vida,
rígida cerca de arame,
na mais anônima célula,
e um chão, um riso, uma voz
ressoma incessantemente
em nossas fundas paredes.

Novas coisas, sucedendo-se,
iludem a nossa fome
de primitivo alimento.
As descobertas são máscaras
do mais obscuro real,
essa ferida alastrada
na pele de nossas almas.

Quando vim da minha terra,
não vim, perdi-me no espaço,
na ilusão de ter saído.
Ai de mim, nunca saí.
Lá estou eu, enterrado
por baixo de falas mansas,
por baixo de negras sombras,
por baixo de lavras de ouro,
por baixo de gerações,
por baixo, eu sei, de mim mesmo,
este vivente enganado, enganoso.

"A ilusão do migrante" - **Carlos Drummond de Andrade**

Acknowledgements

This work represents a part of a dream I've always had. From a very young age, I knew my life wouldn't be tied to one place. My curiosity to explore the world beyond Brazilian borders has guided my decisions and actions. Although I never planned to study in Italy, I can't express enough gratitude for all the choices that have led me here today. The journey hasn't been the easiest. Studying in another country, as amazing as it is, demands a lot from you. Moving to a new country, finding a home, adapting, learning the language, building a life, dealing with bureaucracy, and being far from family and friends – it all takes a toll.

Fortunately, I've never been truly alone on this journey. My first acknowledgement, and to whom I dedicate this work, goes to my family. They have always believed in me, supported me, and cheered for me in every step of my life, and this time was no different. To them, I simply want to say thank you and that they have all my love, no matter where I am. Thanks are also due to my fiancé, Mateusz, who entered my life during this journey and embraced it wholeheartedly, providing tremendous support, especially in these past few months.

I've also been fortunate to begin this journey with a friend who, over the course of these two years, has become part of my family – my masters, Amanda. Without her daily support, shared laughter, moments of tears, love, companionship, and endless gossip sessions while watching high-quality entertainment, life in Italy wouldn't be as enjoyable and manageable. A special thanks for all the love and friendship also goes to Samuel, who was also there at the start of this journey and supported me tremendously along the way.

I'd like to extend my gratitude to my friends in Brazil, whom I miss constantly. They have been the family I've chosen for many years and have always been incredibly supportive. A big thanks to the friends I've made in Padova, especially those from the MA course and my fellow Brazilian friends who shared all the joyful and challenging moments, creating a fantastic support network and place of comfort and friendship.

My sincerest appreciation and thanks go to all the professors, tutors, and staff at the Department of Historical and Geographic Sciences and the Ancient World, especially the professors within the Mobility Studies curriculum who have consistently been supportive, open, and dedicated. A special thanks to Adriana, whom I hold very dear in my heart. Besides all the assistance with academic procedures, problems, and doubts over these two years, she has also, and most importantly, gifted me with her friendship, and for that, I am forever grateful. I'd also like to express my immense gratitude to my supervisor, Daniel Zilio, for accepting my research proposal, and for his patience, support, and guidance throughout this process.

List of Abbreviations

API - Application Programming Interface

BRICS - Brazil, Russia, India, China, South Africa group

CELAC - Community of Latin American and Caribbean States

CNDH - National Human Rights Commission (Comisión Nacional de los Derechos Humanos)

CONARE – National Committee for the Refugees (Comitê Nacional para os Refugiados)

CSV – Comma Separated Value

EU - European Union

FAO - Food and Agriculture Organization

FENAPEF - National Federation of Federal Police Officers (Federação Nacional dos Policiais Federais)

FTAA - Free Trade Area of the Americas

IBAS - India, Brazil, South Africa

ILO - International Labour Organization

IRC - Internet Relay Chat

MERCOSUR - Southern Common Market

MINUSTAH - United Nations Stabilization Mission in Haiti (Mission des Nations Unies pour la Stabilisation en Haïti)

ML - Machine Learning

MSF – Doctors Without Borders

NLP - Natural Language Processing

NGO - Non-Governmental Organization

NLTK - Natural Language Toolkit

OAS - Organization of American States

OECD - Organization for Economic Cooperation and Development

SA - Sentiment Analysis

SNA - Social Network Analysis

SXSWi - South by Southwest Interactive

UN - United Nations

UNASUR - Union of South American Nations

UNHCR/ACNUR - United Nations High Commissioner for Refugees

UNSC - United Nations Security Council

VADER - Valence Aware Dictionary and sEntiment Reasoner

WTO - World Trade Organization

URL - Uniform Resource Locator

Abstract

Technology has reshaped societal interaction and the expression of opinions. Migration is a prominent trend, and analysing social media discussions provides insights into societal perspectives. This thesis explores how events between 2015 and 2020 impacted Brazilian sentiment on Twitter about migrants and refugees. Its aim was to uncover the influence of key sociopolitical events on public sentiment, clarifying how these echoed in the digital realm. Four key objectives guided this research: (a) understanding public opinions on migrants and refugees, (b) investigating how events influenced Twitter sentiment, (c) identifying terms used in migration-related tweets, and (d) tracking sentiment shifts, especially concerning changes in government. Sentiment analysis using VADER (Valence Aware Dictionary and sEntiment Reasoner) was employed to analyse tweet data. The use of computational methods in social sciences is gaining traction, yet no analysis has been conducted before to understand the sentiments of the Brazilian population regarding migration. The analysis underscored Twitter's role in reflecting and shaping public discourse, offering insights into how major events influenced discussions on migration. In conclusion, this study illuminated the landscape of Brazilian sentiment on migration, emphasizing the significance of innovative social media analysis methodologies for policymaking and societal inclusivity in the digital age.

Keywords: sentiment analysis, text analysis, migration, social media, Twitter, Brazil.

Sommario

L'uso diffuso di strumenti tecnologici ha interconnesso i mondi online e offline, sfumando i confini tra di loro e portando a cambiamenti trasformativi nell'impegno con la società e le istituzioni. Le piattaforme dei social media, in particolare Twitter, sono emerse come componenti centrali della vita quotidiana, fungendo da canali dinamici per lo scambio di opinioni ed emozioni, trascendendo i limiti geografici e accumulando un ampio archivio di discorsi pubblici. All'interno di questo dinamico ecosistema digitale, questa tesi si propone di esplorare come una serie di eventi nazionali e internazionali verificatisi tra il 2015 e il 2020 abbiano influenzato il sentimento pubblico brasiliano espresso su Twitter riguardo a migranti e rifugiati. L'obiettivo centrale è stato quello di decifrare l'impatto di questi eventi sui sentimenti espressi dal pubblico brasiliano su Twitter, arricchendo così la comprensione di come i principali eventi sociopolitici si siano riverberati nel paesaggio digitale.

Al centro del perseguimento di questo scopo di ricerca c'erano quattro obiettivi fondamentali: (a) la delucidazione delle opinioni pubbliche prevalenti riguardo a migranti e rifugiati, (b) l'indagine di come vari eventi internazionali e nazionali si siano correlati e abbiano influenzato i sentimenti sfumati espressi su Twitter durante il periodo di tempo specificato, (c) l'identificazione e la quantificazione delle terminologie specifiche utilizzate nel discorso su Twitter riguardo a migrazione e rifugiati, e (d) il discernimento di notevoli cambiamenti nel sentiment, in particolare per quanto riguarda i cambiamenti nella leadership governativa. Per svolgere questo compito, è stata utilizzata *sentiment analysis* con VADER (Valence Aware Dictionary and sEntiment Reasoner) per analizzare i dati dei tweet estratti nell'arco di tempo stabilito. L'uso di metodi computazionali nelle scienze sociali si sta diffondendo sempre di più, in particolare nel campo delle migrazioni, ma finora non era stata condotta alcuna analisi per comprendere i sentimenti della popolazione brasiliana riguardo a questo tema. Pertanto, questa tesi si propone di contribuire alla letteratura sulla *sentiment analysis* applicata agli studi sulla migrazione, evidenziando l'importanza di metodi di ricerca diversi e arricchendo la comprensione delle dinamiche sociali.

Lo studio ha rivelato che, contrariamente alle aspettative, i sentimenti negativi non hanno dominato, anche nel mezzo di un cambiamento verso un governo più conservatore; al contrario, hanno prevalso la neutralità e la positività. L'analisi ha sottolineato il ruolo di Twitter nel riflettere e plasmare il discorso pubblico, offrendo spunti cruciali su come i principali eventi abbiano influenzato le discussioni sulla migrazione. In conclusione, questo studio ha illuminato il panorama in evoluzione del sentimento brasiliano sulla migrazione, sottolineando l'importanza di metodologie innovative di analisi dei social media per la definizione delle politiche e l'inclusività della società nell'era digitale.
Parole chiave: sentiment analysis, text analysis, migrazione, social media, Twitter, Brasile.

Table of contents

Introduction

Advancements in technology and the proliferation of connected socialization environments have led to a significant transformation of human social life in recent decades. The constant and massive flow of data and information, as well as the active participation of individuals across the globe in its production and dissemination, has fundamentally altered human interactions, making us rethink static models to understand society (Bouvier 2015). We can no longer speak of two distinct worlds; online and offline are inextricably connected and influence one another in a variety of ways. As argued by Van Dijck and Poell (2013) "social media platforms have penetrated deeply into the mechanics of everyday life, affecting people's informal interactions, as well as institutional structures and professional routine." Empirical research has demonstrated the importance of understanding how online and offline behaviours interrelate, and how placing distinctions between them significantly limits our capacity to explain and comprehend the communication processes within (Bouvier 2015).

The internet has provided spaces such as social network platforms where people share views, emotions, and sentiments towards products, people, and life in general (Patodkar and Sheikh I.R 2016). These platforms evolved as the most popular environment for users to voice their opinions about events happening around the world (Y. Chen, Sack, and Alam 2021), or about personalities, politicians, products, companies, etc (Luo, Osborne, and Wang 2015). Likewise, people usually share many aspects of their social lives on the platforms, providing a useful, rich, and up-to-date source of information (Inuwa-Dutse, Liptrott, and Korkontzelos 2020). Analysing the massive amount of data produced becomes critical to understanding social behaviours, human interactions, and their outcomes. It has also given rise to whole new fields of social sciences and humanistic research to study various aspects of social phenomena on online platforms (Burgess and Baym 2020).

As social networking, blogging, and micro-blogging websites gain popularity, they have grown into a vast archive of textual data. Officially launched in July 2006 in the United States, Twitter[1] has fallen in the popular taste, being defined as a "nervous system for the planet" and a "global newsroom" (Burgess and Baym 2020). In December 2022 it was claimed to have over 368 million monthly active users worldwide (Dixon 2022). Known as a micro-blogging platform, which is a variant of a blogging platform, Twitter allows users to post quickly short updates (until 2017 constrained to 140 characters in a single tweet), creating an innovative communication method (Ross et al. 2011). It offers a new way of exchanging lightweight information and updates, that is allied

---

[1] This thesis was initiated prior to Elon Musk's acquisition of Twitter, and it will consistently utilize the term "Twitter" to denote the social platform rather than the platform's new name "X."

with its features such as replies, hashtags (#), mentions (@), and retweets (RT), constituting a hybrid environment of blogging, instant messaging, social networking, and status notification (Ross et al. 2011).

Studies on Twitter have placed identity and self-representation at its core, and how they position people against others, as part of evaluating and legitimising broader types of identities and social processes (Bouvier 2015). In her piece on the importance of discourse studies engaging with theories and empirical work on social media, Gwen Bouvier (2015) suggests that these online platforms are used for a combination of different purposes, including identity construction, the maintenance of social relationships, and engagement with more socially relevant matters.

Initially thought as a service to update your friends about your daily activities and thoughts, Twitter also became a place where people get updates on recent news, share information, and actively engage with social issues. Since the early days of Twitter, personalities like politicians, activists, and celebrities, as well as leading international media outlets such as BBC, Al Jazeera, and The New York Times (Burgess and Baym 2020), have been using their accounts to share news headlines, provide updates on government actions, and express opinions. These tweets reach millions of people's timelines within seconds. Users use comments, retweets, and hashtags as tools to engage in specific conversations, spread their opinions and even mobilise for a determined cause. Beginning in the United States and later spreading worldwide, an important example of media activism was the #BlackLivesMatter that from 2013 to 2020, mobilised around 390 million *tweets* (Bianchi 2020). People's participation on Twitter in this case was linked to significant offline events, which made the horrifying racism incidents and protests that were taking place gain more visibility, strengthened the movement, and allowed us to observe how individuals were feeling and what they were thinking about was happening.

With hundreds of millions of people expressing their opinions and involving themselves in the most diverse topics and discussions worldwide, this rich environment provides academic researchers with a plethora of social information with great potential to improve opinion retrieval performance in research (Luo, Osborne, and Wang 2015). Decades of data obtained through conventional survey methods have been and are still being used to understand public opinion, however, we must recognise that successful solutions to issues affecting current society's well-being cannot afford to disregard how the online public views those issues (Inuwa-Dutse, Liptrott, and Korkontzelos 2020). Due to the large amount of data available to be extracted and analysed, it becomes hard or even impossible for this task to be accomplished by humans. Therefore, new tools and methodologies unifying computational sciences and other academic fields have been developed to achieve this aim. Sentiment analysis (SA) has garnered enormous popularity and it is used in many

analytical fields (Ribeiro et al. 2016), being one of the most widely studied applications of natural language processing (NLP) and machine learning (ML) (Patodkar and Sheikh I.R 2016).

Sentiment analysis, also known as opinion mining or contextual mining, is used in NLP, computational linguistics, and text analysis as a process that studies public opinion about an entity, identifying, systematically extracting and quantifying, subjective information (Sharma et al. 2020). It can be applied for example to use data from blogs and social media websites such as Twitter, where people freely express their views about a certain topic and categorize sentiments as positive, negative, or neutral. It has been used in different fields and institutions, such as private companies, to identify people's opinions about a specific product (Jagdale, Shirsat, and Deshmukh 2019; Fang and Zhan 2015). While we cannot disregard most usual and traditional methods of measuring public attitudes, such as interviews and surveys, the voice of public opinion has been enhanced by the development of Web 2.0[2] and debates have gained a new stage inside social media. Surveys can offer a trustworthy national representation but have some pitfalls regarding cost, geographical scales, frequency, and fast data obsolescence (Y. Chen, Sack, and Alam 2022; Rowe et al. 2021). By aggregating new methods for gathering people's opinions in such a dynamic and open space as Twitter, it is possible to reduce some of these methodological constraints of traditional data sources, which are also related to the amount of data that can be obtained, collected, processed, and analysed.

Sentiment Analysis and the COVID-19 Pandemic

More recently, the measurement of public perception on contentious issues including, the COVID-19 pandemic, migration, and climate has emerged as an increasingly important concern for social scientists (Y. Chen, Sack, and Alam 2022). The COVID-19 pandemic has brought about unprecedented challenges and complexities, impacting multiple aspects of life around the world. The use of sentiment analysis has become increasingly relevant in understanding the public's perception and emotional responses during this time. With the internet being an integral part of daily activities, sentiment analysis has become an essential tool for organizations and governments to gain insights into user feedback and make informed decisions. The research conducted during this period was an asset to provide guidance for policy-making that could be implemented in social media platforms, for example, identifying the appropriate level of moderation required to combat misinformation on social media. Numerous studies have been conducted worldwide to investigate various aspects of people's sentiments concerning the COVID-19 pandemic. These investigations include exploring opinions on the pandemic and vaccines (Ansari and Khan 2021), assessing vaccine acceptance (Alabrah et al. 2022; Kapoteli et al. 2022), examining perceptions and feelings about lockdown measures (Gupta et

---

[2] Also known as participatory web or social web. It refers to websites that have predominantly user-generated content, participatory culture, and interoperability (see https://history-computer.com/web-2-0/).

al. 2021), investigating vaccine side effects (Ferawati et al. 2022), as well as comparing the opinions about different vaccines (Marcec and Likić 2021). In addition, a dataset focusing on COVID-19 sentiment was created to facilitate research (Naseem et al. 2021). Specifically, this study created a new extensive sentiment dataset, COVIDSENTI, comprising 90,000 tweets associated with the pandemic, which was generated during the early phases of the outbreak, between February and March 2020.

Further studies were made as literature reviews to specifically analyse the development of the techniques and approaches applied in sentiment analysis during the period. Alromema's (2022) study focuses on reviewing existing applications of sentiment analysis on Twitter data that identify major sentiment trends associated with the push to reopen during the pandemic. His study explores the most popular techniques and approaches in sentiment analysis in research carried out during the pandemic. With a similar aim to better understand the various methods used to collect, pre-process, and classify sentiments the study carried out by Kartikasari Kusuma Agustiningsih, Ema Utami, and Hanif Al Fatta (2021) focuses on reviewing research papers that employ sentiment analysis to study COVID-19 vaccine sentiment using Twitter data.

Sentiment Analysis and Migration Studies

The increasing development of the field of social media analysis, mainly in the last few years, represents an opening possibility for applying these techniques to various subjects of interest. Migration, which is the topic of interest of this study, has become a widely discussed subject over the years due to the wide variety of displacements occurring around the world and how these flows affect global society on social, political, and economic levels. Recent studies have explored sentiment analysis and natural language processing techniques in the context of social sciences research, particularly concerning migration studies, trying to develop a framework to predict migratory flows and grasp people's opinions towards migrants.

Identifying and understanding online attitudes concerning migration can be an effective method for improving the protection of migrants and refugees, combating misrepresentations, and offering different perspectives on how to recognise and undermine hate speech and discrimination. Public sentiment towards migrants is an essential component of migrant inclusion and social cohesiveness, influencing policies and the ability of migrants and refugees to successfully participate in receiving communities. Actions of prejudice, intolerance, and xenophobia can make it difficult for immigrants to find work, housing, and a sense of belonging in their communities, which can lead to greater social problems (Rowe et al. 2021).

Böhme, Gröger, and Stöhr (2020) address the scarcity and inconsistency of migration data, especially in developing countries, by utilizing geo-referenced online search data to measure

migration intentions. Through fixed effects panel models, machine learning, and prediction techniques, the study demonstrates the effectiveness of the approach in predicting international migration flows. Using survey data to support the validity of the measures in reflecting genuine emigration intentions, the findings contribute to novel methods for measuring migration intentions and improving the performance of conventional migration models in prediction tasks. Also tackling the inconsistency of migration flow data and the limitations of traditional census methodologies, Zagheni et al. (2014) consider the availability of geolocated data from online sources, such as Twitter, as an opportunity to track migration patterns and explore the relationship between internal and international migration. The study uses geolocated Twitter data from OECD (Organisation for Economic Cooperation and Development) countries between May 2011 and April 2013 to examine users' geographic movements and the findings demonstrate some promising potential to predict migration trends and improve understanding of migration dynamics, highlighting the relevance of publicly available data for monitoring migration trends and suggests future research directions.

Proposing innovative methods to add to traditional survey methods, which might contain limitations in capturing comprehensive and timely data regarding people's opinions, Rowe et al. (2021) research emphasizes the importance of addressing misperceptions and misrepresentations of immigration, which can lead to intolerance and discrimination. The study uses Twitter data and NLP techniques to measure and monitor sentiment towards immigration, discussing research strategies, statistical representation, data access permissions, sentiment scoring, and ethical concerns related to this kind of research methodology. Along the same line, Inuwa-Dutse, Liptrott, and Korkontzelos (2020) examine public opinions on migrants and refugees expressed on Twitter, focusing on the context of the EU refugee crisis. Their findings reveal a slightly higher proportion of negative sentiments towards migrants, particularly among ordinary users, however, users with a large number of followers and non-governmental organizations (NGOs) tend to express more positive sentiments. Using a wider range of data sources, such as major newspapers and discussion forums along with Twitter, Yantseva's (2020) study analyses media discourse on migration in Sweden between 2012 and 2019. The study tries to comprehend the themes and terminology used in media discussions about migration, especially considering the refugee crisis in Europe. The findings indicate a predominance of negative tones in social media communications and the existence of language that is antagonistic to immigration in media contexts, even if, however, the study does not show an impact of the refugee crisis on the media's discourses. Yiyi Chen, Harald Sack, and Mehwish Alam's (2021) research analyses sentiments and hate speech on Twitter, also to investigate popular opinions toward migration. The study introduces MigrationsKB (MGKB), a knowledge base with annotated migration-related tweets from European countries hosting immigrants between 2013 and July 2021. Through advanced techniques like topic modelling and sentiment analysis, they processed and

annotated the tweets, also considering external databases to identify factors influencing negative attitudes towards migration. The findings supported the hypothesis that countries with more migrants tend to have fewer negative and hateful tweets.

Objectives and Framework of the Research

The field of sentiment analysis and computational text analysis offers promising avenues for studying migration-related topics. By analysing sentiments on social media platforms like Twitter, researchers can gain valuable insights into public attitudes towards migration, media discourse, and prediction of migratory flows. These approaches help address the limitations of traditional data sources and offer real-time monitoring of migration trends. Additionally, they demonstrate how social media may be used to create a safe space for social analysis that combines important political, economic, and social events with in-progress debates on these platforms. As previously mentioned, segregating the world into online and offline spaces ignores the significant influence and importance of social media on modern culture. It also dismisses a wide range of spontaneous perspectives posted online every day. However, it is critical to recognize the representativeness constraints of Twitter data. It would be wrong to assume that all opinions expressed on Twitter, or any other social media platform, represent the views of entire nations or societies. However, one cannot dismiss the significance of these findings and the extensive field of investigation they open up regarding the impact of crises, their consequences, responses to mass media and events, and people's sentiments involving it all.

Building bridges across disciplines enables us to establish a shared set of dependent and independent variables. This, in turn, facilitates the pursuit of explanations and elucidation within the framework of coherent models. These models aim to foster a comprehensive understanding of two primary facets: migrants' behaviours and the reactions of states and societies towards migrants (Brettell and Hollifield 2023). In this sense, this dissertation aims to analyse Brazilian's sentiments towards migrants and refugees, especially examining whether national and international events that occurred between 2015 and 2020 had an impact on the variation of people's sentiments. To answer this, the overarching objectives are fourfold:

(a) To understand public opinions towards migrants and refugees.

(b) To discern whether a spectrum of international events (e.g., the Syrian war, the election of Donald Trump and his anti-immigration policies, the Venezuelan crisis, the European immigration crisis, etc.) and national occurrences (e.g., Dilma Rousseff's impeachment, the election of Jair Bolsonaro, etc.) exerted influence over the nuances of sentiments expressed on Twitter concerning migrants and refugees during the specified period.

(c) To identify and quantify specific terms employed about migrants and refugees within the Twitter discourse.

(d) To ascertain whether discernible shifts occurred, indicating a transition from a predominantly positive sentiment to a negative one, or vice versa, particularly in response to changes in the government.

Aligned with the aforementioned studies using Twitter data and sentiment analysis, this dissertation aims to contribute to the existing literature on sentiment analysis applied to the topic of migration. Existing studies on this subject have primarily focused on Europe, particularly concerning European migration flows and the recent refugee crisis. This research on Brazilian sentiments towards migrants and refugees, contextualized within the broader political landscape of Brazil, is therefore, of significant importance for academia, policymakers, and also for changing the research focus to the Global South. From an academic perspective, the research offers a detailed analysis of the interplay between politics, social media, and public opinion, and sheds light on the complex factors that influence attitudes towards migrants and refugees. The majority of sentiment analysis research combined with social sciences in the Brazilian context has focused on the social media discourses about political themes, such as elections (M. S. Amaral and Pinho 2018; D. K. S. Santos 2023) and corruption (Paiva, Garcia, and Alcântara 2017), the COVID-19 pandemic (Gomes Pessanha et al. 2020; De Melo and Figueiredo 2020), or other social problems and events. Still, there is a dearth of research applying sentiment analysis to migration studies in South America and more specifically, no research was found by now focusing on migration and sentiment analysis in the Brazilian framework.

Furthermore, this research can also contribute to the development of more nuanced and effective policy interventions to support the integration of migrants and refugees into Brazilian society. It can offer valuable insights into the challenges and opportunities associated with managing migration in a rapidly changing political and social landscape inside social media platforms, helping in the development of more effective policy interventions that promote social cohesion, equity, and inclusion for migrant and refugee populations. Moreover, as previously mentioned, social media presents a vast reservoir of information that has yet space to be explored by the field of migration studies. This includes not only the potential application of sentiment analysis and its various methodological approaches but more specifically, the use of this analytical tool to investigate public attitudes towards migrants and refugees on social media concerning a wide range of contexts and social, political, and economic events.

The Importance of Twitter in the Brazilian Social and Political Arena

According to DataReportal (Data Reportal 2023a), an estimated 4.80 billion social media users existed worldwide in April 2023, accounting for approximately 59.9% of the global population. Following the United States, Japan, and India, respectively, Brazil is the fourth on the list ranking the number of Twitter users by country, where there are 24.3 million Twitter users (Data Reportal 2023b). In 2014, Twitter emerged as one of the prominent platforms, alongside Facebook, employed by individuals to mobilize Brazilian society against President Dilma Rousseff's government. Notably, certain hashtags, such as *#vemprarua* ("come to the street") and *#ogiganteacordou* ("the giant has awakened"), inundated the Twitter trending topics for an extended period, fuelling the population to protest against the escalating prices of public transportation and the corruption scandals associated with Operation Car Wash - a wide-reaching investigation into money laundering and bribery.

These movements gained significant interest, capturing the attention of both the national and international media, as they reflected widespread discontent with the government's handling of public funds and exposed deep-rooted issues of corruption within the political landscape of Dilma's and previous governments. Consequently, these movements ultimately led to President Dilma's controversial impeachment in 2016, aligning with the prevailing right-wing turn in Latin America at that time, and facilitated the ascent of Vice-President Michel Temer's interim government, followed by the subsequent rise of Jair Bolsonaro and the so-called *Bolsonarismo*[3] in 2018.

This pivotal juncture in Brazilian political history effectively exemplifies the growing power and influence of social media within society and in the political arena. The events in Brazil during that period were often compared to the developments in the United States, particularly the rise of President Donald Trump. The examination of social media's impact on political outcomes was evident in the context of Trump's election, and similar dynamics could be observed later in Brazil, powering ideologies like Trumpism and *Bolsonarism*.

Migration in Brazil and Social Media

Furthermore, during this time of political instability, Brazil experienced an influx of migrants and refugees. Several macro factors contributed to this phenomenon. Firstly, the international economic crisis that began in 2009, prompted an increase in south-to-south human mobility, with Brazil becoming a destination country (Tonhati, Cavalcanti, and De Oliveira 2022). Secondly, Brazil's social and economic and social development and its international image as an emerging power also served as an incentive for migratory flows to the country.

---

[3] Bolsonarismo (or Bolsonarism), is a term used to refer to the ideology of the Brazilian ex-President, Jair Bolsonaro. Bolsonarism is a far-right movement backed by political and social conservatives and evangelicals, characterized by radical anti-leftist ideals and wild neoliberalism. It has been compared to Donald Trump's ideology and even called the "Brazilian version of Trumpism".

Additionally, specific global events redirected the flows to Brazil. Since 2011, the ongoing Syrian war has forced more than 14 million Syrians to flee their homes (UNHCR 2022), resulting in several asylum requests in Brazil in 2016, despite the geographical distance between the two countries (Manfrinato and Souza 2019). Brazilian consulates in the Middle East introduced simplified visa procedures in 2013, facilitating travel to the country to seek asylum (Langlois 2016). Moreover, the catastrophic earthquake that destroyed Haiti in 2010, displacing approximately 1.5 million people, led many Haitians to search for asylum in Brazil. Political repression in the Democratic Republic of the Congo and Venezuela's political and economic crisis further contributed to south-south migrations towards Brazil (Osava 2018).

These diverse migration flows prompted Brazil to adapt its migration regulations and policies to address the emerging challenges and demands. This transition occurred during this period of political instability, polarisation, shifting ideologies within the government, which created some obstacles to the approval of the new policies. For instance, the New Law of Migration (Law 13.445/2017) which aimed to replace the Foreigner Statute Law (Law 6.815/1980), encountered a three-year approval process. Notably, public discourses against migration gained prominence after President Dilma's impeachment in 2016 (Conectas 2019). Some hashtags such as *#NovaLeideMigraçãoJá* ("new migration law now") and *#MigrarÉDireito* ("to migrate is a right") gained popularity, advocating for the new law. Nevertheless, the law faced substantial vetoes from President Michel Temer.

From Dilma Rousseff's supportive stance on migrants and refugees in 2015, when she published an article in the newspaper Folha de São Paulo expressing Brazil's willingness to welcome refugees, to President Jair Bolsonaro's derogatory views about migrants and refugees in a 2015 interview, referring to them as "the scum of the world" in an interview (Azevedo 2015), there has been a notable shift in the discourse related to migration. This shift became even more evident when Bolsonaro assumed the presidency in 2019, as one of his initial actions was to announce Brazil's withdrawal from the Global Pact of Migration. Bolsonaro argued that migration should be treated as a domestic and security issue rather than a global concern (Aguiar 2020).

Consequently, the convergence of exogenous and endogenous factors resulted in increased migration flows to Brazil, fuelling debates and shaping opinions both in the online realm and the offline world. As the demographic and political landscapes are changing, together with the increased use of social media platforms by individuals to express their feelings and perspectives, examining the online conversations and interactions surrounding migration on social media platforms assumes great importance. Such analysis allows a comprehensive understanding of the multifaceted factors that influence public sentiment towards migrants and refugees in Brazil.

Therefore, this research aims to contribute to the field of sentiment analysis methodology, particularly in its intersection with social research, to enhance our understanding of the intricate interplay between social media, public opinion, and migration in contemporary society. By exploring analytical approaches that leverage sentiment analysis to comprehend public attitudes towards migration and refugees, this study seeks to provide a more nuanced and diverse understanding of the social, political, and media influences that shape public opinion about migrants and refugees in Brazil. The objectives to be achieved draws from the hypothesis that national and international events that occurred between 2015 and 2020 might had an impact on the sentiments expressed on Twitter regarding migrants and refugees in Brazil. This impact resulted in fluctuations of sentiment, potentially shifting sentiments from positive to negative or vice versa, particularly in response to changes in government policies and discourses. The study aims to analyse these fluctuations, understand public opinions toward migrants and refugees, quantify specific terms used about them on Twitter, and identify any patterns or trends in sentiment changes over this period.

To achieve this goal, Twitter data was collected via Twitter's application programming interface (i.e., API). The API enables data exchange between applications and services to be done programmatically (via code) and provides access to tweets, users, direct messages, lists, trends, media, and places in a structured manner (Gottschalk and Pichierri 2022). An initial application was sent to obtain a Twitter developer account to gain access to the data. Once the application has been accepted, it was possible to access the standard product and then apply for the academic research product, which granted access to the Twitter API v2 and allowed one to access historical public tweets without charge.[4] As there is no pre-established dataset with tweets about migration in the Brazilian Portuguese language, a new dataset was created using a query with topic-related keywords. Thus, the tweets of that specific timeframe and with specific keywords were retrieved and stored in a dataset. After the crawling, the data were then pre-processed, and spamming content was identified and eliminated to ensure that it was gathered only from tweets associated with the study, with the aid of selected hashtags. Subsequently, the text was processed and translated to be compatible with the chosen program, VADER (Valence Aware Dictionary and Sentiment Reasoner). VADER is a sentiment analysis tool that is designed to assess the sentiment (positive, negative, or neutral) of text data, primarily in the English language. It uses a pre-built lexicon of words and their associated sentiment scores to analyse the sentiment of a piece of text. VADER is commonly used in natural language processing tasks to gauge the emotional tone of social media posts, reviews, and other forms of text data. Additionally, during the data analysis process, the in-development Portuguese-language

---

[4] Recent news announced that this is about to change, and it will be further explained in the 2nd Chapter.

Python libraries necessitated some personal knowledge of Portuguese to provide inputs to the program for a better textual analysis.

In order to have an organised and comprehensive understanding of the topic and the results, the thesis was divided into four chapter and conclusion. Chapter one provides an introductory overview of recent migration flows to Brazil and situates these trends within a broader context of international events that have contributed to the movement of people to Brazil between 2010 and 2020. This chapter also explores the internal political landscape of Brazil during this period, highlighting key events, laws, and phenomena that have shaped the country's governance under Dilma Rousseff, Michel Temer (transitional government following the 2016 coup), and Jair Bolsonaro. The chapter examines the growing role of social media in Brazilian politics and world migration, with a particular emphasis on Twitter.

Chapter two offers a detailed analysis of the social media platform Twitter, including its history, development, and growth. The chapter explores the platform's most important features and how they facilitate connections between users in a fast-paced social environment, as well as its significance in the propagation of opinions online. Additionally, the chapter discusses academic access to Twitter data, the types of research that can be conducted on Twitter, the opportunities for academic research, and recent modifications. The chapter concludes with a discussion of ethical considerations related to the use of Twitter data.

Chapter three outlines the methodology employed in the research project, beginning with a brief explanation about Python and Jupyter Notebook, the environment chosen to develop this research. It also comprises an explanation about Python libraries used along data pre-processing and analysis, about sentiment analysis methodology and the tool applied to do it. Also, about how the academic access was obtained and the process of data collection.

Chapter four proceeds through the various stages of pre-processing in order to clean the undesired tweets that could spam the results, and the text processing steps related to cleaning the text to obtain the best results from the sentiment analysis. The chapter provides a comprehensive description of the processes and the results obtained, both along the pre-processing part and after the sentiment analysis.

Finally, the conclusion presents the research findings, highlights key insights, and identifies any potential limitations or areas for further investigation. The study reveals negative sentiments did not dominate; instead, neutrality and positivity prevailed. The analysis underscores Twitter's role in reflecting and shaping public discourse, offering crucial insights into how major events influence discussions on migration. This research contributes to sentiment analysis in migration studies, highlights the importance of diverse research methods, and enriches our understanding of societal dynamics. Despite limitations, such as language constraints, it opens avenues for further

interdisciplinary exploration. The chapter further concludes with a discussion of the implications of the research and potential avenues for future research in this field.

Chapter 1: Migration, historical contextualization, and social media

1. International migration, Brazil, and Social Media

The movement of people is not a new phenomenon in a global society. The distribution and mixing of populations throughout history have largely been the result of human migration, a process that continues to persist today and will continue in the future. However, a notable difference between ancient times and the present lies in the various legal, physical, and ideological barriers that hinder individuals' desire to migrate.

In recent years, there has been a significant increase in migration, as people seek better opportunities, safety, and refuge in different regions. From 2015 to 2020, public and media attention was focused on the millions of refugees attempting to enter Europe. In our highly interconnected world, up-to-date information about unfolding events was disseminated through various channels, including websites and social media platforms. At the same time, the countries involved, whether as destinations or origins of these migrations, grappled with the task of making decisions, adapting regulations, and managing the influx of people. Since 2015, many countries have implemented direct and indirect policies to mitigate and discourage international migration, reinforcing security.[5] However, as it is possible to see in everyday news, individuals continue to seek refuge in different countries, fleeing their nations for various reasons, most notably to escape the horrors of war and persecution (citation needed for this). Distressing images of people on the move were communicated by the global media, while systemic barriers were erected in Europe, and physical barriers were being constructed in the United States, as promised by Donald Trump during his 2016 presidential campaign. Social platforms like Twitter became significant mediums for discussing these global events, migratory processes, and their consequences for people and nations.

Not far from the global scenario, Brazil experienced an unexpected influx of migrants and refugees during a five-year period (2015 – 2020). This coincided with a time of political turmoil in the country that reshaped its political landscape and migration policies. As this thesis aims to answer questions about migratory flows to Brazil and the public perception of migrants and refugees, it is pivotal to understand what happened. In this chapter sheds light on Brazil's political and migration scenario, providing an overview of migration flows and major events worldwide between 2015 and 2020, with a specific focus on migratory flows to Brazil during that period. The chapter explores the

---

[5] For more information about the debate of securitization refer to Bello and Léonard (2023).

complexities of events in Latin America that led to South-South migration, examining factors such as economic conditions, political instability, and violence that influenced migration patterns. Furthermore, the chapter analyses the political landscape within Brazil, highlighting the shift from left-wing to right-wing governments. It discusses the wave of protests that arose against the government in Brazil, and the developments of Operation Car Wash, which evolved into the impeachment of Dilma Rousseff's government. The subsequent rise of Michel Temer and the election of Jair Bolsonaro as president will be examined with a special focus on the topic of migration, emphasizing the policies and approaches implemented at that time to manage migration flows.

In addition to political shifts, this chapter investigates the role of social media and Twitter in shaping public discourse and political movements in Brazilian society. It explores the social networks and media strategies employed to establish relationships with the public, as Twitter emerged as a prominent platform for conversations and discussions about governmental events. By examining the intricate relationship between migration, politics, and media in Brazil, this chapter proposes a comprehensive understanding of the migration scenario worldwide, but especially the flows to Brazil and their consequences. This understanding is essential for comprehending the data obtained in the next steps of the research.

International Panorama and Migration Flows

International movements are increasingly facilitated by the advancements of the digital revolution, distance-shrinking technology, and more affordable travel costs (United Nations 2018). This accessibility has opened up new possibilities for people to migrate across borders and as globalization deepens and becomes more intricate, the resulting transformations increasingly shape the dynamics that define people's lives.

These dynamics are interconnected, ranging from individual choices to collective decisions, occurring simultaneously on macro and micro levels. Economic prosperity, inequality, demographic shifts, violence, conflict, environmental changes, and other factors permeate the decision to stay or leave, to relocate or remain stationary, and for how long. While the majority of international migrants relocate for reasons related to work, family, and education (United Nations 2018), some are compelled to leave their homes and countries due to conflicts, persecution, and natural disasters. The complexities of migration and the diverse reasons behind it highlight the intricate tapestry of human mobility in today's world.

In accordance, the emergence of a new international division of labour demands extensive economic restructuring at all global, national, and local levels. This restructuring process requires the development of novel spatial configurations and the restructuring of social phenomena previously restricted to boundaries within nations (Sassen 2007). As the localities are inserted into the global

logic, international migration will tend to correspond to the transformations arising from the restructuring of global capitalism, the insertion of countries in world geopolitics and their migratory policies, and the consequent role that such localities assume in the international division of labour (Baeninger and Peres 2017). For this study, the term "international migration" or "migratory flows" will be used comprehensively, encompassing various forms of human movement across borders. This includes forced migration flows resulting from conflict, persecution, or environmental factors, as well as the movements of asylum-seekers, refugees, and other individuals who have been displaced from their homes. By adopting this inclusive approach, we aim to capture the full spectrum of migration dynamics and understand public opinions about any kind of human mobility.

In the past few years, the world has witnessed significant migration and displacement events. The period within the second decade of the 21st century has been marked by major migration movements and forced displacements that have garnered international attention, such as the conflicts in the Syrian Arab Republic, Yemen, the Central African Republic, the Democratic Republic of the Congo, and South Sudan; the cases of extreme violence, such as with the Rohingya people that live in Myanmar and had to flee to Bangladesh; and economic and politic instability that has driven millions of Venezuelans to leave their country in search of better opportunities and safety (World Migration Report 2020).
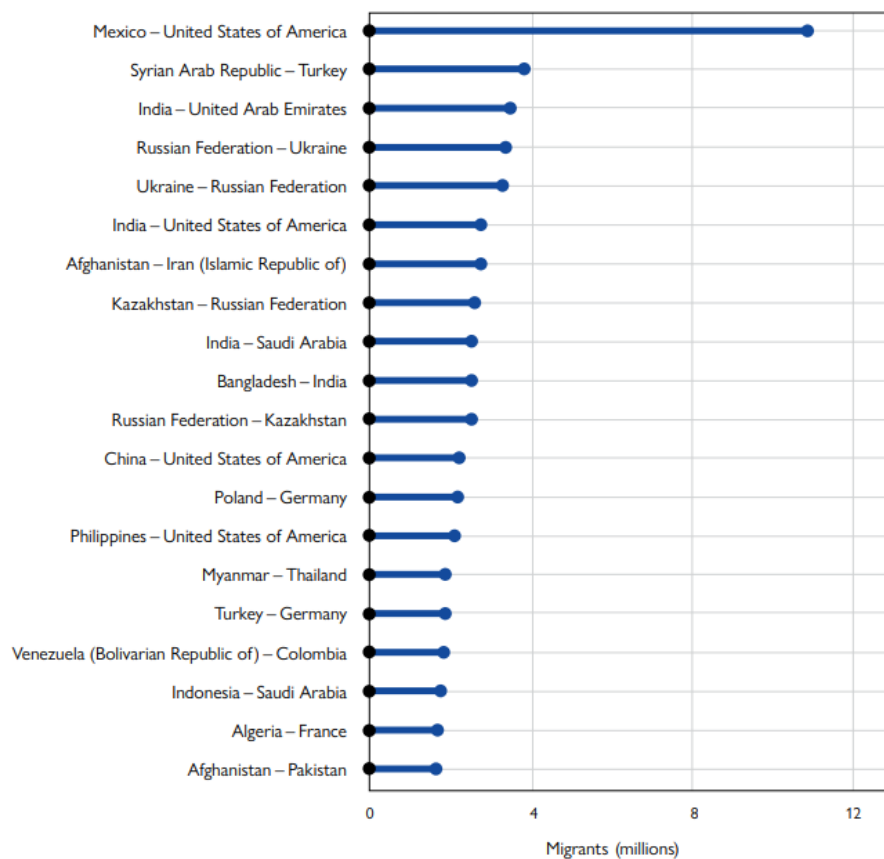

Figure 1. Top 20 international migration country-to-country corridors, 2020. World Migration Report. 2022. International Organization for Migration (IOM).

Through quantitative analysis of international migration data, it becomes evident that migration is not a uniform phenomenon (World Migration Report 2022). Each region exhibits distinct patterns, giving rise to what is commonly referred to as migration "corridors" over time. One significant event that led to the formation of the world's second-largest migration corridor was the migration to Europe in 2015, as highlighted in the World Migration Report 2022. During this time, numerous refugees from countries such as Syria, Afghanistan, Iraq, Eritrea, Somalia, and Nigeria were compelled to leave their homelands due to ongoing internal wars and persecution. They embarked on perilous journeys, seeking to reach Europe by land, primarily through Turkey, or by sea, crossing the Mediterranean (Kingsley 2015; Clayton and Holland 2015).

The magnitude of this crisis captured global attention on multiple fronts. Unprepared nations struggled to accommodate the massive influx of people, leading to significant challenges for the host countries. Border securitization measures were heightened throughout Europe in response to the situation, seeking to regulate the movement of migrants (Hartwig 2020). However, the crisis not only had humanitarian and logistical implications but also had profound effects on the political and public spheres. It sparked intense debates on whether refugees and asylum seekers should be welcomed, fuelling discourses of xenophobia and racism in some contexts (Georgi 2019).

To gain a better understanding and visualization of the most recent migration flows, particularly those relevant to the specific timeframe of this thesis, it is important to examine graphs illustrating migration and refugee trends worldwide. While data on migrant stocks is more readily available, comprehensive knowledge about migration movements, or flows, is considerably limited due to various challenges and difficulties, including measurement tools and data availability. Nevertheless, international organizations such as the United Nations, UNHCR (United Nations High Commissioner for Refugees), the World Bank, the OECD, and regional bodies make efforts to gather and analyse data to provide insights into global migration patterns.

In 2015, the estimated number of international migrants globally was 244 million, accounting for approximately 3.3% of the world's population (United Nations 2018). Internal migration within countries was even more prevalent, with over 740 million people migrating within their own country of birth in 2015. In 2019, the total number of international migrants globally was estimated to be 272 million, accounting for approximately 3.5% of the world's population (World Migration Report 2020). By 2020, the global number of international migrants reached 281 million, representing approximately 3.6% of the world's population (World Migration Report 2022). These statistics highlight the significant presence and impact of international migration, as well as the persistent displacement in various regions of the world.

Below are two graphs illustrating the fluctuations in international immigration over the past five years and the flow of refugees, asylum-seekers, and displaced individuals from 1975 to 2022.

These visual representations provide insights into the changing patterns of global migration and the factors driving forced displacement. The second graph also highlights significant events that have contributed to the displacement of populations, shedding light on the complex dynamics that shape the movement of people across borders.
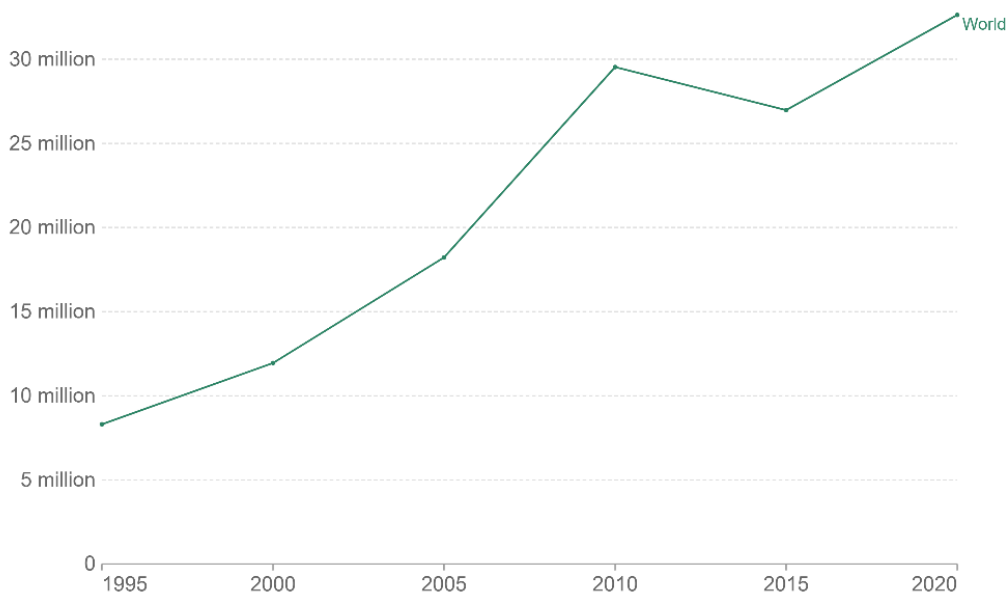

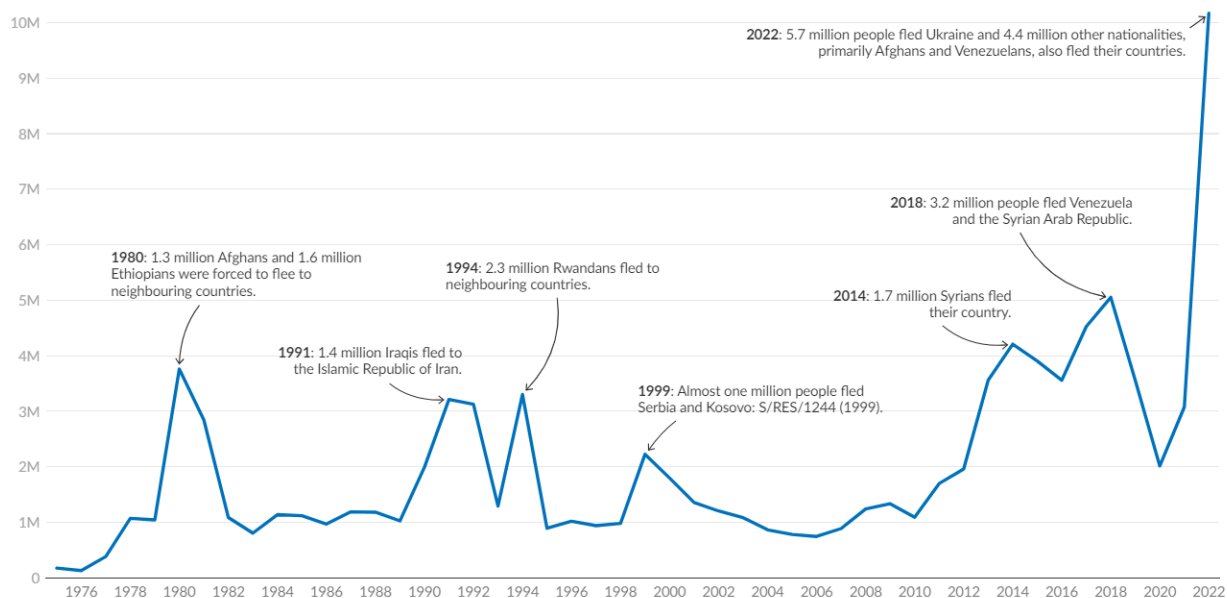Figure 2. Change in the number of international immigrants in the previous five years (Worldwide). Source: https://ourworldindata.org/migration.

Figure 3. Refugees, asylum-seekers, and others in need of international protection were displaced during each year (1975-2022). Source: https://www.unhcr.org/global-trends.

After examining the broader global trends in migration and displacement, attention can now be turned to Latin America, with a particular focus on migration flows to Brazil between 2015 and 2020. By delving into the intricacies of these specific flows, it is possible to gain valuable insights

into the factors that drive migration to Brazil, how the country dealt with the emerging flows during this period and understand the political events that took place at that moment and its importance for this understating public opinion on migration.

1.1. Migration Flows to Brazil and Migration Policies

The period between 2010 and 2020 witnessed significant migration flows to Brazil, both in terms of asylum-seekers and labour migration. The complexities and dynamics of these migratory movements intertwine with political turmoil, economic conditions, violence, and persecution. A multitude of push and pull factors shaped the migration patterns and Brazilian politics regarding migration during this timeframe.

The 2007-2008 international economic crisis[6] led to long-term consequences worldwide and increased south-to-south migration. Between 2010 and 2020, the migratory flows to Brazil reshaped the country's position in the global landscape of contemporary migration (Tonhati, Cavalcanti, and De Oliveira 2022). The phenomenon was not only seen in Latin America, but in 2020 it represented 36% of the global migrant stock (Brettell and Hollifield 2023).

The Brazilian government was positive about the consequences of the economic crisis, after implementing policies to boost domestic consumption and prevent a recession. These measures included reducing interest rates, lowering taxes, and increasing access to credit. However, Brazil's heavy reliance on commodity exports, coupled with declining commodity prices in the following years of the crisis, has posed challenges to the economy (Trevizan 2017). That aligned with the prolonged policies aimed at stimulating consumption led to a deficit in public accounts, affecting the country's credibility and instituting an economic recession during Dilma Rousseff's second government (2015 – 2016).

This first moment of positivism firstly attracted the many people. At that point, the country's economic and in social development in the latest years, its positioning in the geopolitics of the region and the world, the external image of the country as an emerging power, and the organisation of the Olympics and the World Cup (Tonhati, Cavalcanti, and De Oliveira 2022), put Brazil in a spotlight. In the period between 2010 to 2019, there were a total of 1.085.673 migrants registered in Brazil (Cavalcanti, Oliveira, and Macedo 2020).

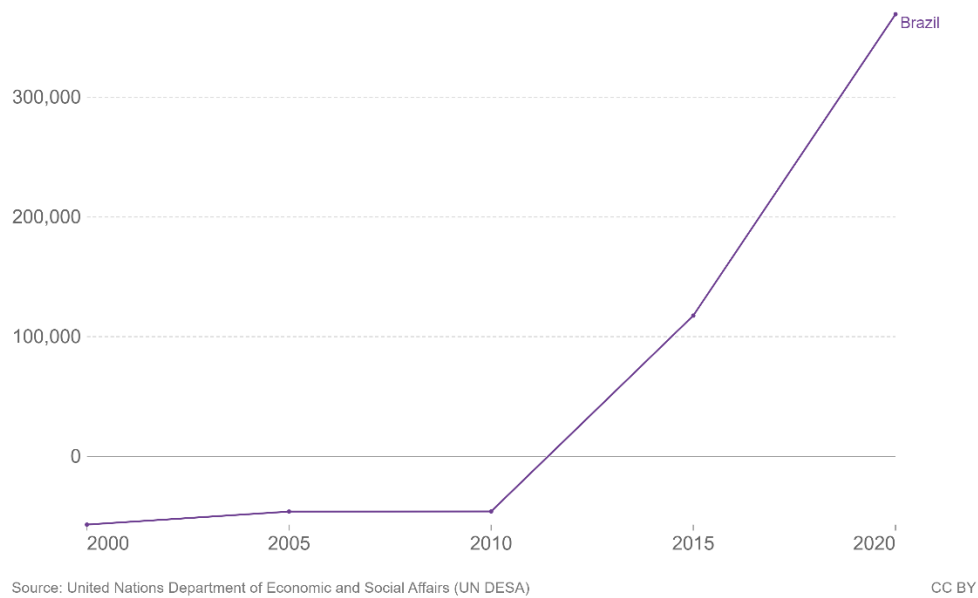[6] For more information about the economic crisis 2007-2008 see: https://www.britannica.com/money/topic/financial-crisis-of-2007-2008

Figure 4. Change in the number of international immigrants in the previous five years (Brazil). Source: https://ourworldindata.org/.

The decade from 2010 to 2020 was marked by an increase in migrants originating from the Global South., such as Haitians, Venezuelans, Bolivians, Senegalese, Congolese, Angolans, Cubans, Bengalis, Syrians, and Pakistanis (Tonhati, Cavalcanti, and De Oliveira 2022). The economic reasons were not the only ones that prompted the movement of people. There was an ongoing series of external events that took place in this decade, making people forcibly move from their home country and find a new home in Brazil, such as the aforementioned conflicts. Two of the biggest flows to the country during the period came from Haiti and Venezuela, due to climate, economic, social, and political reasons.

Immigration to Brazil throughout this decade can be delineated into two distinct phases, according to Simões et al. (2020). The initial phase, spanning from 2010 to 2014, witnessed a substantial upsurge in immigration, propelled by a flourishing economy and burgeoning employment opportunities. In stark contrast, the subsequent phase, from 2015 to 2019, show some fluctuating migration patterns. This oscillation can be attributed to the economic turbulence of 2015-2016 and the influx of new migration currents, particularly from Central America, the Caribbean, and South America, which surged notably between 2017 and 2019.

The 2010 earthquake that struck Haiti, resulting in the reported death of approximately 220.000 people and leaving 1.5 million people homeless (United Nations 2022), led to a significant population movement beyond the country's borders. As noted by Handerson (2015) Brazil was not the first choice of destination for Haitians. However, between 2012 and 2016, the Brazilian Federal Police registered approximately 73,000 Haitians in the country (Milesi 2016). Baeninger and Peres (2017) discuss that the Brazilian leading position on the MINUSTAH (The United Nations

29

Stabilization Mission in Haiti)[7] operations and the presence of military forces in the country, might have prompted the decision to choose Brazil as a place of destination. They also highlight the fact that in a political decision instead of putting Haitians under refugee status, the Brazilian government decided to give them humanitarian visas based on the difficulties the country faced (Baeninger and Peres 2017). The measure was extended for seven years, promoting regular migration and support to immigrants through various programs (Niño and Niño 2018). The humanitarian visa was also later extended to Syrians and Venezuelans (Uebel 2018).

The political and economic crisis that began in 2013 in the Bolivarian Republic of Venezuela has become one of the biggest humanitarian crises in the world, with an estimated displaced population of more than 4,5 million people in 2019 (Martino and Moreira 2020). Before the Russian invasion of Ukraine in February 2022, Venezuela held the unfortunate position of being the largest source of displaced individuals worldwide (Brumat and Geddes 2023). Colombia and Peru hosted the largest number of Venezuelans (World Migration Report 2020), but Brazil also became a place of destination because of its infrastructural facilities, economic development (Galvão, Bruna de Paula Miranda Pereira, and da Costa 2021), and progressive policies to legalise Venezuelan refugees in the country. In recent years more than 800,000 Venezuelans went to Brazil to seek refuge (United Nations 2022). The death of Hugo Chávez in 2013 and the election and re-election of Nicolás Maduro, less supported by various sectors in the government, the fluctuations in the price of oil barrels (Galvão, Bruna de Paula Miranda Pereira, and da Costa 2021) are reasons why the crisis got worst. It has compromised various sectors in the country, such as healthcare and food distribution, mobilised the countries around to deal with the outflow of people from the country (Niño and Niño 2018), and put millions of people in situations of hunger, family separations, structural, physical, and psychological violence.

---

[7] The MINUSTAH operation held in Haiti was established in 2004 and lasted until 2017 after President Bertrand Aristide departed from Haiti in exile due to an armed conflict around several cities in the country. Brazil had an active part in the operations and politics around it, sending around 37,000 military people, holding a leader as the official general of the mission, and being present as a non-permanent member of the Security Council in 2004-2005 and 2010-2011. (For more information refer to: https://peacekeeping.un.org/en/mission/minustah).
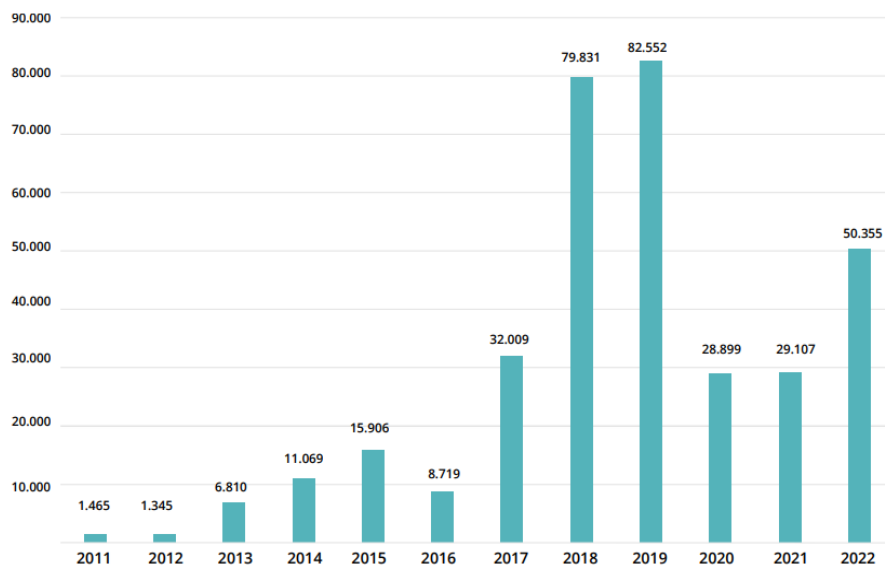
Figure 5. Number of people requesting asylum in Brazil (per year of request) – 2011 to 2022. Source: Observatório das Migrações Internacionais - OBMigra (2023).

To address the challenges posed by the large-scale displacements mentioned above, Brazil had to take political action and adapt its laws and policies. To understand the changes in laws and the actions of the Brazilian government during the analysed period, it is crucial to have an overview of how the country has developed its legal framework to comprehend migratory flows over the years.

Brazil has long been a country of immigrants (forced and not) since its colonization by the Portuguese empire, and has moulded its treatment of migrants in different ways over the years. According to Jubilut (2006), Brazil ratified and adopted the 1951 United Nations Convention on the Status of Refugees in 1961 and the 1967 Protocol on the Status of Refugees in 1972. This made Brazil the second South American country to ratify the Convention and the fifth in the region to ratify the Protocol. Internally, however, Brazil did not meet the international standards on refugee protection (Jubilut 2006) and made limited advancements in the laws regarding migrants and refugees for some years. During the period of the military dictatorship, from 1964 to 1985, Brazil approved the Law 6.815/1980 (Brasil 1980), known as Foreigner Statute Law, which remained valid until 2017, used to define migrants as a possible threat to national security (Tonhati, Cavalcanti, and De Oliveira 2022). The UNHCR office that held clandestine operations in the country during the dictatorship, as its activities were geographically constrained, received the support of the main three catholic NGOs in the country: *Comissão Justiça e Paz* (Justice and Peace Commission), *Caritas Arquidiocesana de Sao Paulo* and *Caritas Arquidiocesana do Rio de Janeiro* (both local branches of Caritas International) (Jubilut 2006). After the re-democratization and the new constitution in 1988, some changes started to appear in the direction of the defence of the rights of migrants and refugees. In 1997, Law 9.474/1997 was passed, creating a new framework of protection for refugees, a basis for the Brazilian

refugee humanitarian commitment (Jubilut and Godoy 2017), and positioning Brazil as a model for other Latin American countries (Brumat and Geddes 2023). The new law, called Refugee Act, established the arrangements for the implementation of the 1951 United Nations Convention on the Status of Refugees and also incorporated a broader definition of refugee based on the Cartagena Declaration. The Declaration was adopted in 1984 by Brazil and 14 other countries in Latin America and the Caribbean, but it was only internalized in national legislation in 1997 (OAS 1984).

The Refugee Act was created through the involvement of the government and representatives of the UNHCR, focusing on establishing parameters for applying the law regarding refugees and establishing the National Committee for Refugees (CONARE) (Brasil 1997). This organization is responsible for refugee status determination and is composed of one government representative, one representative from civil society (role exercised by Caritas), and one representative from the international community (role exercised by the UNHCR).

After years with no substitution for the Foreigner Statute Law, which saw migrants as threats to national security, and the substantial Haitian immigration, Brazil saw a need to change its migratory policy law (Tonhati, Cavalcanti, and De Oliveira 2022; Cavalcanti, Oliveira, and Macedo 2020; Niño and Niño 2018; Baeninger and Peres 2017). The first project of the New Migration Law was evaluated by civil society in 2014 at the 1st National Conference about Migration and Refuge (COMIGRAR). The Law 13.445 (Brasil 2017) was enacted in 2017 and introduced a new and progressive changes in how Brazil would deal with migrants from now on. Some noteworthy aspects of the law include: (1) the law protects any migrant within Brazilian territory, not only those who reside in Brazil; (2) migrants are not addressed as "aliens" and "foreigners" anymore, making the new approach in consonance with human-rights laws of treating non-nationals as humans not threats to the security; (3) it reaffirms the freedom in the country and compared to the previous law it offers more principles and guarantees to migrants in Brazil (de Freitas Castro 2020).

In conclusion, Brazil's evolution of migration and refugee laws has played a crucial role in shaping the country's response to international migration flows. However, understanding the political panorama is equally essential in comprehending the broader context in which these laws have developed. The next part is focused on gaining insight into the political environment, exploring the interconnections between the shifts of politicians inside a political turmoil and the migration and refugee policies in the country.

2. Political Turmoil and New Migration Policies

Brazil faced significant challenges regarding migration during the period between 2010 and 2020. The aftermath of the 2007-2008 international economic crisis reached Brazil slightly later,

leading to an economic downturn. However, the troubles went beyond economic factors, as the country also grappled with a deep political crisis. From 2013, the political arena in the country was marked by a tumultuous period characterized by polarization and political instability. This volatile political climate had profound implications for various aspects of Brazilian society, including migration and refugee policies.

2.1. Protests, Impeachment, and the Conservative-Turn

Since 2015, Brazil has undergone significant changes in its political landscape and governance. After two terms led by the left-wing Workers' Party (PT) under President Luiz Inacio Lula da Silva (from 2003 to 2006 and 2006 to 2010), the country elected its first female president, Dilma Rousseff (same party), in 2011. During the governments of Lula da Silva and Dilma Rousseff, Brazil focused on social and economic policies aimed at reducing poverty, promoting social inclusion, expanding access to education, fostering gender equality, investing in public infrastructure, and stimulating economic growth (Almeida 2014).

Internationally, Brazil's popularity grew alongside the promotion of good relationships with other countries. Lula left the government in 2010 with 87% public approval, and he was succeeded by Dilma Rousseff, who had previously served as Minister of Energy and Chief of Staff during Lula's presidency (Marinho 2010).

Although the government started its work with good approval and a good parliamentary support base, even after a corruption scandal involving the Worker's Party staff, called *Mensalão,*[8] the economy started a period of retraction in 2011. Reductions in taxes were implemented and the country experienced a downturn in its GDP (Gross Domestic Product) and a decline in household consumption, which had a detrimental impact on the labour market. Employment levels decreased, and there was a notable rise in informal employment (Iasulaitis and Vicari 2021).

Furthermore, Dilma Rousseff's challenges in navigating the dynamics with parliamentarians from her political coalition added to the economic issues (Brum, Lima, and Alcantara 2021). In Brazil's coalition presidentialism, the effectiveness of governments relies on the interplay between the executive and legislative powers. However, Dilma faced difficulties in establishing productive dialogues in a highly fragmented Congress (Nunes and Melo 2017).

---

[8] "*Mensalão*" was a political scandal in Brazil during the mid-2000s involving corruption and bribery allegations within the ruling Workers' Party (PT) and its congressional allies. It was named after alleged monthly payments to politicians in exchange for their support of government initiatives, leading to the conviction of several high-profile politicians. For further information about *Mensalão*: /web/20230714094427/https://www.economist.com/the-economist-explains/2013/11/18/what-is-brazils-mensalao

In addition to that, a wave of protests swept the country in 2013, triggered by the increase in public transportation fares in the city of São Paulo, and later expanded to encompass other issues and the most diverse layers of the highly unequal Brazilian society (Uebel and Ranincheski 2017). The protests on the streets and on social media continued at a slower pace until Brazil hosted the World Cup in 2014. During this time, Operation Car Wash was unleashed, exposing corruption scandals involving various politicians and businessmen linked to Petrobras (The Brazilian Oil Company), and also deep-rooted corruption within Brazil's political system. Operation Car Wash, spearheaded by former judge Sérgio Moro, sent shockwaves throughout Brazil. It led to the issuance of thousands of arrest warrants, targeting individuals ranging from high-ranking officials to private companies, and politicians. The involvement of several politicians from the Worker's Party in the Operation Car Wash scandal further tarnished the party's already weakened image, exacerbated by the previous corruption (*Mensalão*).

Amidst this challenging scenario, Dilma was re-elected by a narrow margin of votes in a highly polarised political and social environment, inflamed by the corruption outbreaks, with protests already calling for her impeachment (Nunes and Melo 2017). It was the beginning of an institutional crisis that would soon develop into a political one (Uebel and Ranincheski 2017). Additionally, the operation had significant economic implications for Petrobras, the state-run oil company. The revelations of widespread corruption, money laundering, and embezzlement shattered public trust in the political establishment and ignited calls for accountability and transparency.

Under mounting pressure, exacerbated by corruption scandals and a financial crisis, the impeachment process against President Dilma Rousseff was launched in December 2015. The allegations primarily focused on an accusation of fiscal mismanagement, an infraction to the Fiscal Responsibility Law. In a vote session full of offences, misogyny, and exaltation of the torturers of dictatorship (Barba and Wentzel 2016), the impeachment of Dilma Rousseff took place and resulted in her definitive removal from office in August 2016. As discussed in the political science literature, the impeachment was the outcome of concerted and coordinated action among political parties, elitist social movements, supportive media, and judicial actors (F. Santos et al. 2016). According to Nunes and Melo (2017), the technical arguments in favour of impeachment were not as significant as the widespread consensus among the majority in Congress, society, media, and finance that it was necessary to remove the president. This sentiment was further fuelled by the collaboration between conservative media outlets and the more conservative and corporate faction within the legal establishment (Souza 2016). The conservative-liberal press played a determining role in manipulating public opinion and influencing politicians who voted against Dilma Rousseff (van Dijk 2017; Goldstein 2016). The impeachment campaign was orchestrated through various institutional and non-

institutional channels. While street protests were just one component of this larger effort, they played a crucial role in intensifying the conflict and shaping its outcome (Tatagiba 2018).

Dilma's vice-president Michel Temer also said to be one of the most interested persons in Dilma's impeachment (Uebel 2018), assumed the presidency right after the impeachment and had to build his government in a very delicate situation. Despite the new government's successes in Congress in the first months, achieved through Temer's coalitions, these accomplishments were insufficient to overcome the broader crisis of legitimacy plaguing the administration (Nunes and Melo 2017). Differently from what happened with Dilma, almost nothing was said by the media about the processes of investigations of Michel Temer in Operation Car Wash (Rodrigues and Rodrigues 2018). Some recordings suggested his involvement in bribery schemes within Congress, leading to his warrant in the Supreme Court on charges of passive corruption. However, for the investigation to proceed, the accusation needed approval from the Chamber of Deputies, which did not happen as Temer maintained a supportive network and the case was subsequently archived (Rodrigues and Rodrigues 2018).

Compounding the tumultuous situation, in March 2016, former President Lula da Silva, a still influential figure in Brazilian politics, was formally charged under Operation Car Wash for alleged fraud and money laundering. These charges included allegations of concealing his ownership of a luxurious seaside apartment, which was purportedly acquired through his connections with the Odebrecht construction firm. Lula vehemently denied these corruption allegations, fuelling the already highly charged political climate in Brazil (BBC News 2016). Despite denying ownership of the apartment, Rousseff's appointment of Lula as Chief of Staff was blocked by the federal judge Sérgio Moro. As chief of staff, Lula would have enjoyed legal immunity from federal prosecution and could only be tried in the Supreme Court. The process continued and by April 2017, Lula was arrested and remained imprisoned until November 2019. Reiterating his innocence, Lula vehemently argued that the charges against him were politically motivated, aimed at preventing his candidacy for the presidency in 2018 (Al Jazeera 2019). He contended that it was a deliberate plan to hinder his political ambitions.

One of the reasons for his desire to assume the presidency again was the increased popularity of the far-right deputy and candidate in the 2018 elections, Jair Bolsonaro. Bolsonaro served in the Brazilian Army and later turned his career to politics. The political landscape in Brazil between 2015 and 2020 witnessed a series of significant developments, shaped by a combination of popular protests, the resurgence of the far-right led by conservative and religious movements, strategic utilization of social media, and the spread of fake news and disinformation (M. Botelho et al. 2022). These factors contributed to a deepening political and institutional crisis, culminating in the popularization of Bolsonaro with his inflamed right-wing populist speeches, positioning himself as an "anti-system"

and "anti-political" candidate (M. Botelho et al. 2022). By capitalizing on the prevailing polarisation in the nation, Jair Bolsonaro strategically used social media engagement to dissociate himself from corruption accusations, foster an image of being a superior and better choice for Brazil, and perpetuate the negative sentiments held by a significant portion of the population towards the Worker's Party. This approach ensured his victory in the 2018 election, and he subsequently took office as the president in 2019. It is argued that Bolsonaro's rise to power in Brazil aligns with a broader global trend where far-right movements gain prominence by appealing to traditional moral values and ideals (Iasulaitis and Vicari 2021).

As matter of information, the Supreme Court nullified all charges against Lula in 2021 (BBC News 2021), based on the allegation of biased decision-making by Judge Sergio Moro (Reuters 2021). Lula ran for president once again in the 2022 elections and narrowly won in a tight race against Jair Bolsonaro. In 2022, Dilma Rousseff's accusations were dismissed by the Federal Regional Court (Carta Capital 2022) due to a lack of evidence, confirming allegations of a massive conspiracy to orchestrate the impeachment, which can be understood as a coup (Souza 2016).

Understanding the political landscape and events that unfolded during this period is crucial for comprehending the subsequent changes in migration policies. The political environment influenced the direction and approach of these policies, shaping the government's stance on migration, refugee status, and asylum seekers. In the following section, we will delve into the specific connections between political developments and the evolution of migration policies, shedding light on their interplay and implications for migrants and host communities alike.

2.2. Politics and Policies and Their Implications Concerning Migration

Brazil has been the stage for a series of events that have exacerbated the country's economic, social, and political situation, leading to a transformation in public and social opinions in Brazilian society. Amidst all of this, the country has also become a destination for thousands of immigrants and refugees who, in search of stability, have encountered Brazil in a state of chaos. It is therefore essential to understand how changes in presidents, their positions, and the polarization have influenced the ruptures and continuities of Brazil's migration policies.

A country's migration policy is directly linked to its foreign policy, as understanding and dealing with international migrations is essential to ensure participation in the international arena (Reis 2011). Brazil, therefore, follows this same logic, and as argued by Uebel (2018), in the case of the country, migration policy is conditioned, dependent, and inseparable from foreign policy. For many years, Brazilian foreign policy followed a more moderate approach, focusing on issues related

to Brazilians outside of Brazil (Uebel 2018). Even with consecutive amnesty to migrants without a regular legal status in 1981, 1989, and 1998 (A. P. M. Amaral and Costa 2017).

However, there was a drastic change in the Lula's government, characterized by a resurgence of international assertiveness and humanitarian aid (Uebel and Ranincheski 2017). As observed by Uebel (2018), during Lula's first eight years in government, it was possible to witness several significant developments: the strengthening of Mercosur, the creation of CELAC (Community of Latin American and Caribbean States), UNASUR (Union of South American Nations), BRICS (Brazil, Russia, India, China, and South Africa), IBAS (India, Brazil, and South Africa), Brazil's leadership in the WTO (World Trade Organization), Food and Agriculture Organization (FAO), and ILO (International Labour Organization), the opening of embassies in all Latin American and Caribbean countries and over 90% of African countries, the Tehran agreements, engagement in the Doha Round, and the most important political achievement securing a permanent seat on the UN Security Council. Regarding the treatment given to the migration issue by the State, it was connected to a broader objective on the international stage, particularly during the Lula's administration (Reis 2011). Externally, as part of Brazil's integration into the international stage and the construction of a positive image, Lula's government positioned Brazil as an "El Dorado" and a promoter of peace. Brazil's leadership in MINUSTAH, as well as its participation in United Nations-led humanitarian actions in some African countries and the community of Portuguese-speaking countries, allowed Brazil to establish political and economic relationships, and present itself as a country of opportunities for migration (Uebel 2018). In this context, a policy of attracting immigrants was established, which was intrinsically linked to the country's foreign policy at that time and the maintenance of Brazil's soft power in the international arena through peace-building mechanisms. As argued by Vigevani and Cepaluni (2007) the case of Haiti serves as a prime example of the concept of autonomy through diversification, where diversification entails not only expanding relationships with non-traditional partner states but also having the ability to intervene in areas that may not be immediately beneficial and are related to globally acknowledged public goods.

Concerning Latin America, the relations between the Brazilian government and neighbouring countries were strengthened in pursuit of greater integration among nations, also based on Brazil's strategic projection. It is important to highlight that, considering the arrival of Haitians in Brazil due to the earthquake and the political situation in their country, Lula's government granted a humanitarian visa to the population based on the Refugee Act of 1997, as mentioned before. This decision was later extended to Syrian and Venezuelan migrants, which contributed to an increase in immigration to Brazil (Fernandes and Faria, 2017). This decision regarding humanitarian visas would later become the foundation of the New Migration Law.

Following Uebel's analysis (2018), it is possible to observe that despite Lula's election with great proposals in favour of advancing the human rights agenda and social movements, and despite Brazil being a multicultural country, paradoxically, there was still a significant challenge in integrating these immigrants into Brazilian society and the job market. It is possible to understand that external efforts did not progress in parallel with the country's public policies to ensure the inclusion of these individuals in society. Brazil still operated under an outdated migration law that was developed during the military dictatorship (the Foreigner's Statute Law) and aimed at facilitating the expulsion of foreigners involved in political activities in the country.

Dilma's government began with the perspective of continuing the work of the Lula government, but there were some ruptures related to changes in leadership in the Ministry of Foreign Affairs and the inconsistency of ministers in their positions (Uebel 2018). Later on, there was a major rupture due to political instability and the impeachment process. Initially, the government maintained a policy of incentives (as seen in the continuation of humanitarian visas) and promoted Brazil abroad. The economy was still showing growth, and as mentioned before, Brazil had a positive image when it offered to host the Olympics, the World Cup, and the Confederations Cup. The government facilitated visa acquisition for people coming to the country to attend these events through the Law 12.663/2012 (Brasil 2012). Another example of attraction was the *Mais Médicos* Program, the first government-subsidized immigration process in Brazil in five decades. Its objective was to provide healthcare professionals to regions with the highest unmet demand in the country, leading to the attraction of over 15.000 migrants between 2013 and 2016 (Uebel 2018). All the incentive measures taken by both governments led to a subsequent attraction of other migrants, as well-known in theories of international migration. Among many factors, including economic, labour-related, social, and external factors, the creation of networks among migrants creates various opportunities for new migrations (see Massey et al. 1993).

However, the influx of Venezuelans into Brazil coincides with a backdrop of economic, political, and institutional turmoil in the country (Niño and Niño 2018). This crisis disrupted the progressive migration agenda of Dilma's government, as the focus shifted towards trying to maintain stability within a politically, economically, and socially chaotic scenario and ensuring the president's re-election (Uebel 2018). In this context, the 1st COMIGRAR, a conference intended to discuss the integration of the new migration policy into the government's agenda, was organized by Dilma Rousseff's government in 2014. However, instead of fostering a positive and proactive agenda in support of migration, it was immediately interrupted. The 2nd COMIGRAR, which was supposed to take place the following year (2015), did not happen as planned. As tensions grew, the environment became increasingly hostile and discouraged the involvement of immigrants in political protests against Dilma Rousseff's impeachment. The National Federation of Federal Police Officers

(FENAPEF) released a statement in April of 2016, highlighting the legal prohibition on foreign nationals from participating in political demonstrations in Brazil (A. P. M. Amaral and Costa 2017). In addition to the political and economic turmoil, a labour crisis unfolded, marked by a decline in job prospects, a surge in unemployment rates, and the termination of employment for recently arrived immigrants (Uebel and Ranincheski 2017).

Following the controversial impeachment of President Dilma Rousseff, the interim government led by Michel Temer introduced significant changes in various government policies, including social and foreign policies, which impacted Brazilian immigrants and contributed to a decrease in the intensity of population flows entering the country (Niño and Niño 2018). The government under Temer also blamed the Venezuelan government for the massive emigration flows from that country (Uebel and Ranincheski 2017), and according to the National Council for Human Rights (CNDH), there was an excessive militarization in the humanitarian response to the influx of Venezuelans (Niño and Niño 2018). Although the new Migration Law was approved by the Federal Senate in April 2017, it was significantly altered from its original proposal (Uebel and Ranincheski 2017), with the president vetoing 20 provisions (Niño and Niño 2018). One of the most protested vetoes was the article which would have granted residence permits to migrants who entered the country before July 6th 2016, regardless of their previous migratory situation (Tonhati, Cavalcanti, and De Oliveira 2022). According to Oliveira (2018), the vetoes were influenced by conservative groups who had gained significant influence in Brazilian society and held positions within the government. Consequently, certain aspects of national security and immediate deportations were retained, which could be seen as remnants of the "Foreigners' Statute" (Tonhati, Cavalcanti, and De Oliveira 2022). In 2018, to ensure humanitarian assistance to Venezuelan refugees and migrants in Roraima, the main entry point from Venezuela to Brazil, the federal government established Operation Welcome, which involved the government, UNHCR, NGOs, and civil society.[9]

The presidency of Jair Bolsonaro in Brazil has brought about significant shifts in the country's migration policies and approaches. Bolsonaro's election in 2018 occurred amidst the deepening socio-economic crisis in Venezuela, leading to a large influx of Venezuelan migrants and refugees into Brazil (Mendes and de Menezes 2019). Bolsonaro, known for his ultra-conservative and nationalist views, aligned himself with leaders like former U.S. President Donald Trump, Italian Minister Matteo Salvini, and Hungarian Prime Minister Viktor Orbán, who advocate for strict anti-immigration measures (Mendes and de Menezes 2019). Even before being elected, Bolsonaro referred

---

[9] Operation Welcome is a Brazilian military and humanitarian operation launched in 2018 to assist Venezuelan refugees and migrants who have fled to Brazil due to economic and political instability in Venezuela. It provides essential services and support to these individuals, particularly in the border state of Roraima, in collaboration with international organizations and NGOs. For more information regarding the 'Operation Welcome': https://www.gov.br/casacivil/pt-br/acolhida/sobre-a-operacao-acolhida-2

to migrants and refugees arriving in Brazil as the "scum of the world" (Azevedo 2015). During his campaign and throughout his presidency, he also expressed negative views towards migrants and refugees. In 2018, in an interview with a Brazilian newspaper, he shared his perspective on the exodus of people from Venezuela under Nicolás Maduro, saying that the rich people and middle-class went to the United States, and now just the poor were going to Brazil and that the country had already a lot of issues (Exame 2018).

Bolsonaro called for the repeal of Brazil's new Migration Law, claiming that immigrants' actions and customs were fundamentally different from Brazilian culture and that law would result in an uncontrolled influx of various people into the nation (Brumat and Geddes 2023), proposing the establishment of refugee camps in border regions to accommodate the influx of Venezuelan migrants (Mendes and de Menezes 2019). Bolsonaro's approach reflected a shift towards restrictive measures, aligning Brazil with a more nationalist and exclusionary approach to migration.

In terms of Brazil's foreign relations, the government of Jair Bolsonaro made significant policy shifts. On his first day in office, he announced Brazil's withdrawal from the Global Compact for Safe, Orderly, and Regular Migration, a global agreement aimed at promoting international cooperation on migration issues (Mendes and de Menezes 2019). This decision indicated a departure from Brazil's previous commitment to international migration governance.

Concerns regarding inclusivity and human rights have been highlighted by the government's attitude to immigration. Although Brazil first adopted measures to encourage and accommodate Venezuelan migrants, later steps showed a more restricted posture. The Decision n.666 of July 2019 (Brasil 2019), issued by the Ministry of Justice and Public Security, allowed for the summary deportation of migrants suspected of being involved in criminal activity. The authorities also enacted measures to tighten immigration controls and adopted rhetoric labelling migrants as dangerous.

In January 2020, Jair Bolsonaro delivered a speech expressing support for protecting Venezuelan migrants who had fled their country due to the dire circumstances caused by President Nicolas Maduro's regime (Martino and Moreira 2020). The notion that Venezuela disregards its citizens' fundamental rights and lacks democratic ideals, however, is conveyed by designating these migrants as refugees. To influence Brazil's position toward its neighbouring country in the international community, labelling Venezuelans as refugees was used as part of a political strategy (Martino and Moreira 2020). Despite this rhetoric, when considering Bolsonaro's track record since taking office in 2018, it is highly unlikely to argue that he has transformed into a progressive advocate for the rights of migrants and refugees (Brumat and Geddes 2023). Pereira (2020) asserts that Bolsonaro's government commitment was not to immigrants and refugees, but exclusively to its anti-communist battle.

It is important to note that the COVID-19 pandemic, which began in early 2020, further complicated the migration situation, resulting in border closures and heightened vulnerability for migrants and refugees along the Brazil-Venezuela border (Galvão, Bruna de Paula Miranda Pereira, and da Costa 2021). The government's response to the pandemic and its border policies were influenced by its ideological position, which in turn led to a lack of coordinated actions to effectively address the overall impacts of the pandemic in the country, but also about migrants and refugees.

3.   The Impact of Social Media on Politics and Migration

The rise of social media has had a profound impact on various aspects of society, including politics, migration, and the global refugee crisis. With its widespread accessibility and instantaneous nature, social media platforms have become powerful tools for shaping public opinion, mobilizing communities, and disseminating information.

Online interactions, their content, reach, and influence, have been widely acknowledged as powerful tools for shaping public opinion (de Mello 2019). Recuero (2015) suggests that social media platforms serve as metaphors for the organization of social groups, reflecting the relationships found in offline spaces, being therefore, an extension of the "real-world" public sphere, facilitating the production and circulation of public debates and opinions. Digital platforms have enabled the creation of new channels for distributing and producing information, thereby occupying a space that was once exclusively held by traditional media for many years (Reenlsober and Karhawi 2022). This dynamic embodies both pluralistic and democratic elements, while also posing challenges to accountability and can even be a potential threat to democracy (de Mello 2019). Social media are organized by algorithms that determine what we should see and how many people it will achieve, resulting in an environment filled only with subjects people agree with or are familiar with, and this cannot be overlooked. Bauman (2016) argues that people use the internet conveniently to isolate themselves, creating a comfort zone where only like-minded individuals are admitted and the opinions of those on the "opposite side", are just one click from banishment and oblivion. Although segmentation has its positive aspects, algorithmic logic intensifies this personalization and offers a filtered worldview with personal, cultural, and social consequences that arise when a mass of people starts living a filtered existence (Pariser 2012).

If the algorithm of social media platforms is programmed to offer users any content capable of attracting them and people have also this tendency of isolation in their comfort zones, it can force them to sustain any position, whether reasonable or absurd, realistic or intergalactic, as long as it intercepts their aspirations and fears (Reenlsober and Karhawi 2022). In this sense, Bauman (2016) explores the idea that, in the absence of tangible evidence, people often rely on numbers as a form of

authority to validate their beliefs and assure themselves that they are on the right path of thinking. This dynamic has been widely used as an opportunity inside politics in recent years.

Around the world, the role of social media in the political sphere has been a subject of intense public debate (Boulianne 2015). Digital platforms have been used as tools for division and polarization, particularly by leaders seeking to exploit the issue of migration and downplay its benefits (World Migration Report 2020). The phenomenon of "Twiplomacy," referring to the online interactions of world leaders and quasi-official foreign policy statements on Twitter, has garnered considerable attention in academic research, particularly in the United States with President Donald Trump's communication habits and preferences (de Mello 2019). As a consequence, Twitter turned to be extensively utilized as a data source in the realm of political behaviour, as it is a real-time indicator of political sentiment (Reenlsober and Karhawi 2022). The presidential elections in the U.S. and Brazil, held in 2016 and 2018 respectively, serve as examples of how social networks have diminished the role of traditional media in facilitating meaningful communication between candidates and voters (de Mello 2019). A study by Baptista, Hauber, and Orlandini (2022), analysed the Twitter activities of Donald Trump during the 2016 campaign and Jair Bolsonaro during the 2018 campaign. The findings revealed that both candidates focused less on government proposals and more on direct communication, attacks, nationalism, and depoliticization. They employed personalism to engage with their supporters and positioned themselves as truth-bearers against supposedly deceptive media. According to the research, nationalist values emphasizing freedom, independence, and national sovereignty were also prominent in their rhetoric. Xenophobia emerges as an effective tool for garnering votes by capitalizing on the emotions and frustrations of individuals affected by the crisis. It attributes the presence of foreigners as the root cause and proposes the closure of borders and the exclusion of immigrants as a logical solution, employing the mechanisms of criminal law (A. P. M. Amaral and Costa 2017).

According to de Mello (2019), out of the 140 million active social media users in Brazil, which makes up 66% of the country's population, approximately 43% are Twitter users. In 2017, a report indicated that around 40% of Brazilians spent an average of five hours per day on the internet, and out of those, 72% relied on social media platforms as their primary source of news content (Amaral and Costa 2017). The global users' age distribution data says that 58.2% of people in the platform have between 25-49 years old.

Jair Bolsonaro's campaign's communication strategy centred around the use of social media, particularly Twitter, as a powerful tool to amplify the candidate's values and positions that were not receiving sufficient attention in traditional media channels (Iasulaitis and Vicari 2021). His prominent tweets exhibited a negative tone towards societal issues, and his high number of retweets and likes

indicated that he had already established himself as a major influencer (Reenlsober and Karhawi 2022).

In the context explored earlier, social media has played a significant role in shaping political changes, impacting the country's political landscape, influencing the way public officials operate, as well as shaping people's behaviours. Worldwide the "Arab Spring" serves as a notable example of a transformative development, where voices were amplified, and activists utilized digital platforms to organize and mobilize in political spheres (AlSayyad and Guvenc 2015). In Brazil, the popular protests that started in 2013 sparked a reconfiguration of the far right, led by conservative and religious movements. This reconfiguration involved the strategic use of social media, the spread of fake news, and the dissemination of misinformation, often taking place during popular demonstrations (M. Botelho et al. 2022). Twitter served as a platform for mobilizing users from diverse locations through the use of hashtags, facilitating the dissemination of information from the streets to a wider social media audience, and frequently creating a feedback loop that brought the conversation back to the physical protests (de Mello 2019). Recent studies have conducted comprehensive research on the involvement of social media as a significant factor in shaping public opinion and driving political mobilization in Brazil (see Spyer 2017).

Concerning migration and refugee issues, social media has played a significant role in influencing public discourse, highlighting the experiences of migrants and refugees, and facilitating both support and opposition towards the movement of people. The nature of the migration debate has intensified over the years, driven by fear and division in public and political discourse. Inuwa-Dutse, Liptrott, and Korkontzelos (2020, and Yantseva (2020) show with their findings examining public opinions in social media about migrants and refugees that there is a higher proportion of negative sentiments towards migrants and refugees.

Disruption and disinformation are employed as tactical tools to gain power, impacting public, political, and social media discourse on migration and shaping societal values. The positive contributions of migrants, such as their dynamism and contributions to innovation and entrepreneurship, are then overlooked or ignored (*World Migration Report* 2022). The framing of refugees on social media, as observed in a study analysing Twitter during the refugee crisis, revealed dominant frames related to security, humanitarianism, and at times, explicitly racist hashtags (Siapera et al. 2018).

As shown in the 2022 World Migration Report, the COVID-19 pandemic has further exacerbated the spread of misinformation and disinformation. Coordinated disinformation campaigns to manipulate public opinion, and contributed to the harm caused by hate speech, xenophobia, and anti-migrant policy approaches. The absence of reliable information and the emotional manipulation of fear further contribute to the spread of disinformation. Governments worldwide implemented

various measures and restrictions in response to the pandemic, further impacting migration and its discourse (*World Migration Report* 2022). Bauman (2016) argues that the depiction of migration as a "crisis" creates a state of "moral panic" as news from this battleground leaves a significant impact, embodying complex, and controversial issues, where the imperative of morality confronts the fear of the unfamiliar presented by the masses of strangers at our doorstep (Bauman 2016). As pointed out by Haas, Castles, and Miller (2022), the depiction of the recent mass movements as a crisis might be equivocated as their analysis demonstrate that we might be in an age of migration.

Given the strategic resemblance between Bolsonaro and Donald Trump, as mentioned above, and the well-known positionality of both politicians regarding migrants and refugees, along with the experiences in Europe regarding the sentiment of people towards migrants and refugees, it would not be incorrect to hypothesize that a similar sentiment may have occurred in Brazilian society.

In conclusion, this chapter has provided valuable insights into the main migration flows worldwide from 2010 to 2020, with a particular focus on the migratory flows to Brazil. Additionally, the chapter has examined the political events that unfolded in Brazil during that time and how they influenced the country's approach to migration. The significance of conducting migrant-centric research and analysis has been emphasized, as it allows us to better comprehend the impact of social media activists on unbalanced political discourse, as highlighted in the World Migration Report of 2020.

Looking ahead, the next chapter will provide a comprehensive exploration of Twitter as a research tool. It will delve into its features, opportunities for academia, and the implications of utilizing Twitter data for research purposes. By gaining a deeper understanding of how Twitter operates and harnessing its potential for studying migration, we will have a better comprehension of the results when analysing the data.

Chapter 2: Twitter, a Social Platform, and Data Resource

1.   Twitter, a Revolutionary Platform

In the ever-expanding landscape of social media, few platforms have made as significant an impact as Twitter. Twitter has revolutionized the way people communicate, share information, and engage with each other on the most diverse topics in the world. Launched in 2006, it started as a platform for updating friends about your personal life, thoughts, and location. It turned out to be a global phenomenon, with more than 368 million users in 2022 (Dixon 2022). This chapter briefly explores the history of Twitter, tracing its journey from a simple microblogging[10] website into a social media powerhouse.

The first part analyses the development of Twitter from an idea of a platform that would enable individuals to express themselves succinctly in 140 characters or less, to a platform that transformed the way we communicate and consume information. To better understand how this occurred, it is pivotal to also delve into the platform's main features: the Tweets, Retweets, and hashtags, their role as building blocks of Twitter and as drivers for user engagement.

Beyond its role as a social networking site, Twitter emerged as a valuable archive of contemporary culture and an environment that provides rich insights into the collective mindset of millions of people. Many scientists, including historians, and social, political, and economic, but not only, have already turned their attention to exploring Twitter as an environment of social interaction that can offer a considerable amount of data and perspectives of the past, present, and future. Twitter's growth and evolution provided new options for people to interact, share ideas, and engage in public discussion. The platform evolved into an effective tool for amplifying voices, mobilizing social movements, and influencing public opinion. Its effect surpassed virtual domains, permeating real-world events, and triggering social and political change. Therefore, the subsequent part of this chapter delves into the opportunities and the importance of studying and researching the platform.

Last, it briefly presents the debate about ethics on using Twitter data and its possible implications. The aim is to bring to light the relevance of Twitter in the contemporary world when analysing events and discourses of current issues in society and also explain its features. Like other social media, Twitter offers many historical and social research opportunities. The power of network

---

[10] the activity of sending regular short messages, photos, or videos over the internet, either to a selected group of people or so that they can be viewed by anyone, as a means of keeping people informed about your activities and thoughts (Oxford Learner's Dictionary).

effects and real-time updates shows the importance of developing research and news tools to analyse this constant production of data, but one must not forget about the debate on the ethical implications of it.

1.1.   Twitter: What Exactly Is It?

The initial concept of using a short message service (SMS), allowing people to share their status, location, and thoughts with a small group of friends, was presented by Jack Dorsey, who was a student at New York University at that time. He pitched this idea to Evan Williams, Biz Stone, and Noah Glass, the founders of Odeo, a podcasting venture that aggregated and facilitated the creation, recording, and sharing of podcasts.

Twitter was created by the founders of Odeo, along with Jack Dorsey, in March 2006. The complete version of Twitter was then launched in July 2006 in the United States. It provided a lightweight service that allowed users to update their friends using a minimalist interface. Users could view all updates in a public timeline on the website and also receive them via SMS within the United States (Burgess and Baym 2020). The platform encouraged users to share their thoughts by prompting them with the question "What's your status?". Initially, the primary language of the platform was US English, but it became widely available in other languages after 2010.

Figure 6. Twitter's first interface was in 2006. Source: WebArchive (https://web.archive.org/web/20061203201128/http://twitter.com/).

As a newly debuted platform, people tried to make sense of it, classifying Twitter as a microblogging platform or social networking (or both), and compared it to other existing media, such as Myspace. The tipping point that changed the platform's positionality and popularity was the annual SXSWi (South by Southwest Interactive) conference in Texas in 2007, where the company placed screens in the conference's hallways and streamed the participants' updates and mentions about what was happening during the event (Terdiman 2007). Since then, Twitter has experienced rapid global growth (Haustein et al. 2016) both in the number of users and also in the company's revenue, having still continuous growth as exemplified in the graphs below.
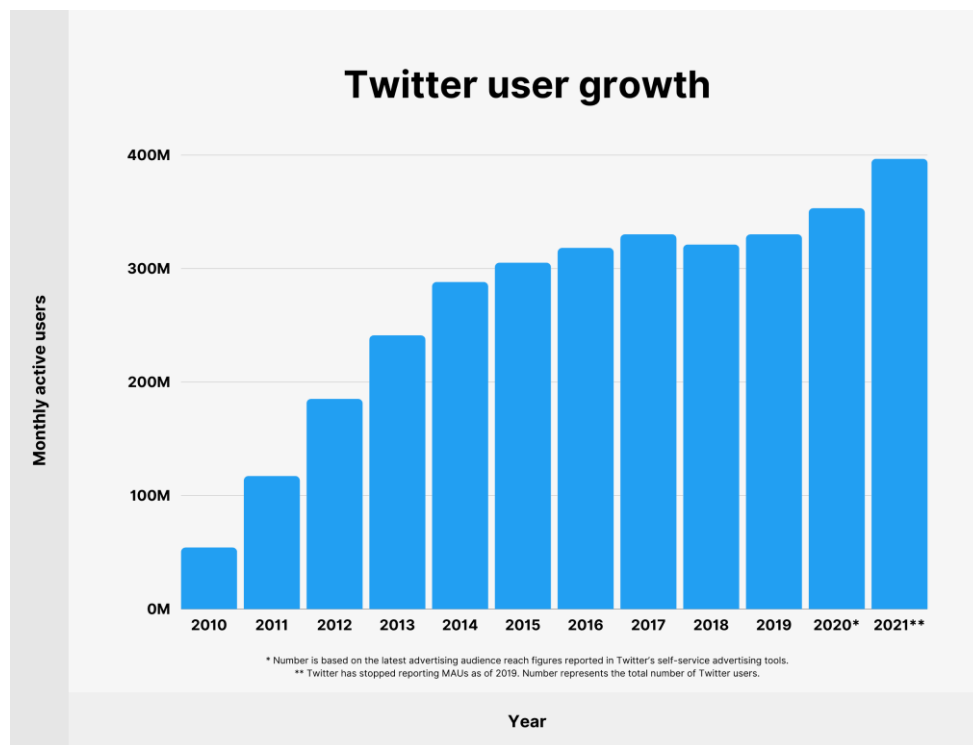
Figure 7. Twitter user growth from 2010 to 2021. Source: https://backlinko.com/twitter-users.

The stable growth of the platform is intricately intertwined with the rapid advancement and proliferation of computers, smartphones, and internet accessibility. Initially launched as a web-based platform, like much of other social media, Twitter had to adapt to the changing technological landscape. Which fortunately for the company, also complied with Twitter's objective of being a platform for instant communication and fast updating.

The way users interacted with the platform, as well as its growing popularity, caused Twitter to evolve in conjunction with advancements in technology, resulting in a harmonious relationship between the two. This iterative process of development and expansion was driven by the evolving needs and expectations of Twitter's users, who sought more interactive features to communicate between themselves.
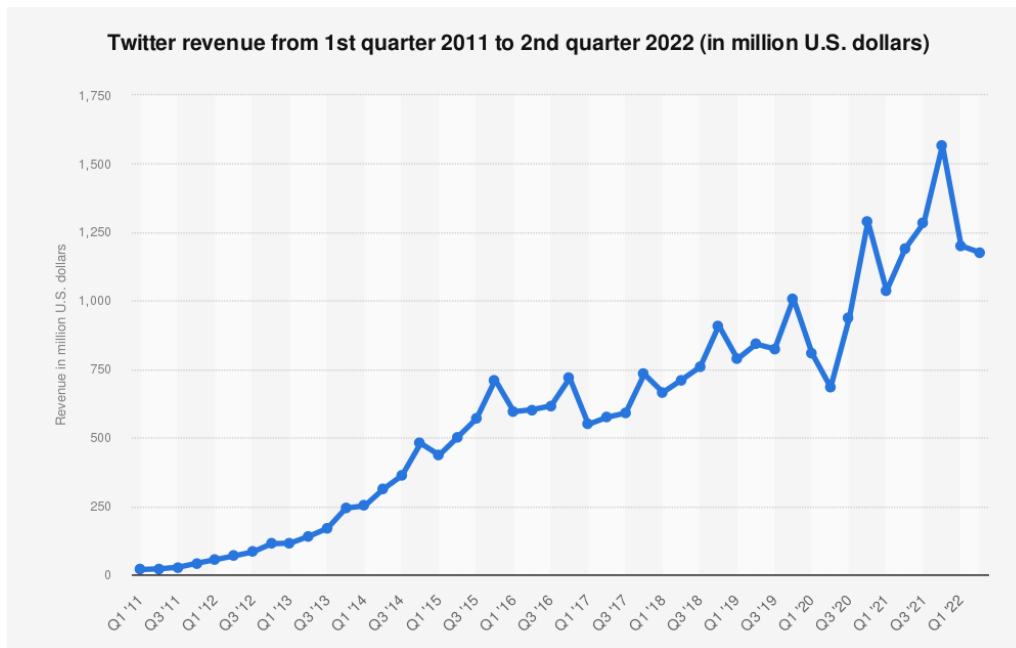
Figure 8. Twitter revenue from 1st quarter 2011 to 2nd quarter 2022 (in million US$). Source: Statista (https://www.statista.com/statistics/274568/quarterly-revenue-of-twitter/).

To meet these demands, Twitter embarked on a journey of innovation, introducing new functionalities that allowed users to make the platform more "conversational." Later on, there were also additional features such as sharing images, videos, and links within their tweets. The integration of multimedia content not only enriched the user experience but also positioned Twitter as a dynamic hub for real-time information and visual storytelling. Additionally, Twitter harnessed technological advancements in data analytics and machine learning algorithms to provide personalized recommendations, trending topics, and targeted advertising, further enhancing user engagement and retention.

The context surrounding Twitter's growth, such as the rise of social media, the increasing connectivity of global populations, and the demand for instant information sharing, acted as catalysts for its technological development. Twitter's ability to adapt and respond to these contextual factors played a crucial role in its evolution, allowing it to stay relevant in an ever-changing digital landscape.

In recent years there have been proposals to acquire Twitter, with several parties indicating interest in doing so, such as the owner of Facebook Mark Zuckerberg (Panzarino 2013) in 2013. However, the one who manage to acquire the company was the well-known entrepreneur Elon Musk, the CEO of organizations like Tesla and SpaceX, in 2022. Musk said that he planned to make new features, make algorithms open-source, combat spambots, and promote free speech (Sato 2022). His positions and actions are controversial whatsoever, and the reactions to them are very polarized. The acquisition was finalized in October 2022 and Twitter Inc. was bought by US$ 44 billion. By April

he has fired more than six thousand employees, which represented 80% of the company's staff (Toh and Liu 2023). Since there, some new features have been added to the platform, such as Twitter Blue[11] (a premium subscription service) and other features related to the interface that also raised some polarized positions from users. Most recently the platform was rebranded as "X" and is constantly under upgrades.

1.2. The Development of Twitter as an Information and News Platform

Another important aspect to mention is the platform's primary focus is information. The aim was to inform your friends and followers about your current status. Soon in 2007, the network was joined by politicians such as Barack Obama and the leading mainstream international media outlets like the BBC, Al Jazeera, and the New York Times (Burgess and Baym 2020). Since then and over the years, many people come to Twitter mainly to stay informed about current news and events. The platform emerged as a hub for significant discussions from local to global news, entertainment to politics, there is an infinite range of topics that people write about every day. It even promoted the use of the platform to this intent with the launching of Twitter for Newsroom in 2011. The newsroom along with new conversational settings enables journalists and news media to host panel discussions about specific topics.

Described as an "awareness system," Twitter plays a pivotal role in identifying emerging trends or issues that have not yet gained significant attention in mainstream news (Hermida 2010), Twitter's involvement and rapid dissemination of news are exemplified through significant global incidents. These occurrences firmly establish Twitter as a platform for cultivating global awareness. For instance, how the developments during the Arab Spring were highly commented on Twitter and the role of the platform in organising and informing people about what was happening (Aday et al. 2013). Another important event was on January 15, 2009, when a commuter ferry passenger Tweeted a picture of the US Airways aircraft 1549 landed on the Hudson River in New York City. This Tweet attracted users' interest and rapidly travelled across the internet, raising awareness of the unfolding catastrophe. In particular, the photograph showing passengers safely disembarking from the aeroplane that had landed in the water was published to Twitpic.com, a third-party photo-hosting site for Twitter users, adding to the immediacy and significance of the news relayed via Twitter (Britannica 2023).

---

[11] For more information about Twitter Blue access Twitter Help Centre (https://help.twitter.com/en/using-twitter/twitter-blue).

During that period, Twitter did not support the inclusion of images within tweets. However, it was not until June 2011 that they introduced an integrated photo-sharing service (Kumar 2011), enabling users to upload and attach photos directly from Twitter.com. This development marked a significant milestone as it allowed users to share visual content alongside their Tweets. Subsequently, Twitter continued to enhance its image-sharing capabilities, enabling users to post multiple pictures and provide accompanying descriptions in subsequent years.

Following a blog post shared by Twitter News (2022), Twitter analysed in collaboration with third-party research groups to explore the habits and preferences of its users regarding news consumption. The analysis drew upon surveys conducted in the United States. The findings indicate that a substantial majority, specifically 83% of Twitter users, actively engage in Tweeting about news topics. Moreover, three out of four individuals who turn to Twitter for news do so daily, highlighting the platform's significance as a regular source of news updates.

Additionally, the research reveals that 55% of people specifically choose Twitter as their preferred platform for accessing news, prioritizing it over other social media platforms. Nevertheless, Twitter also serves as a valuable tool for discovering news outlets, journalists, and individuals whose Tweets are deemed worthy of following. This feature enhances the platform's appeal and utility in curating personalized news experiences.

Importantly, Twitter holds particular importance within the journalism profession. Approximately 83% of young journalists report that Twitter is either their most utilized social media platform or ranks as the second most frequently employed platform in their professional endeavours. This underscores the platform's relevance as a prominent medium for journalistic engagement and dissemination of information.

From a personal perspective but also adding information to the data, as an avid Twitter user since 2009, I have frequently immersed myself in the platform to stay informed about current events. Over the years, my morning routine has consistently involved accessing Twitter to peruse the headlines shared by the newspapers I follow. Notably, the 2022 Brazilian presidential election, spanning two rounds, garnered significant attention and became a focal point of discussions on Twitter. Consequently, I closely monitored real-time updates on the election results, alongside the accompanying commentary from individuals worldwide, using the platform as a primary source of information. From more personal updates to emerging information about politics, the ease and instantaneity of the platform, made Twitter grow as a sharing news media.

Place of information, space of interaction

The inherent public nature of tweets, coupled with the interactive capabilities of the platform, presents journalists with a convenient and valuable space to gather information. The substantial

presence of influential individuals and celebrities who actively utilise the platform to disseminate information, express opinions, promote their brand, and engage with others, makes Twitter even more appealing for journalists. In comparison to other social networks, Twitter has achieved notable success in bridging the gap between ordinary users and prominent figures, including those who wield significant popularity, power, wealth, and influence (Marwick and Boyd 2011).[12]

Lefebvre (1974) put forth the argument that social space is constructed through communication, with artists, writers, and philosophers occupying the realm of "representation of space." The advent of social media has transformed people's perception of space within their social and geographical contexts, facilitated by the proximity and co-location of information (Edwards and Usher 2007). Within social media, the dynamics of social, political, and economic power transcend spatial and temporal boundaries, prompting a re-evaluation of the various scales of place and mobility in society (Edwards and Usher 2007).

Without any limitation to following each other or any requirement of mutual acceptance, Twitter eliminates the barriers to connecting users on the platform and contributes to the ease of engagement. The number of followers can even give one some type of legitimacy and put this person in a better "position" in comparison to the ones who have fewer followers. The contemporary spaces of communication and influence go far beyond philosophers and writers, it is also today deeply occupied by celebrities, entrepreneurs, digital influencers, bloggers, etc. (Lasorsa, Lewis, and Holton 2012). Within contemporary societies, the exercise of social power relies substantially on social networks (Castells 2011), highlighting the active role and influence of digital platforms in the social fabric of modern society, and how they (and the programmers within the company) have the power to shape the people's interaction.

The digitalization and widespread adoption of social media has led to the transformation of Twitter from a place to a dynamic space (Broersma and Graham 2013). In this context, "place" embodies a multifaceted and subjective experience deeply influenced by local discourses and contextual factors. Conversely, "space" represents the realm of tangible and potential movements involving individuals, entities, information, and communication (Edwards and Usher 2007).

Twitter has evolved from a straightforward information platform to a multifaceted hub that combines both information dissemination and social networking. Users not only passively consume news but actively participate by engaging with the content, expressing their opinions, sharing their interests, and establishing connections that align with their preferences. Several studies have found that the informational aspect of Twitter takes precedence over its communicative functions (Dijck

---

[12] From the data extracted from to research, only 1.23% of the user accounts were verified.

2011; Kwak et al. 2010). Regardless of whether Twitter is perceived as a social network platform or an information-sharing platform, it fulfils significant roles in both aspects.

2.   Twitter's Features

As previously mentioned, the conversational nature of Twitter, along with its aggregated features, was not initially planned by the developers. Unlike face-to-face communication, the platform enables asynchronous and geographically unconstrained dialogues (Boyd, Golder, and Lotan 2010). Users have the freedom to choose whom they follow, be it celebrities, politicians, a large number of people, or only close friends. They can engage in conversations, and promote, and share posts according to their preferences. The timeline is constructed based on individual interests, with tweets displayed in reverse chronological order on the first page.

The presence of Twitter applications on mobile devices allows users to tweet and read tweets from virtually anywhere, indicating the high level of mobility associated with the platform. Consequently, Twitter exhibits a flexible and unpredictable temporal structure. A tweet can be seen within seconds of its posting or discovered days, weeks, or even years later.

The incorporation of additional features into the platform, driven by its increasing usage and popularity, has facilitated the establishment of connections. Twitter has been acknowledged by public relations professionals and scholars as a valuable tool for nurturing relationships (Evans, Twomey, & Talan, 2011), despite some counterarguments highlighting the distinctions between face-to-face and social media interactions. Nevertheless, social media interactions played a significant role in fostering socialization during the unparalleled circumstances, as for the coronavirus pandemic.

1.1.   The Tweet, Reply and Mention (@)

Few terms have become as influential as the tweet. From what is primarily defined as "a short, high-pitched sound made by a small bird" (Collins Dictionary 2023), it transcended the definition of the simple sound of a bird, drawing into people's vocabulary also as a verb and a noun. It now encompasses a vibrant ecosystem of ideas, conversations, and trends.

In its essence, a tweet is a concise message that users share on the Twitter platform. Initially limited to 140 characters and now expanded to 280 characters, Tweets now have the possibility of adding other types of information such as images, gifs, videos, links, and pools. Its impact, however, expands far beyond its character limit. Each tweet can represent a unique range of expressions, momentary thoughts, opinions, and experiences.

Following the company's definition, the mention (@) "is a tweet that contains another person's username anywhere in the body of the Tweet" (Twitter Help Centre n.d.). A person can

directly notify and involve that user in the conversation. When you mention someone in a tweet, they receive a notification, and the tweet appears in their mentions tab. It allows for direct communication and helps connect users across the platform.

Although the symbol traces back to older IRC[13] practices, the @ on Twitter was first applied by users in the relatively early days of the platform's life. Applying the new feature in early 2007, Twitter transformed the central informational aspect of the platform into a more discussion and ongoing interactional framework, enabling people to mention and respond to each other. Its meaning has been regularly tweaked, contested, and reinterpreted by both users and the company in the years afterwards, but exemplifies users' persistent trust in their ability to rearrange to socialize (Burgess and Baym 2020).

A reply on Twitter is distinct from mentioning someone in your tweet, although it is represented by the same symbol (@). It serves as a direct response to another person's tweet and can be added by clicking on the dedicated "reply". Replies hold great significance in shaping conversations on the platform, as they enable users to engage in meaningful discussions and contribute to the dialogue in a relevant, credible, and safe manner. Twitter employs complex algorithms to tailor responses to individual interests, aiming to present content that users find interesting and valuable.

Over time, Twitter introduced the feature of aggregating replies and mentions, allowing users to easily view the interactions they have received. Subsequent redesigns incorporated a 'notifications' section that provides metrics related to mentions, replies, retweets, and likes. This tool enhances seamless interaction among users, enabling them to engage not only with familiar individuals but also with strangers (Burgess and Baym 2020). These features are integral to fostering connections, collaborations, and feedback, contributing to the dynamic nature of the platform.

2.2.   The Hashtag (#)

The hashtag (represented by the symbol "#") is a metadata tag, that helps to describe an item and allows it to be found (Hashtag Org. 2012). On social media, the hashtag is used as a form of user-generated tag that associates cross-referencing of content by topic or theme (Chang and Iyer 2012). Beyond that, hashtags also became a form of organising information for users on many social media, such as Facebook, Instagram, YouTube, and Twitter. Officially adopted in 2009 by Twitter, users use the hashtag to categorize Tweets by keyword, using the symbol # before a word or phrase making

---

[13] Internet Relay Chat (IRC) is a text-based chat system for instant messaging. The system used the @ to refer to a message. Source: (https://history-computer.com/irc-guide/)

them appear more easily in Twitter searches. By clicking on a word with the hashtag, the user sees instantaneously other tweets that include that hashtag.

The first person to post a hashtag on Twitter was an American blogger and product consultant Chris Messina, who tweeted in August 2007 a suggestion to adopt the # (pound) for aggregating messages in groups. From this first suggestion, Messina (2007) wrote an extensive description of his intentions of how hashtags would work. The core idea was that hashtags were used as metadata to provide extra information about a tweet, like what event the person is referring to. The design of the interface ensures that any character preceded by a hash symbol becomes a clickable link (as long as there is no space between the symbol and the character).

Yang (2016) highlights that the rise of hashtag activism is a captivating development in digital activism. It serves as an instrument to expand communication the use of communication and ensure its democratic nature, in the sense that everyone has a means of expressing their ideas (Ta'amneh and Al-Ghazo 2021). Hashtag activism involves using hashtags on social media to unite and organize expressive protests, where a word, phrase, or sentence becomes a collective voice. It has become a popular and influential approach that has made a significant impact on political and social developments globally. According to Goswami (2018), the hashtag movement has proven to be a catalyst in socio-political events, acting as a knowledge tag to advocate for causes and connect social media users across platforms through concise and impactful messages.

In 2007, the use of hashtags gained prominence during the San Diego County fire, where the hashtag #sandiegofire enabled users to easily track updates about the incident on Twitter. Subsequently, in 2009, hashtags became integral to the protests following Iran's presidential election, often referred to as the "twitter revolution." Twitter itself embraced hashtags in 2009 and from that on, many other events have gained worldly awareness over the years (Panko 2017), making hashtag activism is to be considered one of the most popular strategies with a substantial impact on political and social changes around the world (Ta'amneh and Al-Ghazo 2021).

Originally designed to facilitate content search, filter out noise, and group topics, Scott (2015) proposes that users can also use hashtags as a means of emphasizing a specific topic or theme in a tweet, regardless of its potential to aggregate any other knowledge. She suggests that users have transformed the symbol into a highlighting tool (Scott 2015), expanding its original and primary function of making tweet content searchable. It has also been noted that hashtags have become a prevalent and widely adopted phenomenon for building public opinion and creating awareness (Müller-Prove 2008). When a significant number of users post tweets with the same hashtag within a brief period, that hashtag is considered to be trending. Users can browse and view current trends in real-time on the platform.

According to research, hashtags can also play a valuable role in sentiment analysis. A study conducted by Maynard and Greenwood (2014) revealed that some individuals choose to include a hashtag that reflects their emotions or sentiments towards a statement. As a result, the words used in hashtags can serve as a strong indicator of whether a statement is sarcastic or not, which is typically more challenging to predict using computational tools.

Since its adoption, the hashtag has been widely embraced by users, contributing to the rise of Twitter as a platform for news, information, and professional promotion. Apart from the aforementioned impact on amplifying the interactional aspect of Twitter, it has become deeply integrated into the institutional fabric of mainstream media, companies, brands, events, and political campaigns. The study of hashtags has emerged as a significant subfield in social media research, reflecting the growing importance and impact of these digital markers (Burgess and Baym 2020).

2.3.　The Retweet (RT)

In the realm of Twitter, participants engage in various forms of conversation, whether it be actively commenting or simply acknowledging their presence as listeners. Each act, from spreading a tweet to retweeting, serves to bring people into the conversation (Boyd, Golder, and Lotan 2010). A retweet is essentially a way for users to publicly share a tweet, forwarding news or interesting content. It provides the option to include personal comments and media, known as a "quote tweet," to gather more information. Beyond being a quick method to share a message, it also serves as a way of referencing the author and may even spark additional conversations based on the main message, without directly addressing the author. In this sense, retweeting is not only sharing the message but also validating the content and engaging with others (Boyd, Golder, and Lotan 2010). Users can view retweets in their timelines, and if someone Retweets their own Tweet, it will be visible in the notification tab (Twitter Help Centre n.d.).

The retweet button was introduced in 2009 on the platform by a team of developers led by Chris Wetherell. Before its introduction, users would manually type "RT" and copy-paste the text from the original tweet to share it. In a 2019 interview, Wetherell expressed regret over inventing the feature, as they had no idea how it would significantly impact the way people use the platform. However, since the button's invention, the power to swiftly spread any kind of message has multiplied. This speed and convenience of sharing information has become the main concern for developers, as it often leads to a lack of critical thinking about the shared content. Interestingly, a person's decision to retweet is rarely influenced by factors related to the information source (Shi et al. 2017).

Wetherell's original intention was to provide a means of amplifying the voices of those who lacked representation. However, the retweet feature inadvertently became a vehicle for the dissemination of extreme content, misinformation, fake news, polarized content, and misogynistic attacks. One example is the Gamergate harassment campaign, where retweeting played a significant role in organizing and perpetuating attacks against women in the game industry (BBC News 2019).

Recent research shows that Retweets, which also include quote tweets, constitute roughly 50% of all tweets (McClain et al. 2021). As a result, several studies have been carried out to create conceptual frameworks for comprehending the behavioural factors underlying people's tweeting habits. According to Abdullah, Nishioka, and Murayama's (2016) research, people retweet during disasters to disseminate information, express their feelings, receive feedback, and warn others. Statistics show that during large crises and emergencies, including fires, flooding, and elections, among others, the percentage of retweets dramatically increases (Hughes and Palen 2009).

Another study investigates the variables affecting a person's decision to Retweet (Shi et al. 2017). It points out five key elements: (1) the information source, which includes elements like source reliability; (2) the information richness of the stimuli; (3) the topical inclinations of the information receiver; (4) the connection among the source and the receiver, such as the force of the social tie; and (5) contextual elements, such as the bandwagon effect[14]. As a result of their research, it is shown that the receiver's topical relevance and social tie strength stand out among these elements as the most important ones, followed by the stimuli's information richness and the bandwagon effect.

Besides spreading awareness about some specific topic, comment, update, product or news, the retweet also proved to contribute to the conversational ecology where conversations are formed through a dynamic interplay of public voices, generating an emotional connection, and fostering a shared understanding of the context (Boyd, Golder, and Lotan 2010). The nature of these conversations on Twitter is incredibly diverse, with retweeting catalysing different types of discussions that can range from small and localized exchanges to sprawling threads that capture a massive audience.

As users navigate the constraints of Twitter, their approaches to retweeting unveil their values regarding specific messages and the platform as a conversational environment. With varying social and informational goals, users adopt different retweeting practices that align with their objectives. Regardless of the reasons behind choosing to retweet, users become integral parts of larger conversations that are not orderly exchanges of interactions, but rather multiple spheres of simultaneous conversations (Boyd, Golder, and Lotan 2010).

---

[14] The bandwagon effect is the tendency for people to adopt certain behaviours, styles, or attitudes simply because others are doing so.

Moreover, companies have recognized the value of the retweet button on Twitter and have established stronger connections with the platform as a result. By utilizing this feature, advertised posts gain enhanced visibility when shared or liked by users. This amplification of reach enables companies to effectively promote their products or services by facilitating the easy sharing of specific posts. As a result, these shared posts have the potential to gain widespread attention, increasing their popularity and potential impact (Kantrowitz 2019).

3. Twitter as an Environment for Academic Research

Twitter holds significant importance in the realm of academic research due to its enormous accessibility, user base, real-time nature, and diverse content. Since 2008, Twitter has become mainstream (Zhao and Rosson 2009), and it provides researchers with a unique platform to study and analyse various aspects of human behaviour, social dynamics, and communication patterns.

The readily available source of data, real-time insights, worldwide reach, and a wide range of topics on Twitter make it an attractive space for multiple fields of study. Its accessible data can be examined to discover important things about a broad range of subjects, including historical data. Real-time data collection enables researchers to gain instantaneous insights into public opinion, social movements, and breaking news. The study of numerous communities and cultural phenomena is made possible by Twitter's database, which spans geographical and cultural borders. The aforementioned components help to a thorough understanding of social dynamics, cultural interactions, and human behaviour, offering scholars many opportunities and a big dataset for researchers.

The introduction of the Twitter API back in 2006 has opened up a world of opportunities for academic researchers to delve into a wide range of topics by utilizing data from public conversations available on the platform. In 2020, Twitter released its latest version of the API, known as Twitter API v2, along with the Twitter developer platform, providing researchers with better tools and resources. In that year also, Twitter has made significant updates to its developer policy and built a dedicated webpage specifically for Academic Research (Tornes 2021).

In 2021, Twitter took a step further by launching the Academic Research product track on the new Twitter API, expanding the possibilities for the academic community. To gain access, researchers need to apply for their research projects and obtain approval from Twitter. This product track offered qualified researchers access to all v2 endpoints, released up to that date, granting them free access to the complete history of public conversations through the full archive research endpoint. An endpoint, in the context of API communication, is the point of interaction where two systems interact. It serves as a location or address through which an API accesses the necessary resources from another system. Essentially, it is a doorway or access point that allows users to perform specific

actions or retrieve the desired information. When requesting an API endpoint, users are essentially asking the API to carry out a specific action or provide specific data. Additional information in the form of parameters or data payloads may be required by endpoints to fulfil requests accurately. These parameters refine the results or provide essential details for the requested action (TechTarget n.d.).

Previously, access to these API endpoints was restricted to paid premium and enterprise accounts, but now researchers enjoyed an increased Tweet volume cap of 10 million per month, allowing for a more comprehensive data retrieval. The v2 endpoints also provided enhanced filtering capabilities, enabling researchers to focus on the specific data relevant to their studies and simplifying the data-cleaning process. Twitter further supported researchers with technical and methodological guides.

As a result of these advancements, academic researchers became one of the largest user groups of the Twitter API in 2021 (Tornes 2021), benefiting from the expanded features and resources provided by Twitter for their research purposes. This development has further solidified Twitter as a valuable platform for academic investigation and analysis.

However, in February 2023, Twitter made an announcement stating that it would introduce charges for third-party access to its platform data (Center for an Informed Public - University of Washington 2023). This decision created a significant barrier for academic researchers, as the cost became unaffordable. Additionally, even for those willing to or can pay, the number of retrievable tweets was reduced. This measure received widespread criticism from researchers, as it severely limited important studies across diverse fields. The lack of response from Twitter thus far has left many projects in a compromised state, with no clarity on the future of accessing essential data. The uncertainty surrounding Twitter's business strategy, particularly under the administration of Elon Musk, adds to the concern and anticipation for what lies ahead.

3.1. Research Opportunities on Twitter

With wide-ranging data, research on Twitter can be a fruitful environment for various academic fields. In a study conducted in 2013 entitled "What people study when they study Twitter: Classifying Twitter-related academic papers" by Shirley Ann Williams, Melissa Terras, and Claire Warwick, they aimed to classify and analyse over a thousand research papers published between 2007 and 2011 that focused on Twitter-related topics. Based on their findings, there was an impressive increase in the number of published papers related to Twitter from 2007 to 2011. In their early work they found research concerning the platform as related to the domains of user and message (e.g., the number of tweets; the number of followers and followings; times of posting; location), and also linguistic and semantic analysis of Tweets and conversations (Williams, Terras, and Warwick 2013).

In this study, they provide an extensive review of the literature and a classification of the different scholarships that have done research regarding the topic so far.

This literature review reveals that researchers consider four key aspects when studying microblogging. According to them, these aspects are:

(1) The message: the tweet itself, and all the metadata the person decides to add to their content (hashtags, URLs, mentions, etc.).

(2) The user: who this person is, if it is displayed any location data or any other detail that might be considered (e.g., number of followers, number of likes or mentions related to a tweet).

(3) The technology: how to interact with Twitter API (e.g., different applications and programs to interact with the Twitter interface, how to retrieve data).

(4) The concept: overviews, discussions, and reviews (e.g., about scholarships, methodologies, ethics).

In addition, they also affirmed that researchers also consider: the domain (the different perspectives that research is undertaken), the data (size, depth, and length of studies, how it is collected and stored), and the method (which methodology will be applied; from data retrieving and pre-processing to data analysis.

It is also worth mentioning the set of approaches they found that researchers usually undertake using the same data set of publications from 2007 to 2011.

(1) Analytic: researchers employ various types of analysis, such as content analysis, semantic analysis, data analysis, and social network analysis, using both quantitative and qualitative approaches.

(2) Design and development: the creation of proposed systems to explore, experiment, demonstrate, model, predict, or simulate specific outcomes.

(3) Examination: authors utilize methods such as review and surveys, biography, case studies, ethnography, evaluation, interviews, and community approaches to examine and study their research topics.

(4) Knowledge discovery: researchers utilize established methods from fields such as artificial intelligence, mathematics, and statistics to perform tasks like data mining, text mining, and natural language processing. Also, advancements in novel algorithms to apply in other spheres.

The outcomes presented above give us a picture of what has been done from the beginning of Twitter, until 2011. Furthermore, from 2013 to 2023, there have been significant advancements and expansions in research topics related to Twitter. If you simply type "Twitter" in Google Scholar, nearly 8 million results will come up from the most diverse academic scholarships and topics. Of

course, not all necessarily will analyse Twitter or use Twitter data to do analysis, but one can get the importance and the reach of the platform in academic research.

For instance, researchers have explored a wide range of topics including communication patterns in specific contexts in the political sphere (Adi, Erickson, and Lilleker 2014; Waters and Williams 2011; Golbeck, Grimes, and Rogers 2010; Aragón et al. 2013), the influence of hashtags on specific contexts (Bruns and Stieglitz 2012), group polarization (Yardi and Boyd 2010), crisis communication (Flew et al. 2014), the impact of celebrities on social media (Masuda, Han, and Lee 2022; Yuan and Lou 2020; Moraes et al. 2019), protests (Theocharis et al. 2015; Earl et al. 2013), and even scholarly communication (Holmberg and Thelwall 2014). They have also delved into social interactions and their effects on human behaviour (Fischer and Reuber 2011), semantic analysis of social movements (Xiong, Cho, and Boatwright 2019), information diffusion (Firdaniza et al. 2021; J. Yang and Counts 2010), interpersonal communications within mediated environments (Wu et al. 2011), and as the main objective of the present dissertation, opinion mining (or sentiment analysis).

These examples demonstrate how computational sciences have increasingly collaborated with disciplines in the humanities, such as history, geography, anthropology, politics, economics, and cultural studies. Moreover, health sciences, including psychology, communication, and media studies, have also witnessed a surge in research collaborations with Twitter. This interdisciplinary approach has fostered the development of diverse tools and methodologies, facilitating a fruitful synergy between computational sciences and the humanities.

An interesting and contemporary study that produced thousands of publications in the last years focus is the COVID-19 pandemic. Ever since the coronavirus rapidly spread across the globe, Twitter has become a hub for billions of conversations revolving around the virus. In early 2020, the Twitter development team pondered how they could contribute to the situation and subsequently created a remarkable resource—a unique dataset that offered researchers real-time access to public discussions about the COVID-19 pandemic. This dataset proved invaluable in facilitating expedited and simplified exploration, enabling researchers to make significant discoveries more rapidly (A. Chen 2021).

This initiative played a vital role in various areas of research, including tracking the spread of the disease, understanding the dissemination of misinformation, crisis management, emergency response, community communication, public sentiment towards the disease and vaccines, as well as the development of machine learning and data tools to assist the scientific community in addressing key questions related to COVID-19. Reflecting on this, the former CEO and co-founder of Twitter, Jack Dorsey, tweeted: "Public conversation can help the world learn faster, solve problems better and realize we're all in this together. Facing a devastating global pandemic brings that, and Twitter's role, to light" (Tornes 2020).

This example vividly illustrates the interconnectedness between disciplines when it comes to engaging with and comprehending a subject of global significance. It also underscores the importance of Twitter data in facilitating swift responses to critical events. It serves as just one compelling instance of the numerous opportunities for academic research within the realm of social media analysis. Furthermore, the dynamic and fast-paced nature of social media generates an abundance of daily content that demands continuous analysis.

As for the humanities and social sciences, since the internet is one large network, social network analysis (SNA) has been extensively used to understand the connections within social media networks. The roots of social network research date way before the invention of the digital platforms we have today, going back to the classical approaches related to relation and structural sociology (J. Scott 2012). It can be defined as a "broad strategy for investigating social structures," where it is prioritized the social context of the actor, its relationships, its ties, the characteristics of these ties, and communities (Otte and Rousseau 2002). Twitter has become a valuable platform for studying and analysing historical events, social phenomena, and cultural trends.

Contemporary, it is of extreme importance to integrate any kind of social media to social network analysis. Like its counterpart social media platform WhatsApp, Twitter serves as a paradigmatic illustration of how a specific technological tool evolves into an infrastructure that supports and moulds various aspects of everyday life, ranging from personal to economic, spiritual to political. This phenomenon exemplifies what we refer to as "life technologies" (Cruz and Harindranath 2020), which profoundly shape nearly all dimensions of social existence. In a similar vein to WhatsApp, Twitter has become an inseparable mediator of social existence, facilitating crucial daily interactions among its users. Twitter's network structure and user interactions make it a valuable platform for conducting social network analysis. Researchers can examine the relationships between users, identify influential individuals or groups, and explore the Twitter network's diffusion of information and ideas.

Furthermore, not only the network aspect but Twitter is also a rich source of data and insights for advancing historical research. A phrase from the historian and professor Igor Lyman during a training session,[15] "What happened seconds ago is already history and it can be analysed" made me think about how we need to think about the use of historical data to narrate historical contemporary events, and we must think all the implications around it, such as data storage, new tools, etc. Twitter's real-time nature provides historians with the ability to capture and analyse historical events as they unfold, amplify marginalized voices that may have been historically underrepresented, shed light on

---

[15] Prof. Lyman gave training sessions in the 2nd WEMov Training School organised by the international network "Women on the Move", funded by E-Cost (CA19112). The event occurred at the University of Tuzla, in Bosnia and Herzegovina, from 6 to 9 June 2023.

alternative narratives, challenge dominant historical interpretations, and promote a more inclusive understanding of the past. Also, as it was highlighted before, it is a powerful tool for studying social movements and activism.

Beyond that, scholars also use Twitter to connect and build a network where they can share and promote their publications, events, and insights and connect with other people around the globe that might have the same research interest as them, particularly in the social sciences and humanities (Mohammadi et al. 2018). As a "push model", researchers can put their research in a space to be seen, not wait for people to find it. There is evidence that suggests a correlation between social media mentions and download and citation count, demonstrating higher popularity among those who had social media mentions (Shuai, Pepe, and Bollen 2012), especially Twitter (Klar et al. 2020; Eysenbach 2011). Without much clear data yet, there might be a possible association of Twitter being a means to connect scholars to the non-academic public, making this difficult interaction slightly more feasible and expanding research findings with the public.

3.2.   The Debate on the Ethics of Using Twitter Data and Misinformation

The use of Twitter data in research has sparked a significant debate regarding the ethical implications involved. Several key ethical considerations arise when utilizing Twitter data, including privacy concerns, informed consent, data ownership, and potential harm to users. However, there is an absence of consensus in web science and related fields over how to solve those issues (Webb et al. 2017). Many of the discussions go around "big data" and the use of automated processes to collect large amounts of data and its implications to participants, privacy, identification, etc.

Discussions also focused on the handling and archiving of these data. There is no agreed best practice around it, but the debate is often related to different levels of risk associated with the data sources. Users normally are not approached directly to solicit their informed consent to take part in research, instead, consent is often assumed to have been given by the user's acceptance of Twitter's Terms of Service. It states that

> "by submitting, posting or displaying Content on or through the Services, you grant us a worldwide, non-exclusive, royalty-free license (with the right to sublicense) to use, copy, reproduce, process, adapt, modify, publish, transmit, display and distribute such Content in any media or distribution methods now known or later developed (for clarity, these rights include, for example, curating, transforming, and translating). This license authorizes us to make your Content available to the rest of the world and to let others do the same."[16]

---

[16] Twitter Terms of Service (Available at: https://twitter.com/en/tos)

However, researchers should address the data collected carefully to not put any of the subjects of research in danger. Researchers need to engage in ongoing discussions surrounding the ethical implications of using Twitter data and collaborate with experts in fields such as data ethics, privacy, and social sciences to develop best practices and guidelines for responsible data use. As shown here, many fields of research have been conducting relevant research using social media, and more specifically, Twitter data. Those fields that deal with populations in fragilized or dangerous situations must take special care to protect those people, their stories, families, and identities from other greater dangers. It is not because the data is available that it alters the standard obligations relating to the minimisation of harm. In terms of general data access and management, a just ethical behaviour would involve adhering to the same norms for subject data and Twitter data. Additional precautions should be taken to reduce risks in the case of data concerning Subjects, such as never revealing individual-level Subject information. The permission of Twitter to conduct research using its platform is demonstrated in its provision of a Developer API and the terms that go with it: in essence, Twitter offers its platform for research (but only under certain conditions, e.g., not for commercial use if you do not have this type of access). Furthermore, with the introduction of the new academic product track, researchers are required to inform Twitter of their intentions and obtain approval for their work. Also, is made clear to the users and developers that

> "one of Twitter's core values is to defend and respect the user's voice. This includes respecting their expectations and intent when they delete or modify the content, they choose to share on Twitter… We believe that business consumers that receive Twitter data have a responsibility to honour the expectations and intent of end users."[17]

Moreover, if the researcher needs to share Twitter content that he/she/they obtained, Twitter suggests that the best way to do so is by sharing tweet IDs, direct message IDs, and/or User IDs, which the end user of the content can then rehydrate (i.e., request the full tweet, user, or direct message content) using the Twitter APIs. This helps ensure that end users of Twitter content always get the most current information. Although challenging, is therefore necessary to continue discussing ethics to improve the quality and value of the research produced.

Twitter offers many opportunities and also poses many challenges to academic research. Debates related to the distribution of misinformation and how to overcome it had a spotlight on presidential elections in the United States and Brazil. The validity and liability of data in determined cases and situations, also regarding methodologies and their challenges and pitfalls. Biases and representativeness, as dealing with Twitter data represent a specific segment of the population, and their views may not be fully representative, where researchers should be mindful of these biases and

---

[17] Twitter Compliance Firehouse API (Available at: https://developer.twitter.com/en/docs/tweets/compliance/overview)

limitations when concluding Twitter data, ensuring that their findings are contextualized and interpreted with caution.

Twitter's dynamic and interconnected nature provides researchers with a unique environment for historical and social research. By utilizing appropriate methodologies, addressing ethical considerations, and acknowledging the challenges and opportunities presented by Twitter data, scholars can unlock valuable insights into an extremely wide range of subjects, topics, and themes.

Chapter 3: Methodology and Data Extraction Process

In the preceding chapters, we delved into an exploration of historical contextualisation through the lens of the historical, political, and migration framework. This intricate tapestry led us through the ebbs and flows of human movements internationally and the flows to Brazil, exposing the intricate sociopolitical underpinnings that influenced the movements and the laws in the country. It was possible to understand the role of social media, particularly Twitter, in shaping public discourse and opinions, catalysing political movements within Brazilian society. This initial chapter, a nexus of migration, politics, and media in Brazil, laid the groundwork for a comprehensive grasp of migration dynamics, with a particular focus on migratory patterns converging towards Brazil and the far-reaching consequences thereof.

Subsequently, the second chapter meticulously unfurls a panorama of Twitter in its capacity as a potent instrument for scholarly inquiry. This segmented immerses itself in exploring Twitter's attributes, its manifold prospects for academia, and the consequential ramifications engendered by harnessing its expansive reservoir of data for scholarly investigation. Such a discerning immersion, by concretizing the operational modus of Twitter and harnessing its intrinsic potential in the context of social studies, constitutes an indispensable backdrop for the forthcoming discourse on sentiment analysis.

Within this strategic narrative arc, the last chapter before the data analysis delves into a comprehensive understanding of the tools necessary and chosen to analyse the tweets. It serves as a pivotal precursor to the endeavour of this thesis – use sentiment analysis and NLP methodologies to grasp people's opinions and discourses about migration. By contextualizing our exploration within the broader narrative portrayed before, we enhance our ability to extract meaningful insights from the data analysis.

This thesis is an entanglement between social sciences and computational sciences, and as mentioned in the literature review, is gaining some room in different spheres of social science. Texts have been analysed for years now in the offline world, reports, books, diaries, songs, letters, etc., have been material for many analyses and unleashed many discoveries to the world. The discussions on social media cannot be put behind, and they prompted the academia to explore new methodologies capable of investigating the large number of texts posted every day. The following lines contain a brief description and explanation of the tools, programs, systems, and libraries used to develop the data analysis. Even though a comprehensive description of the steps will be demonstrated for an

overall understating, the chapter will be focused on the tools used to understand what will be demonstrated in the analysis and does not have all the content or an extensive explanation of the tools. For further information about it, please refer to their official sources.

1. Python

Programming is the act of writing instructions for a computer to do some specific tasks, whereas normally an algorithm is designed as a step-by-step set of instructions/procedures/rules for solving a problem or accomplishing some end. Algorithms can be encoded in any of the existing programming languages, such as Java, Python, C++, and more.

Python is a high-level object-oriented programming language that empowers developers to articulate tasks they wish a program to execute in a lucid and uncomplicated manner. Subsequently, these instructions are automatically converted into executable commands by the computer's processor. Python, conceived in the early 1990s by Guido van Rossum, distinguishes itself through its less convoluted and organized structure, setting it apart from other commonly utilized languages such as Java, C, and C++ (Horstmann and Necaise 2016).[18] This simplicity renders it more approachable for beginners to grasp and immerse themselves in.

Figure 9. Python logo. Source: www.python.org.

In the context of this thesis, Python was selected as the programming language due to its simpler and cleaner syntax (given the author's non-programmer background), its widespread recognition, and the presence of an extensive range of libraries that facilitate the analysis. The Python version employed in this research was 3.10.3.

Among the multitude of options for conducting the analysis, Anaconda was chosen as the interpreter for running and interpreting the program, while the Jupyter Notebook application facilitated the creation and editing of code using the Python scripting language. Jupyter, in its role as the editor, meticulously scrutinizes the program's code, ensuring the precise sequence of each step, and promptly flags any errors requiring correction. This choice was guided by the advice of the supervisor and the author's familiarity with the tools.

---

[18] While many often speculate that the name "Python" may allude to the snake, it actually originates from a vintage BBC television comedy sketch show known as "Monty Python's Flying Circus."

Before diving into the step-by-step process of analysing the data until the sentiment analysis, it is necessary to present a comprehensive list of Python libraries and tools that were used throughout the analysis. The next topics covers a brief explanation of the libraries used for text pre-processing, data manipulation, analysis, and visualization.

1.1. Python Libraries

A Python standard library is a vast collection of built-in modules and packages. These modules offer access to system functions, such as file operations that are normally beyond Python's reach. Additionally, there are Python-written modules that provide standardized solutions for common programming challenges. Some of these modules are designed to make Python programs more portable by abstracting away platform-specific details.[19] They are usually created by a community of programmers and made free for anyone over the internet. Instead of writing many lines of code, the libraries are already there to execute the task, and each of its modules serves a specific purpose, as will be demonstrated next. Generally, it is necessary to install and import these libraries through specific processes or lines of code, the way to do it depends on the interpreter and editor one decides to use.

Pandas

Pandas is one of the most important Python libraries for data manipulation and data analysis. It can read and manipulate tabular data (known as DataFrames) and has practical tools for aggregating, analysing, and cleaning data. *Pandas* have many functionalities, offering a wide array of tools essential for data analysis. It enables data transformations, from sorting and selecting subsets to performing crucial calculations and merging DataFrames.

In the realm of data analysis, *pandas* can perform tasks such as importing datasets from diverse sources like databases and spreadsheets, ensuring data cleanliness by handling missing values, and reshaping datasets to optimize analysis. Beyond this, it efficiently aggregates data, unveils insights through metrics, and can visualize datasets enhancing its utility, enabling the identification of patterns and trends that might otherwise remain hidden. *Pandas* is not restricted to numerical data, it extends to time series analysis and the exploration of textual data, broadening its scope and applicability.

---

[19]   For   more   information   about   Python   libraries   please   refer   to:
https://docs.python.org/3/library/intro.html#:~:text=The%20%E2%80%9CPython%20library%E2%80%9D%20contains%20several,such%20as%20numbers%20and%20lists.

Furthermore, *pandas* integrates well into the Python ecosystem, has efficient and concise code, employs a user-friendly model for representing data, has extensive functionality, and performs well with large datasets.[20]

Csv

Csv enables users to read CSV[21] data into Python data structures, like lists or dictionaries, and also writing data from Python structures back to CSV files. The CSV file format is a commonly used data format while exporting/importing data to/from spreadsheets and data tables in databases.[22]

NLTK

The Natural Language Toolkit (NLTK) library stands out as a prominent tool for natural language processing (NLP) within the Python programming ecosystem. NLP, a realm within computer science and artificial intelligence, strives to empower computers with the capability to comprehend written texts and spoken language. Specifically, it focuses on decoding unstructured textual data, often found in vast and varied datasets from sources like social media, including tweets (Hovy 2020). This unstructured nature implies that the data lacks predetermined categories, and standardized lengths, and is typically replete with misspellings, characters, emotions, punctuation marks, URLs, and other forms of what is commonly referred to as "noise." These noise elements are irrelevant and can significantly undermine the accuracy of any subsequent analysis.

The NLTK library, therefore, assumes the role of an interface facilitating a spectrum of tasks necessary for conducting sentiment analysis. A pivotal step in this process is data preprocessing, aiming to purify the data by removing the aforementioned noises. This critical procedure involves tokenization, stemming, lemmatization, and the parsing of textual data, all of which can be adeptly executed using the NLTK library.

NLTK stands out as an exceptional learning resource, primarily due to Python's swift learning curve. It comprehensively encompasses an array of NLP tasks, rendering it a comprehensive solution. Its user-friendly nature further adds to its appeal, making it approachable and amenable to work. Consequently, NLTK has gained prominence as one of the most widely embraced libraries within the NLP community (Hardeniya 2015).[23]

---

[20] For more information about pandas: https://pypi.org/project/pandas/
[21] CSV (Comma-Separated Values) is a widely used file format for storing structured data. It uses commas to separate values into rows and columns, making it easy to exchange and manipulate data between different software applications.
[22] For more information about Csv library: https://docs.python.org/3/library/csv.html .
[23] For more information about NLTK library: https://www.nltk.org/

Regex

Regular expressions, often abbreviated as *regex*, are powerful and flexible patterns used to match, search for, and manipulate text strings in various programming languages and applications. They provide a concise and standardized way to describe complex search patterns within strings, through a combination of characters and metacharacters that define a specific pattern to be matched within a text. These patterns can be simple, such as matching a specific word, or highly complex, such as extracting specific data formats from a large body of text. Regular expressions are widely used for tasks like text validation, data extraction, text substitution, and more.

Regex are a valuable tool for tasks that involve text processing and manipulation, as they allow developers to efficiently identify and work with specific patterns within strings. However, they can be complex to learn and master due to their powerful capabilities and diverse syntax.[24]

Codecs

Codecs, a name derived from "coder-decoder," facilitate the exchange and manipulation of text data encoded in different character encodings or formats. These libraries or modules serve as the linguistic bridges, allowing developers to navigate the intricacies of character encodings, transcending the barriers that arise when dealing with text data from diverse sources. Character encoding, fundamentally, defines how characters are represented as bytes, making possible to encode text into bytes for storage or transmission, decode bytes into human-readable text, and manage errors that may arise during these transformations. This capability ensures the harmonious integration of text data from various origins, fostering compatibility and interoperability in data processing workflows. In practical terms, encoding and decoding operations with codecs are invaluable for data handling. [25]

Matplotlib

Matplotlib is a library for creating data visualizations and plots. It serves as a versatile and comprehensive tool to generate a wide array of static and interactive graphs and charts. *Matplotlib* is renowned for being flexible, allowing users to exercise precise control over nearly every aspect of a graph's appearance. It offers an extensive repertoire of chart types, spanning simple line plots and bar charts to intricate scatter plots and heatmaps. Furthermore, Matplotlib is integrated with Jupyter Notebook and stands as a foundational tool in the Python ecosystem for data visualization.[26]

---

[24] For more information about regular expression and how to employ it: https://regexr.com/
[25] For more information about Codecs: https://docs.python.org/3/library/codecs.html .
[26] For more information about Matplotlib: https://matplotlib.org/ .

Seaborn

Seaborn is a Python data visualization library built on top of *Matplotlib* and it is designed to provide a high-level interface for creating informative and visually appealing statistical graphics. Seaborn is particularly well-suited for working with complex datasets and is often used in data exploration and analysis, as well as in data-driven storytelling. It is simple and generate aesthetically pleasing visualizations with minimal code. The library supports the creation of statistical plots such as histograms, kernel density plots, regression plots, and also simplifies the process of visualizing complex relationships between variables through functions like pair plots, joint plots, and heatmap visualizations.[27]

Ploty

Plotly is a library for creating interactive and visually appealing data visualizations. Unlike static plots, *Plotly* generates dynamic, web-based charts and graphs that can be explored and manipulated by users. It supports a wide range of chart types, including scatter plots, line charts, bar charts, heatmaps, and 3D plots, making it suitable for a variety of data visualization needs. It has and interactive nature and allows users to zoom, pan, hover over data points for details, and even export plots as standalone HTML files for sharing online. [28]

Cufflinks

Cufflinks is a library that acts as a bridge between *Pandas* data frames and *Plotly*, integrating the capabilities of both libraries to simplify the process of creating interactive visualizations. It essentially extends the data frame objects to enable easy generation of interactive plots and charts directly from the stored data. It eliminates the need for writing complex code by providing a high-level and intuitive interface, making it easier to users who are already familiar with *Pandas*.[29]

Translators

The *Translators* library is a Python package designed to streamline the process of translating text between different languages using various translation services and APIs. It acts as a versatile interface for multiple translation providers, including Google Translate, Microsoft Translator, Bing, and others, simplifying the complexities of making API requests and handling responses by providing a unified and user-friendly interface. The flexibility and support for multiple translation services give

---

[27] For more information about Seaborn: https://seaborn.pydata.org/index.html
[28] For more information about Ploty: https://plotly.com/
[29] For more information about Cufflinks: https://www.analyticsvidhya.com/blog/2021/06/advanced-python-data-visualization-libraries-plotly/

users the freedom to choose the most suitable provider for their specific needs, enhancing the efficiency of multilingual data processing and analysis workflows.[30]

TQDM

This library is a Python package that provides a simple way to add progress bars to various iterations and loops within code. It offers a visually appealing and informative progress meter that helps users monitor the progress of lengthy operations, such as data processing, file downloads, or complex calculations, in real-time.[31]

WordCloud

A word cloud is a popular visualization technique that displays words from a text dataset, with the size of each word proportional to its frequency in the text. The *WordCloud* library is a tool for generating visually striking word clouds from textual data. It provides a visually way to identify the most prominent and recurring terms within a body of text. Word clouds are frequently used in data analysis, natural language processing, and text mining to gain insights into the key themes and topics present in textual data.[32]

1.2. Data Pre-Processing: Steps and Terms

Preprocessing text data is a crucial step in sentiment analysis, as it helps clean and prepare the text for analysis, improving the accuracy of sentiment classification. As mentioned before, some of the libraries can help filter and clean the data to achieve the determined objective. The preprocess usually involves a series of techniques that transform the raw data into a form that is better for doing the analysis. The process may vary regarding the dataset and can be modified to fit one's needs.[33]

- Tokenization: Split the tweet into individual words or tokens (each word, element, and sign, is a token).
- Removing Special Characters and Punctuation.
- Removing Stop Words: removing common and irrelevant words that are unlikely to convey much sentiment. Stop words are words that are very common in a language and do not carry much meaning, such as "and", "the", "of ", and "it". These words can cause

---

[30] For more information about Translators: https://pypi.org/project/translators/
[31] For more information about TQDM: https://pypi.org/project/tqdm/2.2.3/
[32] For more information about Wordcloud: https://pypi.org/project/wordcloud/
[33] For more information about the preprocessing steps to data analysis refer to: https://www.datacamp.com/tutorial/text-analytics-beginners-nltk or "Text analysis in Python for social scientists: discovery and exploration", by Drik Hovy (2020).

noise and skew the analysis if not removed. The NLTK library provides a list of stop words in several languages and one can also add any other that might be out of the provided list.

- Removing URLs and User Mentions.
- Removing hashtags
- Translation

Some other steps in data filtering were also necessary to eliminate rows that contained words and subsequently, subjects that were not useful for the analysis. Each of these steps will be demonstrated and further explained with the codes in the next session.

2. Sentiment Analysis

Sentiment analysis, also referred to as opinion mining, is a technique utilized to ascertain the emotional tone or sentiment conveyed within a given text. This involves harnessing the power of natural language processing (NLP) to systematically discern, gauge, and dissect emotional states and subjective information. It empowers the computational determination of text polarity—whether the underlying sentiment of an opinion is positive, negative, or neutral. Sentiment analysis finds an extensive array of applications, including the monitoring of social media, the analysis of customer feedback, and conducting market research, and call also tackles a wide range of issues that can interest researchers from different fields, such as sociology, psychology, economics, and political science (Hutto and Gilbert 2014).

One of the chief challenges encountered in sentiment analysis lies in the intricate nature of human language. Textual data frequently embodies elements of sarcasm, irony, and other forms of figurative expression that can prove demanding to interpret through conventional methodologies. However, recent research and methodology development permitted to performance of sentiment analysis on large datasets with a high degree of accuracy (Ribeiro et al. 2016).

In contemporary times, there exist numerous methodologies for conducting sentiment analysis on textual data, and these techniques exhibit a spectrum of complexity and accuracy. The principal and extensively employed techniques can be categorized into three distinct groups: lexicon-based and rule-based approaches, a machine learning-based approach, and a deep learning approach. Lexicon-based and rule-based approaches encompass the utilization of predefined guidelines to ascertain the sentiment within a given text. These guidelines are often rooted in external linguistic knowledge. Such approaches tend to offer relative simplicity in terms of implementation and interpretation. The

analysis hinges upon lexical and syntactic attributes of the text, encompassing factors like the presence of positive or negative words and phrases (Zong, Xia, and Zhang 2021).

Conventional machine learning techniques entail the training of a model to discern the sentiment expressed in a text fragment, relying on a collection of labelled training data. Diverse ML algorithms can be employed to train these models (Zong, Xia, and Zhang 2021). This approach can yield heightened accuracy compared to rule-based analysis, particularly when tackling intricate textual content. However, it necessitates a substantial volume of labelled training data, potentially resulting in greater computational demands and time requirements.

Deep learning methods have undergone significant evolution in recent years. They have the capability for feature representation and an integrated end-to-end learning architecture (Zong, Xia, and Zhang 2021). These methods leverage pre-trained models that have been fine-tuned on extensive volumes of text data. Operating through intricate neural networks, these models encode the context and essence of the text, enabling them to attain cutting-edge accuracy across a diverse spectrum of NLP tasks, which encompass sentiment analysis.

Nevertheless, it's imperative to note that the last two models necessitate substantial computational resources and might not be pragmatic for all conceivable use cases. For this thesis, it will be used a lexicon-based approach, following the recent applications of relevant research regarding the topic of migration (see Rowe et al. 2021; Freire-Vidal, Graells-Garrido, and Rowe 2021; Inuwa-Dutse, Liptrott, and Korkontzelos 2020). Following their methodology and at the same time deploying a gold-standard tool on lexicon-based and ruled-based methodology, this thesis will use VADER (Valence Aware Dictionary and sEntiment Reasoner) to conduct the sentiment analysis.

2.1. VADER

VADER (Valence Aware Dictionary and sEntiment Reasoner) is a lexicon and rule-based sentiment analysis tool designed for processing and understanding the sentiment (positive, negative, or neutral) expressed in text. It is a gold-standard sentiment lexicon specifically tuned for analysing sentiments in a micro-blogging context, which often includes informal language, slang, contractions, negations, and emoticons (Hutto and Gilbert 2014). It is available as part of the Natural Language Toolkit (NLTK) library in Python and has gained popularity for its simplicity and effectiveness in sentiment analysis tasks involving short and informal text (Ribeiro et al. 2016).

Opting for VADER over other methodologies is rooted in (1) its apt suitability for handling social media data, (2) its proficiency in comprehending nuanced sentiment across a wide spectrum of post types, (3) its robustness through training on diverse datasets of varying sizes, (4) its recognition as an effective approach in prior studies, including comparisons to other more sophisticated

methodologies (see Ribeiro et al. 2016; Hutto and Gilbert 2014), (5) its heuristics formulated a qualitative analysis of textual attributes that perceive the intensity of sentiment in the text, and (6) its user-friendly implementation and cost-effectiveness.

Regarding the heuristics, VADER includes treatments for: punctuation (e.g., distinguishing between "!" and "!!!" in a text), capitalization (e.g., distinguishing between "amazing" and "AMAZING"), degree modifiers (e.g., "this is extremely good"), the constructive conjunction 'but' for polarity alteration, and tri-gram examination for negation identification (e.g., "The weather isn't really good today") (Ribeiro et al. 2016).

Concerning sentiment classification,

> "VADER provides a normalised, weighted composite score which captures the polarity and intensity of individual tweets. The score ranges from -1 to +1, representing the most extreme negative to most extreme positive sentiment respectively. Intuitively, to derive the score, VADER assigns a score to each word in a tweet, ranging from -4 (extremely negative) through 0 (neutral) to +4 (extremely positive) based on positive and negative text features identified in the text. These scores are then aggregated and normalised to range between -1 and +1" (Rowe et al. 2021).

As demonstrated it is possible to extract valuable insights from textual data and gain a deeper understanding of people's opinions and attitudes. The tools and processes mentioned so far will be demonstrated during the analysis. After preprocessing the data, the overall step-by-step process of performing sentiment analysis using VADER can be summarised as follows:

- Loading and Preparing Data: Importing the text data and applying the pre-processing steps outlined earlier.
- Sentiment Scoring: Applying VADER to assign sentiment scores to each text document or sentence, producing polarity scores for positivity, negativity, and neutrality.
- Interpreting Results: Analysing the sentiment scores and exploring patterns in the data. Demonstrate techniques to visualize sentiment distributions and uncover insights from the sentiment analysis.
- Comparative Analysis: Showcasing how sentiment analysis can be used to compare sentiment across different categories, periods, or groups within the data.

3. Twitter Developer Access and Data Extraction

The initial step to accomplishing sentiment analysis involves creating a well-formed database that aligns with the research objectives. For this study, the database was constructed using data accessible through the Twitter API v2 via the Twitter Developer Account.

To collect Twitter data, it was necessary to use Twitter's application programming interface (API). As aforementioned, the API facilitates data exchange between applications and services through code, granting access to various data types, such as tweets, users, direct messages, lists, trends, media, and places, in a structured manner. Obtaining a Twitter developer account is mandatory to access this data. This account provides access to various API functionalities, such as retrieving tweets, posting tweets, accessing user information, and performing searches. The Twitter API is the official source of getting tweets for research to stay compliant with our developer policy.

The general steps to obtain a Twitter developer account can be further found on Twitter's Developer Website, however, it will be exemplified by the steps taken by the author.[34]

Obtaining access to Twitter's API v2 through the Academic access:

(1) It is necessary to create a personal Twitter Account (for this thesis, the author's account was used to apply for a developer account).

(2) Access the Twitter Developer Portal (https://developer.twitter.com) and sign in with the Twitter account.

(3) Apply for a Developer Account: Once logged in, you need to apply for a developer account. Twitter may ask you to provide additional information about your use case and how you plan to use the API. Be clear about the purpose of your application and how you intend to comply with Twitter's Developer Agreement and Policy.

(4) Create a Twitter App: After the developer account is approved, it is possible to create a new Twitter App. This app will be associated with the developer account and will have a unique set of API keys and access tokens. They are essential for making the API requests and unique to your login. After generating API keys and tokens, it is wise to make sure they are saved and secure.

(5) Applying for the Academic Research Product: After getting the Developer Access, it was necessary to apply for the academic research product. This is an access made for professional academic research which implies having a specific research purpose with the data.

---

[34] The Academic Developer Access for this research was obtained on 31st August 2022. Please note that those instructions and steps refer to this period. Since April 2023 the rules to obtain access to the API have changed. Please refer to the Official Website for the current process and most recent updates.

a. It is requested to explain the project, the reasons to do it and how and where the research will be applied. The application is manually reviewed by Twitter.

b. This product track includes elevated features and functionality tailored for academics: it grants access to the Twitter API v2; it enables access to get historical Tweets from the entire archive of public conversation on Twitter, dating back to 2006 (using the full-archive search endpoint); it allows the retrieval of 10 million tweets per month; it has more advanced filter options to return relevant data.[35]

The examples of information available to researchers in a public tweet include the tweet text, the tweet ID (that uniquely identifies a Tweet), the time at which the tweet was created, public metrics associated with the tweet (e.g., number of retweets, number of likes etc.), public user information (username, user ID, user bio, profile image, URL, etc.), and tweet annotations (some tweets are annotated based on the topic that they are about and the named entities present in those tweets).

3.1. Data Extraction

From the inception of this research, our central focus has revolved around delving into the intricate nuances of sentiments linked to the topic of migration in Brazil. The initial inquiry centred on devising a geospatial framework to capture tweets emanating from Brazil. As evidenced by prior literature (Rowe et al. 2021; Huang and Carley 2019), only a minute fraction of tweets are geotagged, indicating precise location coordinates. As it will be demonstrated in the next chapter, only a small amount of the retrieved tweets on this research contained their geolocation as well. Therefore, a strategic decision was made to leverage the language itself (Portuguese) and use Portuguese keywords for tweet retrieval. The query formulation for tweet extraction was orchestrated through the inclusion of the following keywords: '*imigrantes*' (migrants), '*imigrante*' (immigrant), '*imigração*' (immigration), '*migração*' (migration), '*refugiado*' (refugee), '*refugiados*' (refugees), '*crise migratória*' (migratory crisis), and '*ACNUR*' (UNHCR). The establishment of a well-decided crafted filter assumes great importance, as it serves as an assurance that the acquired data aligns with your specific use case and faithfully mirrors the discourse you intend to scrutinize. Furthermore, it reduces time in the subsequent data cleaning process and ensures adherence to access limits.

Data extraction for this research involved the use of a custom script tailored to retrieve tweets containing the specific keywords mentioned above to design the query. An essential part of this

---

[35] As the instructions and possibilities have changed, this page was saved in WebArchive for further knowledge: /web/20230810144812/https://developer.twitter.com/en/blog/product-news/2021/enabling-the-future-of-academic-research-with-the-twitter-api

process included the implementation of effective time management techniques to prevent potential timeout errors. This time management strategy ensured the uninterrupted operation of the script, allowing it to continuously fetch tweets for extended periods. It's worth noting that the crawling process spanned several hours to ensure thorough data acquisition.

Chapter 4: Data Pre-Processing and Analysis

1. Data Pre-Processing

Prior to delving into data analysis, a pivotal step is to undertake pre-processing, also referred to as cleaning. This procedure is critical in elevating data accuracy, quality, and overall efficacy, facilitating a more refined analysis. This chapter will provide a comprehensive overview of the steps taken to fulfil the objectives of this thesis, presenting results across various stages of data pre-processing, analysis, and concluding with sentiment analysis. As detailed in the preceding chapter, Anaconda Navigator and Jupyter Notebook served as the chosen platforms for the entirety of the dataset's cleaning, preparation, and analysis process. The libraries used were also introduced in the previous chapter and are referenced in this chapter without delving into their specific functions. This segment encompasses the removal of special characters, punctuation, URLs, and irrelevant words, all geared toward fostering consistency and elevating text quality. The objective herein is to rectify tweet syntax as necessary, with the overarching aim of enhancing data comprehensibility for machines and undermining ambiguity during feature extraction.

It is worth emphasizing that, as previously elucidated, the academic access not only facilitated the extraction of tweets but also granted access to a diverse array of supplementary features. This comprehensive inclusion holds the potential to enrich and augment the findings of this study significantly. Consequently, in addition to the core data frame housing the tweet content, the compilation of several distinct data frames was conducted, each meticulously stored in separate CSV files. Notably, the text processing procedures were exclusively applied to the main data file, as it is the one that contains textual content. The other data frames encompass diverse information such as location, user details, and metrics like likes, retweets, quotations, replies, hashtags, and annotations.[36] This multi-faceted approach ensures that a holistic spectrum of information is harnessed, thereby fostering a more comprehensive and nuanced analysis.

---

[36] Tweet annotations offer a way to understand contextual information about the tweet itself, separating the tweets by topic. Twitter analyses all tweets but annotates only a portion of them.
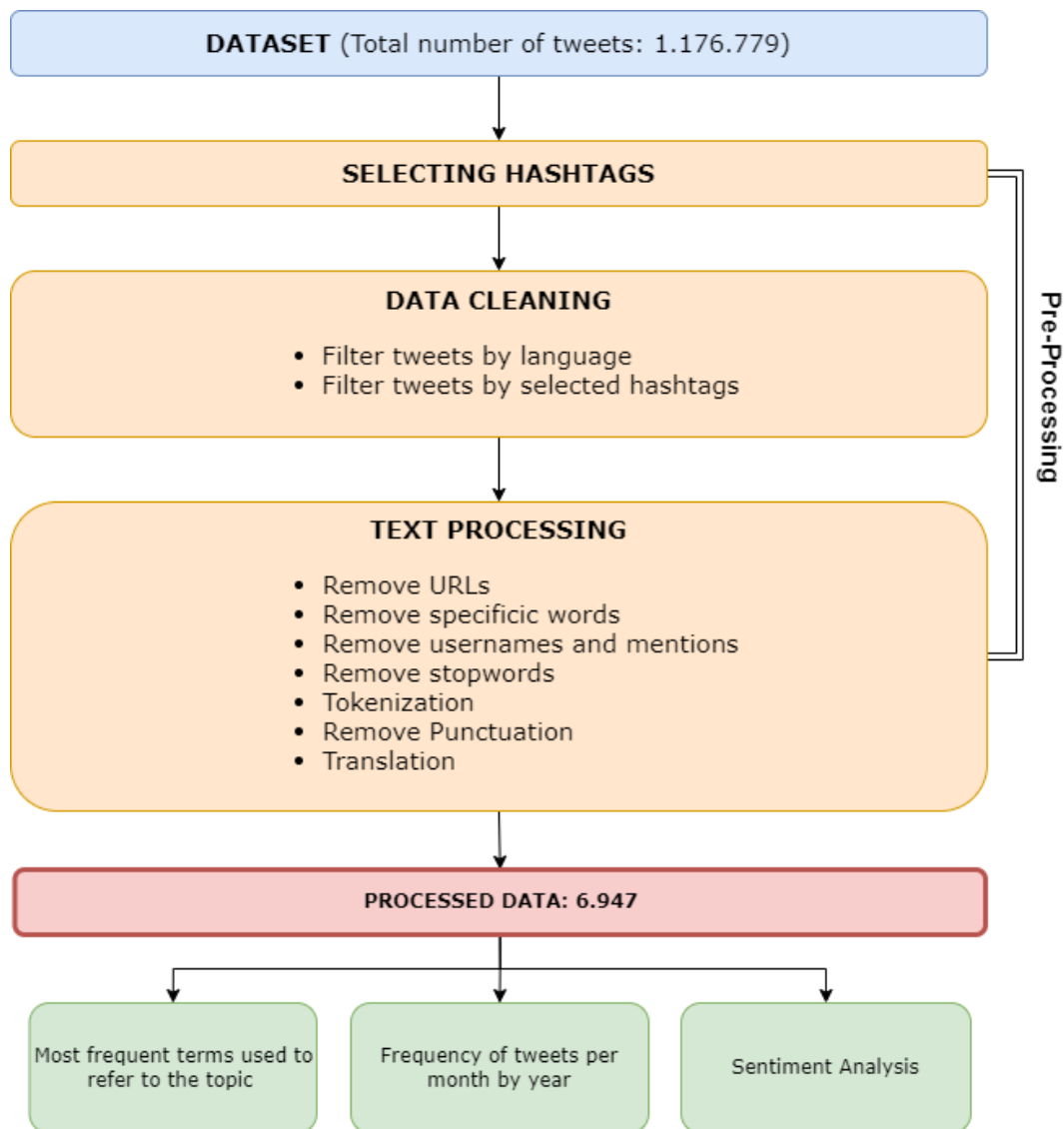
Figure 10. Steps to data pre-processing and sentiment analysis. Source: Author.

As aforementioned, the scarcity of available data regarding where tweets originate, and the locations of their users presents a common challenge in research. Similarly, this study also faced this challenge when attempting to sort tweets by location after gathering them to create the dataset. Out of all the collected tweets, only a small portion—only 1.2%—contained location information. This indicates that a very small number of users allowed their devices to include geographic details when they posted tweets. However, even with this limited geographical data, we found that a significant majority—specifically, 80.05%—of the gathered tweets were associated with Brazil. This discovery was made possible by using specific Portuguese keywords during the process of collecting tweets.
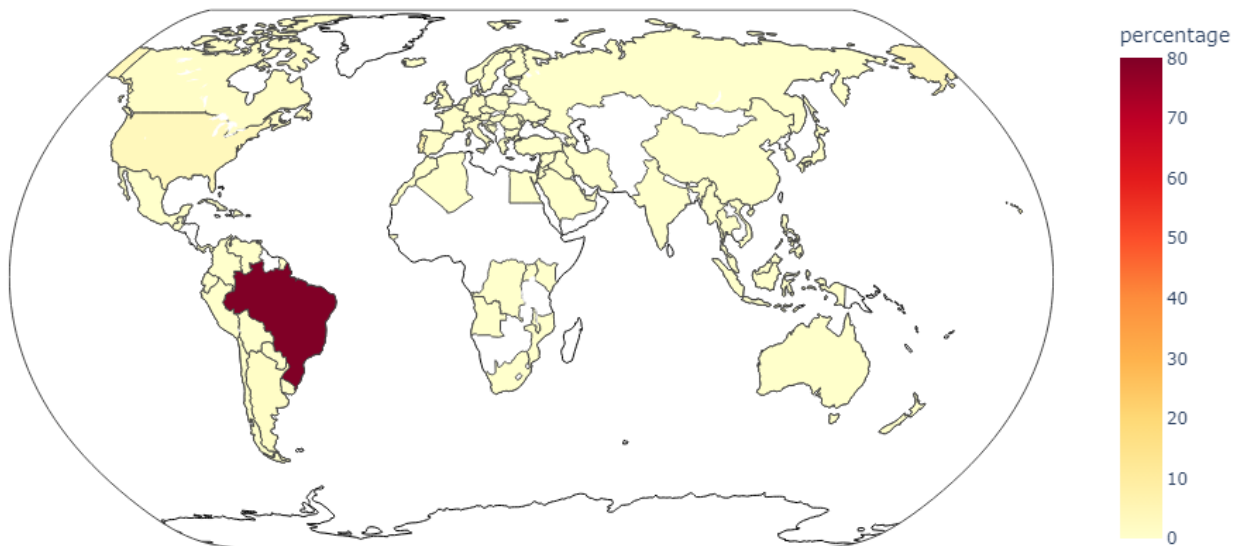
Figure 11. Percentage of tweets by country. Source: Author.

The attempt to employ users' locations to geographically delimit the research revealed that 70.3% of the users displayed such information, a figure exceeding the proportion of geolocated tweets. However, while this data can be automatically provided by enabling Twitter's access to the geolocation of the user's device, it can also be manually updated. A cursory analysis of the location data indicates that while many individuals use this field in their profiles to indicate their current whereabouts, a considerable number also repurpose the field for alternate purposes, such as incorporating emojis (e.g., ' 💙 💚 ') or making statements (e.g., 'On Twitter in support of WikiLeaks & Julian Assange :-)'), among others. Given that this field does not refresh automatically, one cannot deduce the frequency of updates or whether individuals indeed update their stated locations.

Hence, employing the location data would not imbue this study with a robust quality filter for the intended data analysis. Consequently, it was determined that neither user locations nor tweet locations would serve as filters in the pursuit of gathering the perspectives of Brazilians on migrants and refugees. Nevertheless, the observation that 80.05% of the tweets with enabled location information originated from Brazil does suggest that the endeavour engaged with the desired dataset.

1.1. Data Cleaning

Proceeding to the pre-processing phase, the primary data frame initially encompassed a total of 1,176,779 tweets collected using the following keywords: '*imigrantes*' (migrants), '*imigrante*' (immigrant), '*imigração*' (immigration), '*migração*' (migration), '*refugiado*' (refugee), '*refugiados*' (refugees), '*crise migratória*' (migratory crisis), and '*ACNUR*' (UNHCR). Despite the usage of

Portuguese words as conditional statements, other languages were detected in the 'text' column of the dataset. This information becomes accessible during the tweet extraction process. In an effort to refine the dataset further, a decision was made to utilize this column to filter data containing exclusively Portuguese tweets. Recognizing that even if Brazilian users might tweet in languages other than Portuguese, filtering results based on the native language of the population aids in acquiring more accurate data regarding the sentiments of the country's populace. Following the application of this language filter, the total number of tweets was reduced to 775,236.

While a level of filtration pertaining to the topic had been instituted during the query-based extraction, an initial analysis of the outcomes brought to light the presence of certain tweets that tangentially related to the intended subject matter. These tweets aligned with the employed keywords in the query, leading to their inclusion. However, they addressed distinct contexts such as discussions about traffic congestion on a Brazilian road named '*Imigrantes*'. Acknowledging the lack of relevance of such content to the study's central focus, a decision was taken to exclude these extraneous tweets. After a primary analysis of the words in the tweets, a list of the most common words to be extracted came out:

```python
specific_words = [
    'trânsito', 'rodovia', 'bitcoin', 'tráfego', 'AnchietaImigrantes',
    'ecovias', 'Anchieta', 'Radial', 'radial', 'dengue', 'copa', 'Leste',
    'anchieta', 'TIM', 'tim', 'Kaysar', 'novela', 'Bandeirantes',
    'operacaobetalab', 'bbb', 'transito', 'trânsito', 'orcars', 'oscar',
    'grammy', 'grammys', 'kaysar', 'timbetalab', 'masterchefbr', 'enem',
    'futebol', 'bundesliga', 'corinthians'
]
```

Figure 12. Code to remove specific tweets. Source: Author.

All the tweets containing these words were extracted from the data frame, resulting in a total of 659.003 tweets. After cleaning and text processing the column that contained the text of the tweets, it was possible to have an idea of the main terms used to refer to migrants and refugees and their frequency during the period.

Figure 13. Wordcloud with the most frequent terms used to refer to the topic (DataFrame with 659.003 tweets). Source: Author

While the initial filtering successfully reduced a significant amount of noise and allowed for a more focused analysis of the research topic, a subsequent review of certain rows revealed that many entries were still unrelated to the subject. In an effort to enhance the accuracy of the analysis and optimize efficiency, a strategic decision was made. Instead of outright excluding unrelated topics, the chosen approach involved gathering the top 100 most frequently used hashtags each month. These hashtags were then manually screened to select those explicitly related to the subject of migration. Subsequently, tweets containing these chosen hashtags were filtered, followed by the preprocessing and analysis of the data.

The decision to employ hashtags in this approach was not arbitrary, as mentioned earlier, hashtags have evolved into essential tools in social media, and mainly on Twitter. They aid in content categorization, simplifying post organization, and enabling users to discover content aligned with their interests. Moreover, hashtags enhance discoverability, allowing users to search for specific topics, trends, or discussions. They play a vital role in event promotion, research, and activism, empowering users to rally support and raise awareness about social issues or global events.

It's essential to acknowledge that this approach may result in the exclusion of some tweets discussing migration. However, considering the constraints of time and resources, filtering by hashtags was deemed the most pragmatic approach to achieving the research objectives while working with a more focused subset of the dataset. Furthermore, this approach streamlined the translation process for future steps, as it is a necessary step when using VADER. This is due to the current unavailability of software capable of displaying sentiment analysis results in Portuguese, a

point that will be elaborated on in subsequent sections. The list of selected hashtags by year and month, and their respective frequency:

Table 1. Selected hashtags by year and month. Source: Author.

YEAR	MONTH	HASHTAGS AND FREQUENCY
2015	January	acnur (16), brasil (13), imigrantes (12), imigração (6), refugiados (4), eumigrante (4), expressentry (3), sírios (2), imigrante (2), imigracao (2), somostodosmigrantes (2).
	February	refugiados (34), eumigrante (20), onu (12), acnur (11), migrante (11), imigração (10), imigrantes (9), opentosyria (9), maismédicos (5), migración (4), imigrante (3).
	March	refugiados (75), acnur (38), syria (25), onu (23), imigração (22), imigrantes (21), imigracao (8), imigrante (7), eupertenço (7), frontex (5), ibelong (4).
	April	refugiados (69), imigrantes (55), imigração (23), acnur (19), onu (13), migration (10), imigrante (6), xenophobia (6), migrants (5), migrante (5), migrantes (4).
	May	refugiados (46), imigrantes (45), onu (19), imigração (10), imigrante (8), opentosyria (6), acnur (5).
	June	refugiados (72), diamundialdorefugiado (26), onu (23), imigração (18), imigrantes (18), refugiado (14), xenofobianão (13), todossomosrefugiados (12), migrantes (5), diamundialdosrefugiados (5), worldrefugeeday (4), acnur (4), refugees (4).
	July	refugiados (88), síria (29), onu (26), imigração (26), imigrantes (17), refugiado (13), migrante (7).
	August	refugiados (70), líbia (40), diamundialhumanitário (35), imigração (26), imigrantes (26), sharehumanity (25), onu (14), refugiado (8), haitianos (7), sírios (7), migrantes (5), compartilhehumanidade (5), migrante (4).
	September	compartilhehumanidade (220), refugiados (163), refugeeswelcome (78), refugeecrisis (13), imigrantes (11), refugees (9), refugiado (7), acnur (6), refugeescrisis (5), welcomerefugees (5), somostodosimigrantes (4), imigrante (4).
	October	refugiados (186), imigração (27), imigrantes (22), refugeeswelcome (14), compartilhehumanidade (8), refugees (7), crisederefugiados (7), imigracao (6), migración (6), refugiado (5), migrante (5), migrantes (5).
	November	refugiados (165), refugeecrisis (33), refugeeswelcome (25), imigrantes (20), openeuborders (13), imigraçãogerainovação (13), fechaasfronteirasportugal (12), ibelong (8), imigração (7), openborders (6), crisederefugiados (5), estamosxlosrefugiados (5), haitianosficamnobrasil (5), migrantes (4), somostodosimigrantes (4).
	December	refugiados (200), imigrantes (27), refugees (17), refugiado (17), compartilhehumanidade (15), elrefugiado (10), soumigrante (8), somosmigrantes (8), derechoshumanos (7), imigrante (6), acnur (6).
2016	January	refugiados (163), imigrantes (16), migrantes (9), naufrágio (9), acnur (7), refugees (7), imigração (6), migrante (6), deportaci (6), refugeeswelcome (5), crisedosrefugiados (5), refugeecrisis (5).
	February	refugiados (164), refugees (21), imigrante (18), migrantes (11), imigrantes (10), migração (9), welcomerefugees (9), safepassage (8), imigração (7), acnur (7), corridoiumanitari (7), expressentry (6), sirios (6), refugeecrisis (5).
	March	refugiados (372), refugeeswelcome (46), operativomigrante (33), imigrantes (17), withsyria (16), refugioderechoshumanos (12), refugees (10), racismo (10), prouracisme (10), ueturkey (9), uemata (8), migrações (7), imigração (7), refugeecrisis (7), ueturquia (7), crisis (6), refugiado (6), migrantes (5).

	April	refugiados (186), imigrantes (24), refugiado (21), refugiats (19), yorefugio (16), overthefortress (13), migrante (11), imigração (10), refugeeswelcome (9), refugeecameras (9), acnur (8), imigrante (7), openeuborders (7), refugees (6), migrantes (6), crisisrefugiados (6), minutorefugiados (5).
	May	refugiados (235), cambioclimático (44), imigrantes (41), refugeeswelcome (27), acnur (22), jantarhumanitario (13), unhcr (13), eusourefugio (13), imigracao (13), jeansforrefugees (8), refugiadosnobrasil (7), refugees (7), migrações (7), imigração (5), migrante (5).
	June	refugiados (127), comosrefugiados (46), imigrantes (43), acnur (34), teamrefugees (29), withrefugees (26), imigração (15), imigracao (12), worldrefugeeday (12), diamundialdorefugiado (11), sírios (6), nobordersnonations (6), iamanimmigrant (5), refugees (5), diamundialdosrefugiados (5), mediterrâneo (4).
	July	refugiados (144), comosrefugiados (63), acnur (32), teamrefugees (30), imigrantes (21), refugiado (20), imigração (15), opentheborders (13), refúgio (10), refugiats (9), imigracao (8), refugiadas (6), migrante (6), prayforsyria (6), refugiada (5), refugees (5), migrantes (5), moçambicanos (5).
	August	comosrefugiados (400), teamrefugees (285), refugiados (106), refugeeolympicteam (45), imigrantes (27), imigração (25), teamrefugee (19), imigracao (13), migrantes (6), atletasrefugiados (6), crisisderefugiados (6), refugiado (5).
	September	comosrefugiados (164), refugiados (117), imigração (27), teamrefugess (24), imigrantes (23), acnur (22), imigracao (17), onu (16), withrefugees (12), un4refugeesmigrants (11), migrantes (9), teamrefugees (5).
	October	refugiados (119), imigrantes (35), imigração (29), imigracao (24), migrantes (20), welcomerefugees (15), euacolho (14), acnur (12), migrante (12), caritasrj (11), refugees (6), un4refugeesmigrants (5), inmigración (5)
	November	refugiados (135), comosrefugiados (43), imigração (34), imigrantes (27), imigracao (23), withrefugees (18), migrantes (11), refugiado (9), migrante (8), imigrante (6), immigration (6), rifugiato (5), migrarédireito (5).
	December	refugiados (132), euaceitorefugiados (55), nãoaceitamosrefugiados (51), meuamigorefugiado (48), imigração (47), imigracao (29), imigrantes (21), naoaceitamosrefugiados (20), migrantes (18), refugiado (12), conlosrefugiados (10), migramundo (9), comosrefugiados (6), migrarédireito (6), imigracaoitaliana (6), migrante (6), prayforaleppo (5), refugees (4), acolhidas (4), imigracaononordeste (4), migrantsday (4).
2017	January	refugiados (113), imigrantes (46), imigração (36), refugees (19), migrante (17), imigracao (16), imigracaoitaliana (7), migrantes (7), euaceitorefugiados (6), imigracaononordeste (5), muslimban (5), comosrefugiados (4), refugeeswelcome (4), acnur (4).
	February	refugiados (59), volemacollir (59), imigração (34), imigrantes (23), imigracao (18), refugees (8), refugiats (7), vistos (6), acnur (5), refugeeswelcome (5).
	March	refugiados (91), imigração (47), imigrantes (24), migramundo (14), onu (14), migranti (13), imigracao (13), amorsemfronteiras (9), refugiadas (7), acnur (7), direitoshumanos (6), syriacrisis (5), withrefugees (5), migrantes (4), refugees (4).
	April	vetatemer (334), leidamigraçãonão (49), refugiados (42), migracaonao (39), imigração (38), rejeitemigracao (26), aloysiomigracaoagoranao (25), leidemigracaonao (25), imigrantes (22), vistos (13), acnur (13), rejeiteleidemigracao (13), vetamigracaotemer (12), migramundo (11), vetatemerleiimigracao (5), imigracao (4), unhcr (4), aloysiomigraçãoagoranao (4), migraçãonão (4), refugiado (3), refugees (3), migrações (3), leidemigraçaonao (3), migrarédireito (3).

	May	vetatemer (260), migracaoveta (73), refugiados (45), imigrantes (42), terroristaleimigracao (41), imigração (33), vetatemermigração (28), leidemigracaonao (21), iwelcome (9), euacolho (9), migrações (9), acnur (7), comosrefugiados (7), vetaimigracaotemer (6), migramundo (6), naoaleidemigracao (4).
	June	refugiados (102), imigração (65), imigrantes (61), refugeeday (26), comosrefugiados (20), diamundialdorefugiado (17), acnur (12), withrefugees (11), imigracao (8), díamundialdorefuxiado (7).
	July	imigrantes (97), refugiados (87), exodus (56), imigração (47), imigrante (6), acnur (6), migramundo (5), migrarédireito (5).
	August	imigrantes (61), refugiados (50), acnur (33), imigração (29), racismo (12), exodus (10), refugees (8), migramundo (7), migrantes (5), imigracao (4)
	September	imigrantes (71), refugiados (56), onu (28), imigração (22), acnur (18), refugiadosvavel (13), refugees (8), migrantes (6), migramundo (6).
	October	acnur (109), imigrantes (85), refugiados (52), imigração (32), migramundo (12), migrante (6), migrantes (6), refugees (6), refugiado (6).
	November	imigrantes (63), refugiados (46), imigração (35), imigrante (20), acnur (7), migramundo (6), humanflow (5), migrante (5), migranti (5), migrantes (4).
	December	refugiados (257), diadelosderechoshumanos (204), imigrantes (65), imigração (22), migrantes (21), acnur (16), conlosrefugiados (9), acogida (5), refugees (4), corredoreshumanitarios (4), imigrante (3), direitoshumanos (3).
2018	January	refugiados (74), imigrantes (65), imigração (24), acnur (19), migrantes (17), refugiado (13), migrante (11), refugiadoslgtb (6), accoglienza (5), rifugiato (5), refugees (5)
	February	refugiados (61), imigrantes (41), imigração (8), migrantes (8), migramundo (4), migrante (3), comosrefugiados (3), acnur (3), pactoglobalsobrerefugiados (3).
	March	refugiados (80), imigrantes (60), migrantes (47), acnur (8), semfronteiras (8), operaçãoacolhida (7), humanrights (5), imigrante (5), imigração (4), artemigrante (4), crisehumanitária (3), direitoshumanos (3), derechoshumanos (3), fraternidadesemfronteiras (3).
	April	refugiados (58), freeopenarms (49), operaçãoacolhida (44), imigrantes (32), imigração (20), direitoshumanos (10), migramundo (5), unhcr (4), acnur (4), imigrantebrasileiro (3), migrantes (3), refúgioemnúmeros (3).
	May	imigrantes (51), refugiados (40), imigração (19), imigrante (9), respeitaosrefugiados (9), acnur (7), operacaoacolhida (6), migrantes (5), venezuelanos (5), migraciones (4), fronterasur (4), migramundo (3), direitoshumanos (3),migrante (3).
	June	refugiados (51), imigrantes (44), imigração (22), refugeeswelcome (11), acnur (11), migrantes (9), direitoshumanos (8), meuamigorefugiadocopa (7), imigrantesdevolta (6), refugeeswelcomegfbpa (6), xenofobia (5), racismo (5), stopracismo (5), bienvenidos (5), worldrefugeeday2018 (5), refugiats (5), comosrefugiados (4).
	July	refugiados (29), imigração (22), imigrantes (10), criminiimmigrati (8), refugeeswelcomegfbpa (6), acnur (5), oim (5), migração (5), migrantes (5).
	August	refugiados (32), imigrantes (20), criminiimmigrati (17), imigração (14), acnur (9), venezuelanos (7), venezolanos (4).
	September	refugiados (52), imigração (16), xenofobia (9), imigrantes (9), migrantes (8), venezuelanos (6), direitoshumanos (4), criminiimmigrati (3), refugiado (3), acnur (3), deportação (3).
	October	refugiados (54), caravanamigrante (20), acnur (18), imigrantes (16), imigração (11), migrantes (6), venezuelanos (6), refugiado (5), direitoshumanos (4), migrante (3), imigrante (3).

	November	acnur (30), imigração (19), imigrantes (19), venezuela (19), refugiados (18), caravanamigrante (7), imigrante (6), migrantes (4), criminiimmigrati (4), fronteira (4), imigracao (3), venezuelanos (3), orepelosrefugiados (2), direitoshumanos (2)
	December	refugiados (19), imigrantes (16), direitoshumanos (10), imigração (9), acnur (7), sonbienvenidos (5), migrantes (4), migrações (3), pactoglobalimigracao (2), imigracao (2), refugee (2), diainternacionaldosmigrantes (2), reassentados (2)
2019	January	operaçãoacolhida (186), imigrantes (28), refugiados (23), acnur (16), imigração (7), refuxiados (6), migrantes (5), venezuelanos (5), criminiimmigrati (3), caravanamigrante (3), direitoshumanos (2), imigracao (2), conlosrefugiados (2), refugees (2), venezuelanas (2), refugiadas (2).
	February	refugiado (28), refugiados (27), imigrantes (21), acnur (18), operaçãoacolhida (17), imigração (11), migrantes (6), migration (5), migrante (5), direitoshumanos (4), imigracao (4), refuxiados (3), humanity (2), refugees (2).
	March	operaçãoacolhida (43), refugiados (35), imigrantes (10), migrantes (9), acnur (7), comosrefugiados (6), imigração (5), venezuelanos (4), migrante (3), imigracao (3), standwithiraqirefugees (3), migração (2), formigration (2), refuxiados (2), refugees (2), crisenavenezuela (2), imigraçãoitaliana (2)
	April	operaçãoacolhida (56), acnur (39), refugiados (25), imigrantes (17), migrantes (12), resettlement4iranian (12), imigração (10), imigrante (5), migrante (5), migração (4), cáritasbrasileira (3), eumigrante (3), muros (2)
	May	refugiados (45), xenofobianão (39), imigrantes (31), operaçãoacolhida (22), migrantes (14), imigração (12), acnur (12), migrante (8), refugiado (8), imigrante (6), xenofobia (4), visto (4), direitoshumanos (4), criminiimmigrati (3).
	June	refugiados (47), imigrantes (31), migrantes (23), imigração (21), comosrefugiados (16), operaçãoacolhida (14), diamundialdorefugiado (7), operacaoacolhida (6), visto (6), semfronteiras (6), imigrante (6), festadoimigrante (5), acnur (5), imigracao (5), venezuelanos (5), refugees (5), diadorefugiado (5), migrante (4), direitoshumanos (3), estrangeiros (3), semananacionaldosrefugiados (3).
	July	refugiados (41), imigrantes (26), imigração (18), migrantes (17), operacaoacolhida (13), migrante (9), refugiado (8), cáritas (4), venezuelanos (3), acnur (3), seawatch (2), migração (2).
	August	refugiados (42), acnur (16), imigrantes (13), migrantes (13), operaçãoacolhida (12), imigração (9), diamundialhumanitario (5), migrante (4), integração (2), xenofobia (2), imigracao (2).
	September	refugiados (63), migrantes (14), imigrantes (11), imigração (11), operaçãoacolhida (11), acnur (7), naosetrataapenasdemigrantes (6), refuxiadas (5), imigrante (4), venezuelanos (3), refugees (3), opacolhida (3), refugee (3), refugeeswelcome (2), migrante (2), xenofobia (2), aldeiadosimigrantes (2)
	October	operaçãoacolhida (76), refugiados (19), opacolhida (12), imigrantes (9), migrantes (8), operacaoacolhida (7), imigração (6), migrante (6), refugiado (5), acnur (4), caritas (4), refugiadoscristaos (3), oim (2), xenófobas (2), discriminação (2), imigrante (2).
	November	refugiados (24), assistência (17), imigração (16), imigrantes (9), opacolhida (9), operaçãoacolhida (8), acnur (8), imigracao (7), migrante (6), migrantes (5), refugeeforum (3), direitos (3), diversidadcultural (3).
	December	refugiados (23), venezuela (22), operaçãoacolhida (11), refugeeforum (10), acnur (8), imigrantes (8), migrantes (7), migrante (7), imigracao (3), onu (3), imigração (2).

It is imperative to underscore the intricate nuances that become apparent when examining the usage of specific hashtags within defined timeframes, particularly in response to significant events,

and their monthly frequency. Remarkably, a pronounced consistency emerges in the utilization of particular hashtags in addressing this topic, including '*refugiados*' (refugees), 'refugees', '*imigrantes*' (immigrants), '*imigração*' (immigration), and '*migrante*' (migrant). These hashtags shed light on the evolving nature of discussions surrounding refugees and immigrants over the years. Moreover, the prevalence of these neutral hashtags signifies a burgeoning collective awareness concerning migration and refugee-related issues, with a consistent upward trajectory over the observed period.

Hashtags linked to specific countries or nationalities, such as Syria, Libya, and Haiti, indicate a global consciousness of events worldwide. The appearance of hashtags like 'opentosyria,' '*compartilhehumanidade*' (share humanity), '*refugeeswelcome*', '*diamundialdorefugiado*' (World Refugee Day), '*comosrefugiados*'(with refugees), '*migrarédireito*' (migration is a right), and '*todossomosrefugiados*' (we are all refugees) denotes a more positive and welcoming sentiment toward refugees, signifying a growing global willingness to embrace and support those seeking refuge. Additionally, hashtags related to specific crises, such as 'prayforsyria' and 'overthefortress', demonstrate how social media serves as a platform to respond to and raise awareness about urgent humanitarian situations.

The emergence of the Refugee Olympic Team at the Rio Olympics in 2016, symbolized by 'teamrefugees', showcased refugees' potential to excel and inspire on the global stage. Hashtags like 'openeuborders' and 'opentheborders' highlighted ongoing debates about immigration and border policies in Europe and beyond. The presence of hashtags like 'onu' (United Nations) and 'un4refugeesmigrants' indicated the involvement of international organizations in addressing refugee and migrant issues.

In 2016, a diverse range of perspectives on migration became evident. Hashtags like '*naoaceitamosrefugiados*' (we don't accept refugees) highlighted opposing views on refugee acceptance and resettlement, underscoring the complexity of the migration debate and the need for constructive dialogue.

In subsequent years, despite the evolution of supportive hashtags indicating a positive sentiment and a welcoming attitude, there was also growing resistance, particularly concerning the approval of the New Migration Law, as seen in hashtags such as '*vetatemer*' (veto Temer), '*naoaleidemigracao*' (no to the immigration law), '*rejeitemigracao*' (reject immigration), and '*vetaimigracaotemer*' (veto immigration Temer).

The hashtags '*operaçãoacolhida*' and '*opacolhida*' (Operation Welcome) gained prominence in 2018 and 2019, signifying Brazil's efforts to manage the influx of Venezuelan migrants, also emphasizing the society engagement with the topic. Furthermore, there was a noticeable increase in hashtags related to immigration policies, such as '*pactoglobalsobrerefugiados*' (global pact on

refugees), indicating a shift toward policy-focused discussions. Additionally, some overall rejection towards migration also appeared, indicating some polarisation of opinions and negative feelings.

Analysing hashtags can provide some valuable insights into the evolving conversations and sentiments surrounding migration and refugees. It is an indicator of the collective awareness, changing attitudes, and even policy implications over time.

Following the filtering of the dataset with only the tweets that contained the hashtags above from the list, we achieved a more focused dataset containing 9.114 tweets. After that it was also employed the filter of the column 'lang' mentioned before, resulting in a data frame of 6.947 tweets.

The analysis of tweet frequency by month and year provides valuable insights into the evolution of the topic over specific time periods, offering a temporal perspective on its popularity. Notably, the topic exhibited robust engagement during the initial three years under examination, gradually waning in the final two years. The graph allows for the identification of particular months that coincide with events influencing the topic's popularity, shedding light on the dynamic interplay between events and public interest.

By combining this frequency analysis with sentiment analysis by month, we can construct a more comprehensive understanding of the potential correlations between national and international events during this period, the frequency of tweets, and the nuances of sentiment expressed. This holistic approach promises to unveil intricate connections and patterns, enriching our comprehension of the societal responses to the events of the time.
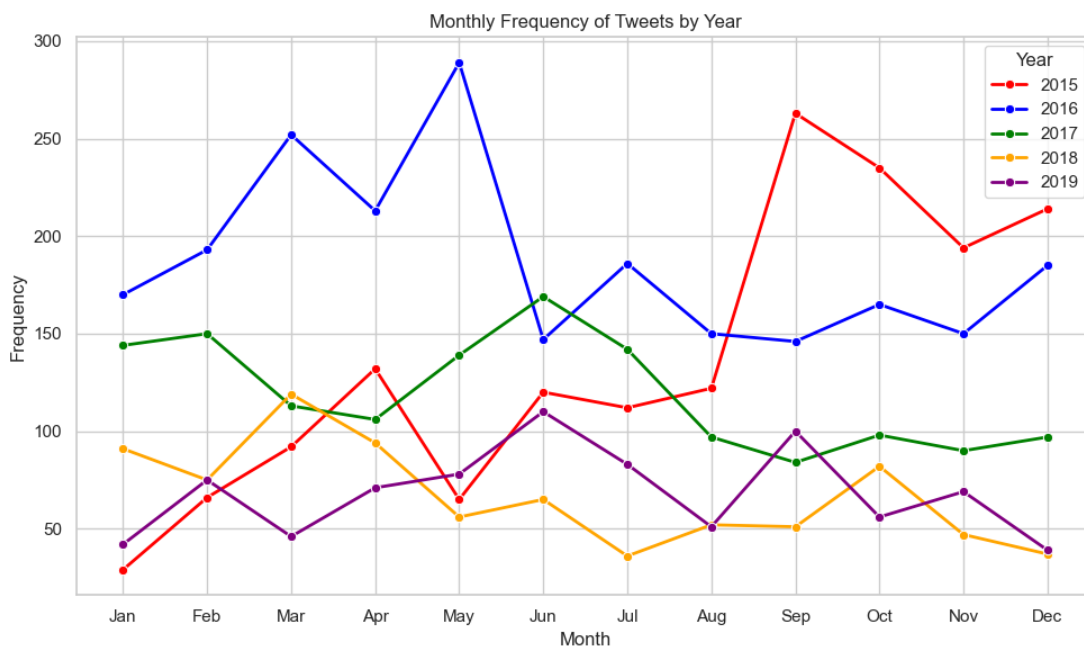
Figure 14. Monthly frequency of tweets by year. Source: Author.

1.2. Text Processing

Having achieved a more focused dataset, the following steps are related to cleaning the column 'text' of the dataset that contain the text of the tweets. To primarily analyse the content of those tweets and get the most used words and terms to refer to migrants and refugees, and then do the sentiment analysis, we proceeded to the text processing. In the context of data analysis and text mining, it is often necessary to preprocess textual data to remove irrelevant or sensitive information, as mentioned in the methodology chapter. The first and a common preprocessing step involves the identification and removal of URLs (Uniform Resource Locators) from text data. URLs are web addresses that may not provide meaningful information in certain analytical tasks, such as the one we aim to accomplish.

```
# Define a regular expression pattern to match URLs
url_pattern = r'http[s]?://(?:[a-zA-Z]|[0-9]|[$-_@.&+]|[!*\\(\\),]|(?:%[0-9a-fA-F][0-9a-fA-F]))+'

# Remove URLs from the 'text' column using regular expressions
df_filtered['text'] = df_filtered['text'].str.replace(url_pattern, '', regex=True).copy()
```

Figure 15. Code for Removing URLs. Source: Author.

In this part, the code accomplishes the task of identifying and removing URLs from a data frame (df_filtered) using a regular expression pattern ('url_pattern') that is defined to match URLs present in the text data. This pattern is designed to recognize the presence of "http://" or "https://", representing web protocols, with the '[s]?://' part allowing for the optional 's' in "https". The subsequent '(?: ...)+' non-capturing group captures various URL components. Within this group, '[a-zA-Z]' matches alphabetic characters, both uppercase and lowercase, '[0-9]' captures single digits, and '[$-_@.&+]' accounts for common special characters like the dollar sign, hyphen, underscore, at symbol, period, ampersand, and plus sign. Additionally, '[!*\\(\\),]' matches other special characters such as exclamation marks, asterisks, parentheses, and commas, while '(?:%[0-9a-fA-F][0-9a-fA-F])' accounts for percent-encoded characters commonly found in URLs. The '+' outside the non-capturing group allows for the repetition of these matched components, effectively identifying URLs within text data with a comprehensive approach.

The second line refers to the URL removal of the identified URLs from the 'text' column of the data frame using the '.str.replace()' method with the specified 'url_pattern'. This effectively replaces the matched URLs with an empty string ('') and cleans the text data. The result is then assigned back to the 'text' column.

```
# List of specific words to remove
specific_words_to_remove = ['RT', 'q' , 'vc', 'htt...', 'http...', 'http', 'https', 'htt', 'htt', 'p', 'pq', 'd', 'in',
                            'tá', 'km', 'n', 'aí', 'ai', 'R', 'h', 'etc', 'la', 'lá', 'at', 'e']

# Regular expression pattern to match the specific words
pattern = r'\b(?:' + '|'.join(specific_words_to_remove) + r')\b'

# Remove the specific words from the 'text' column using regular expressions
df_filtered['text'] = df_filtered['text'].str.replace(pattern, '', flags=re.IGNORECASE, regex=True).copy()
```

Figure 16. Code for removing specific words. Source: Author.

The next step is to remove some specific words from the text. The code begins by defining a list of those specific words to be removed, encompassing terms such as "RT," "http," "https," and various other frequently encountered but non-informative words. Subsequently, it was also used a regular expression pattern to leverage the word boundary \b to ensure that only complete words are matched for removal. The use of the flags '=re.IGNORECASE' option makes the pattern matching case-insensitive, facilitating the removal of specific words regardless of their capitalization. Finally, the code eliminates these words from the 'text' column of a data frame, enhancing the quality and relevance of textual data.

```
# Define a regular expression pattern to match mentions (usernames)
mention_pattern = r'@[\w\d]+'

# Remove mentions from the 'text' column using regular expressions
df_filtered['text'] = df_filtered['text'].str.replace(mention_pattern, '', flags=re.IGNORECASE, regex=True).copy()
```

Figure 17. Code for removing mentions/usernames (@). Source: Author.

In the realm of textual data, mentions of usernames, typically encountered in social media and user-generated content. This next code begins by defining a regular expression pattern (mention_pattern) crafted to identify mentions, which conventionally commence with the '@' symbol followed by alphanumeric characters and digits. The pattern consists of two elements. Firstly, the "@" symbol is included to match the universally recognized symbol that signifies the beginning of a mention. Secondly, the [\w\d]+ component acts as a character class, specifying a set of valid characters for usernames. Specifically, \w covers all word characters, including letters in both uppercase and lowercase, digits, and underscores, ensuring the inclusivity of various username formats. The\d further extends this inclusivity by allowing for the presence of digits within usernames. The + quantifier ensures that the regular expression captures the entire username, rather than just one character. The subsequent step employs the .str.replace() method to systematically erase these mentions from the 'text' column of the data frame (df_filtered). The inclusion of the flags =re.IGNORECASE argument ensures a case-insensitive matching process, guaranteeing the removal of mentions regardless of their capitalization. The resultant text is then seamlessly integrated back into the 'text' column.

```python
# Tokenize the text in the 'text' column using NLTK
df_filtered['tokens'] = df_filtered['text'].apply(word_tokenize)
```

Figure 18. Code for tokenization. Source: Author.

The tokenization is a fundamental process in NLP and text analysis, where the objective is to break down a continuous text into discrete units called tokens. It leverages the '.apply()' method to apply the NLTK's 'word_tokenize' function to each element in the 'text' column, thus tokenizing the text into individual words or tokens (e.g., "I am here" → ['I', 'am', 'here']). Consequently, a new column named 'tokens' was created within the data frame, storing the tokenized representation of the original text.

The next step is to remove stop words. While they are essential in the language, they often do not carry significant meaning in various text analysis tasks. The NLTK library automatically offers a list of stop words in Portuguese that is shared in the list below. However, the author found necessary to complement the list with an additional list of stop words that can be observe in the following figure along with the code.

```
['a', 'à', 'ao', 'aos', 'aquela', 'aquelas', 'aquele', 'aqueles', 'aquilo', 'as', 'às', 'até', 'com', 'como', 'da', 'das', 'd
e', 'dela', 'delas', 'dele', 'deles', 'depois', 'do', 'dos', 'e', 'é', 'ela', 'elas', 'ele', 'eles', 'em', 'entre', 'era', 'era
m', 'éramos', 'essa', 'essas', 'esse', 'esses', 'esta', 'está', 'estamos', 'estão', 'estar', 'estas', 'estava', 'estavam', 'est
ávamos', 'este', 'esteja', 'estejam', 'estejamos', 'estes', 'esteve', 'estive', 'estivemos', 'estiver', 'estivera', 'estivera
m', 'estivéramos', 'estiverem', 'estivermos', 'estivesse', 'estivessem', 'estivéssemos', 'estou', 'eu', 'foi', 'fomos', 'for',
'fora', 'foram', 'fôramos', 'forem', 'formos', 'fosse', 'fossem', 'fôssemos', 'fui', 'há', 'haja', 'hajam', 'hajamos', 'hão',
'havemos', 'haver', 'hei', 'houve', 'houvemos', 'houver', 'houvera', 'houverá', 'houveram', 'houvéramos', 'houverão', 'houvere
i', 'houverem', 'houveremos', 'houveria', 'houveriam', 'houveríamos', 'houvermos', 'houvesse', 'houvessem', 'houvéssemos', 'iss
o', 'isto', 'já', 'lhe', 'lhes', 'mais', 'mas', 'me', 'mesmo', 'meu', 'meus', 'minha', 'minhas', 'muito', 'na', 'não', 'nas',
'nem', 'no', 'nos', 'nós', 'nossa', 'nossas', 'nosso', 'nossos', 'num', 'numa', 'o', 'os', 'ou', 'para', 'pela', 'pelas', 'pel
o', 'pelos', 'por', 'qual', 'quando', 'que', 'quem', 'são', 'se', 'seja', 'sejam', 'sejamos', 'sem', 'ser', 'será', 'serão', 's
erei', 'seremos', 'seria', 'seriam', 'seríamos', 'seu', 'seus', 'só', 'somos', 'sou', 'sua', 'suas', 'também', 'te', 'tem', 'té
m', 'temos', 'tenha', 'tenham', 'tenhamos', 'tenho', 'terá', 'terão', 'terei', 'teremos', 'teria', 'teriam', 'teríamos', 'teu',
'teus', 'teve', 'tinha', 'tinham', 'tínhamos', 'tive', 'tivemos', 'tiver', 'tivera', 'tiveram', 'tivéramos', 'tiverem', 'tiverm
os', 'tivesse', 'tivessem', 'tivéssemos', 'tu', 'tua', 'tuas', 'um', 'uma', 'você', 'vocês', 'vos']
```

Figure 19. List of Portuguese stop words provided by NLTK library. Source: Author.

The code begins by defining an extended list of Portuguese stop words, and then combining the two sets of stop words: the one commonly used NLTK stop words and the additional set. Subsequently, the code applies this combined list of stop words to the tokenized text data within the data frame (df_filtered), employing a lambda function to iteratively filter the 'tokens' column, creating a new column named 'filtered_tokens' that stores the tokenized text with the stop words removed. This process enhances the quality and relevance of the text data for further analysis, as it eliminates non-essential words that may introduce noise or disrupt analytical tasks.

```
additional_stopwords = ['a', 'à', 'agora', 'ainda', 'além', 'algo', 'algumas', 'alguns', 'ali', 'ano', 'anos', 'antes', 'ao',
    'aos', 'apenas', 'apoio', 'após', 'aquela', 'aquelas','aqui', 'aquele', 'aqueles', 'aquilo', 'área', 'as', 'às', 'assim',
    'até', 'atrás', 'através', 'baixo', 'bastante', 'bem', 'boa', 'boas', 'bom', 'bons', 'breve', 'cá', 'cada', 'catorze',
    'cedo', 'cento','certamente', 'certeza', 'cinco', 'coisa', 'coisas', 'com', 'como', 'conselho', 'contra', 'contudo',
    'corrente','cuja', 'cujas', 'cujo', 'cujos', 'da', 'dá', 'dão', 'daquela', 'daquelas', 'daquele', 'daqueles', 'daqui',
    'daquilo','das', 'de', 'debaixo', 'dela', 'delas', 'dele', 'deles', 'dêm', 'demais', 'dentro', 'depois', 'desde', 'dessa',
    'dessas', 'desse', 'desses', 'desta', 'destas', 'deste', 'destes', 'deve', 'deverá', 'dez', 'dezenove', 'dezesseis',
    'dezassete', 'dezoito', 'dia', 'diante', 'diz', 'dizem', 'do', 'dona', 'donas', 'dono', 'donos', 'dos', 'doze', 'duas',
    'dúvida', 'e', 'ela', 'elas', 'ele', 'eles', 'em', 'embora', 'entre', 'então', 'entanto', 'era', 'eram', 'éramos',
    'é', 'essa', 'essas', 'esse', 'esses', 'esta', 'estamos', 'está', 'estão', 'estar', 'estará', 'estas', 'estás', 'estava',
    'estavam', 'este', 'estes', 'esteve', 'estive', 'estivemos', 'estiveram', 'estiveste', 'estivestes', 'estou', 'eu',
    'exemplo', 'falta', 'fará', 'favor', 'faz', 'fazeis', 'fazem', 'fazemos', 'fazer', 'fazes', 'fez', 'fim', 'final', 'foi',
    'fomos', 'for', 'fora', 'foram', 'formos', 'fosse', 'fossem', 'foste', 'fostes', 'fui', 'geral', 'grande', 'grandes',
    'grupo', 'hoje', 'hora', 'horas', 'ir', 'irá', 'isso', 'isto', 'já', 'lá', 'lado', 'ligado', 'local', 'logo', 'longe',
    'lugar', 'maior', 'maioria', 'maiorias', 'mais', 'mal', 'mas', 'máximo', 'me', 'meio', 'menor', 'menos', 'mês',
    'meses', 'meu', 'meus', 'mil', 'minha', 'minhas', 'momento', 'muito', 'muitos', 'na', 'nada', 'naquela', 'naquelas',
    'naquele', 'naqueles', 'nas', 'nem', 'nenhuma', 'nessa', 'nessas', 'nesse', 'nesses', 'nesta', 'nestas', 'neste', 'nestes',
    'ninguém', 'no', 'nos', 'nós', 'nossa', 'nossas', 'nosso', 'nossos', 'nova', 'novas', 'nove', 'novo', 'novos', 'num',
    'numa', 'número', 'nunca', 'nuns', 'o', 'obra', 'obrigada', 'obrigado', 'oitava', 'oitavo', 'oito', 'onde', 'ontem',
    'onze', 'os', 'ou', 'outra', 'outras', 'outro', 'outros', 'para', 'pra','parece', 'parte', 'partir', 'paucas', 'pela','pelas'
    'pelo', 'pelos', 'perto', 'pode', 'pôde', 'podem', 'poderá', 'podia', 'pois', 'ponto', 'pontos', 'por', 'porque', 'porquê',
    'pouca', 'pouco', 'poucos', 'primeira', 'primeiras', 'primeiro', 'primeiros', 'própria', 'próprias', 'próprio', 'próprios',
    'quáis', 'qual', 'qualquer', 'quando', 'quanto', 'quarta', 'quarto', 'quatro', 'que', 'quem', 'quer', 'quê', 'quinta',
    'quinto', 'quinze', 'relação', 'sabe', 'são', 'se', 'segunda', 'segundo', 'sei', 'seis', 'seja', 'sejam', 'sempre', 'sendo',
    'ser', 'será', 'seu', 'seus', 'sexta', 'sexto', 'sim', 'sistema', 'sob', 'sobre', 'sois', 'somos', 'sou', 'sua', 'suas',
    'tal', 'talvez', 'também', 'tanta', 'tantas', 'tanto', 'tantos', 'te', 'tem', 'têm', 'temos', 'tendes', 'tenho', 'tens',
    'ter', 'terá', 'terão', 'terceira', 'terceiro', 'teu', 'teus', 'teve', 'ti', 'tido', 'tinha', 'tinham', 'tive', 'tivemos',
    'tiveram', 'tiveste', 'tivestes', 'toda', 'todas', 'todo', 'todos', 'trabalho', 'três', 'treze', 'tu', 'tua', 'tuas',
    'tudo', 'última', 'últimas', 'último', 'últimos', 'um', 'uma', 'umas', 'uns', 'usa', 'usar', 'vai', 'vais', 'valor',
    'veja', 'vem', 'vens', 'ver', 'verdade', 'verdadeiro', 'vez', 'vezes', 'viagem', 'vindo', 'vinte', 'você', 'vocês',
    'vos', 'vós', 'vossa', 'vossas', 'vosso', 'vossos', 'zero', 'via']

# Combine both lists of stopwords
portuguese_stopwords = stop + additional_stopwords

df_filtered['filtered_tokens'] =
    df_filtered['tokens'].apply(lambda tokens:[word for word in tokens if word.lower() not in portuguese_stopwords])
```

Figure 20. Code for removing stop words code. Source: Author.

The next code begins with the reassembly of individual tokens into phrases or sentences. The custom function 'join_tokens_into_phrase' combines these tokens, restoring the context and readability of the text. This initial step is akin to reassembling the pieces of a puzzle, making the data more comprehensible and context rich.

```
def join_tokens_into_phrase(tokens):
    return ' '.join(tokens)

# Define a pattern to remove punctuation
punctuation_pattern = r'[^\w\s]'

# Define a list of specific punctuation signs to remove
specific_punctuation = [':', ';', '.', ',', '?', '&']

# Remove the specified punctuation from filtered tokens
def remove_specific_punctuation(tokens):
    return [re.sub(punctuation_pattern, '', word) for word in tokens if word not in specific_punctuation]

df_filtered['cleaned_tokens'] = df_filtered['filtered_tokens'].apply(remove_specific_punctuation)

# Combine cleaned tokens into a list
df_filtered['cleaned_text'] = df_filtered['cleaned_tokens'].apply(join_tokens_into_phrase)

# Remove hashtags from the 'cleaned_text' column
df_filtered['cleaned_text'] = df_filtered['cleaned_text'].str.replace(r'#\w+', '', regex=True)
```

Figure 21. Code for joining the text, removing punctuation, removing hashtags. Source: Author.

Next, the code addresses the issue of punctuation. A regular expression pattern, 'punctuation_pattern', is introduced to discern and excise punctuation marks that often clutter the text without adding significant meaning. A curated list, 'specific_punctuation', allows for the selective removal of specific punctuation symbols. The cleaned tokens are then separated in a new column ('cleaned_tokens') within the DataFrame (df_filtered). Subsequently, the code orchestrates the amalgamation of the cleaned tokens into textual form, and the result is deployed in the 'cleaned_text' column. Finally, the code removes hashtags from the 'cleaned_text' column, by using a regular expression pattern (r'#\w+').

The last step before translation of the processed text is reorganizing the data frame to ensure that it is structured to avoid any issues on the following procedures.

```python
def reorganize_dataframe(df):
    # Reset the index
    df.reset_index(drop=True, inplace=True)

    # Rearrange columns in the desired order
    new_column_order = ['id', 'text', 'created_at', 'author_id', 'lang', 'in_reply_to_user_id', 'tokens',
                        'filtered_tokens', 'cleaned_tokens', 'cleaned_text']
    df = df[new_column_order]

    return df

# Call the reorganize_dataframe function on the pre-processed dataframe
df_filtered_org = reorganize_dataframe(df_filtered)
```

Figure 22. Code for reorganizing the DataFrame. Source: Author.

The process commences by resetting the index of the data frame, denoted by the variable 'df'. The 'reset_index' function is called with 'drop=True' to reset the index while discarding the previous index, and 'inplace=True' is used to modify the data frame in place, ensuring that the data is sequential and indexed correctly. Then it rearranges the columns to align with the desired order. The variable 'new_column_order' defines the precise order of columns and the modified data frame, with its index reset and columns reordered, is returned from the 'reorganize_dataframe' function.

Following data cleaning and rigorous filtering procedures, the analysis reveals the prevailing terms commonly employed in reference to migrants and refugees throughout this specified timeframe. The terms depicted in the word cloud below distinctly revolve around the sphere of migration, reaffirming the research's alignment with its predefined objectives.

Figure 23. Word cloud of the most frequent terms used to refer to the topic (focused Data Frame). Source: Author.

1.3. Translation

With the text now pre-processed, the subsequent phase involves translating the content within the 'cleaned_text' column, which houses the filtered text content of the tweets. Translation is a necessary step, especially when utilizing certain sentiment analysis tools like VADER, as they are

primarily designed to work with English texts. Thus, the translation step serves as a bridge to make the text compatible with the chosen sentiment analysis tool, highlighting the multifaceted challenges and considerations involved in text data preprocessing and analysis.

While performing sentiment analysis directly on Portuguese language text is feasible, as demonstrated by the pre-trained NLP model FinBertPTBR, specifically tailored for analysing sentiment in Brazilian Portuguese financial texts (L. L. Santos, Bianchi, and Costa 2023), the approach differs from the one chosen for this thesis. The alternative method involves machine learning and demands significant time, expertise, and resources for building, training, and testing a model tailored specifically for the migration topic. Nonetheless, the rationale for selecting VADER, as explained earlier, remains compelling to our purposes, and thus, translation becomes a necessary step.

It is crucial to acknowledge that when it comes to translating text within Python, certain constraints and limitations come into play. Beyond the fundamental reasons discussed earlier for opting to filter the dataset using specific hashtags, resulting in a more concise yet targeted data frame for analysis, the reduced size of the data frame has also proven advantageous in streamlining the translation process. These constraints primarily pertain to the use of translation libraries and are associated with both character limits and the rate of translation requests. Many translation libraries, while powerful, have restrictions on the maximum number of characters that can be translated in a single request or often impose rate limits on the number of requests that can be made within a specified time frame. Consequently, when dealing with substantial volumes of text to be translated, it is essential to consider the potential delays imposed by rate limits and plan accordingly, whether by staggering translation requests or utilizing alternative approaches to optimize efficiency. Moreover, it is worth noting that using certain translation services, especially those offered through APIs, may incur costs. These costs can accumulate, particularly when translating extensive datasets. Hence, budgetary considerations, along with time, was also thought out during the translation process. Given these constraints, having a smaller data frame has been a strategic advantage. It not only aligns the aim of having a sentiment analysis of a very specific topic, but also work with the character limits imposed by translation libraries and optimizes efficiency needed for the work.

2. Sentiment Analysis

Transitioning now to the final and pivotal aspect of this research, the sentiment analysis process deserves further explanation, including the steps leading to our results. The journey of conducting sentiment analysis for this research began with data collection, filtering, and translation,

as highlighted in the preceding discussions. Empowered by Python and a plethora of specialized libraries and tools, the research team undertook the ambitious task of transforming multilingual Twitter data into a coherent and translatable format. The translation step, as previously elucidated, was necessitated by the choice of the VADER sentiment analysis tool, which predominantly operates with English text. In the upcoming sections, it will be explored the intricacies of the sentiment analysis process, from the utilization of VADER to the evaluation of results, shedding light on the methodology employed and the insights gleaned along the way and at the end of the process.

Focusing on the practical implementation of sentiment analysis using the VADER, the first step involved importing the necessary resources for this analysis. The code snippet begins by importing the SentimentIntensityAnalyzer from the NLTK's VADER module. Additionally, the 'vader_lexicon' is downloaded, which is a lexicon-based resource containing sentiment scores for a vast array of words and phrases. This lexicon serves as a crucial reference point for VADER in determining the sentiment of text.

```
import nltk
from nltk.sentiment.vader import SentimentIntensityAnalyzer
nltk.download('vader_lexicon')
sia = SentimentIntensityAnalyzer()
```

Figure 24. Code for initializing VADER. Source: Author.

Subsequently, the 'SentimentIntensityAnalyzer' is initialized as 'sia'. This analyser is a pre-trained model designed to evaluate the sentiment of textual data, particularly in the context of social media content like tweets this research aimed to analyse. VADER employs a lexicon and a set of grammatical rules to assign sentiment scores to words, phrases, and entire sentences. These scores capture the polarity (positive, negative, or neutral) and the intensity of sentiment expressed in the text.

```
def classify_sentiment(compound_score):
    if compound_score >= 0.05:
        return 'positive'
    elif compound_score <= -0.05:
        return 'negative'
    else:
        return 'neutral'
```

Figure 25. Code defining the function to classify the sentiment scores. Source: Author.

Within the sentiment analysis process, determining the overall sentiment of a text is crucial to distil meaningful insights from the data. The 'classify_sentiment' is introduced to convert the compound sentiment score generated by the VADER into categorical sentiment labels. These labels help categorizing the sentiment as either positive, negative, or neutral. The function takes a

96

'compound_score' as its input, which is a numerical representation of sentiment ranging from -1 (extremely negative) to 1 (extremely positive), with 0 indicating a neutral sentiment. The function employs a set of conditional statements to categorize the sentiment based on the compound score. When the compound score is greater than or equal to 0.05, the function classifies the sentiment as 'positive.' Conversely, if the compound score is less than or equal to -0.05, it labels the sentiment as 'negative.' These thresholds are chosen to define the boundaries between positive and negative sentiments, with scores closer to 0 denoting neutral sentiments and allows for a nuanced classification of sentiment, capturing not only polarity but also the intensity of sentiment expressed in the text.

One of the paramount considerations when employing VADER for sentiment analysis is the establishment of threshold values that classify sentences into distinct sentiment categories: positive, neutral, or negative. These thresholds serve as a crucial tool for researchers and analysts, facilitating the categorization of sentiment within text. While the choice of thresholds can vary, there are standardized values often used in the literature: a compound score of 0.05 or higher signifies positive sentiment, a score below -0.05 indicates negative sentiment, and scores between -0.05 and 0.05 are regarded as neutral.[37]

To illustrate the practical application of these sentiment thresholds, below is demonstrated examples drawn from the dataset. Consider the following sentences as instances of different sentiment categories:

- Positive: "Crucial private sector solutions to overcome long-term challenges facing refugees host countries UNHCR unite to promote social and economic integration refugee population Brazil learn" (Compound: 0.5574).
- Neutral: "Lebanon ikmr refugees" (Compound: 0.0).
- Negative: "African refugees deported to Israel ended up dead" (Compound: -0.6486).

By applying these threshold values and utilizing VADER's sentiment analysis capabilities, it possible to effectively categorize sentences within the dataset, providing valuable insights into the prevailing sentiments associated with the migration topic.

---

[37] The creators of VADER (Hutto and Gilbert 2014), suggest this value and reinforce the current use of it by researchers: https://github.com/cjhutto/vaderSentiment

3.   Data Analysis and Results

In this comprehensive study of tweets related to migration issues, the capabilities of sentiment analysis were harnessed to delve deeper into the Brazilian's public feelings and attitudes towards these critical migrants and refugees. The analysis has thoughtfully categorized these tweets into three distinct sentiment categories: positive, negative, and neutral.

The pie chart gives a representation of the sentiment nuances within the filtered dataset during the whole period analysed (2015-2019). The prevalence of neutral sentiment implies that a substantial portion of the discourse maintains a tone
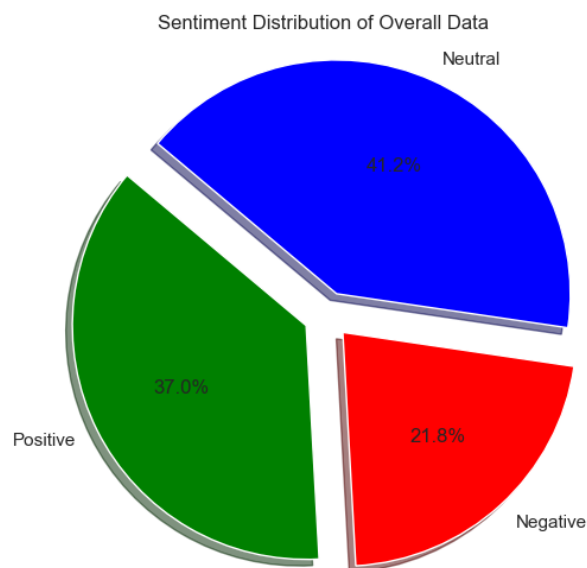
Figure 26. Sentiment distribution of the overall data. Source: Author.

of factual reporting, impartial information dissemination, or objective discussion of the thematic. Neutral tweets often revolve around sharing scientific findings, news updates, or the promotion of general awareness campaigns.

It is possible to see that the positive sentiments (37%) prevailed over the negative (21.8%). This substantial positivity suggests that many individuals within the Twitter community express hope, support, endorsement, and encouragement regarding efforts to address migration challenges. Positive sentiment can emerge from discussions related to successful migration, adaptation stories, resilience-building initiatives, support to institutional actions and cooperationinn. This sentiment can indicate a shared optimism surrounding the capacity to address migration issues within the country and some kind of empathy to the national and international events that unfolded along those five years.

The balance between positive and neutral sentiments emphasizes the nuanced and informative nature of conversations surrounding the topic, encapsulating a substantial majority of the dataset, approximately 78.2%.

Conversely, the smallest segment, representing 21.8% of the dataset, pertains to tweets classified as having a negative sentiment. While a minority, this component indicates that there are Twitter users who voice concerns, critiques, or frustrations regarding various aspects of migration. Negative sentiment could emanate from conversations surrounding many points such as responses to forced displacement, insufficient policy responses, instances of mobility injustices, disagreements with (in)actions of national and international entities and figures, xenophobia, or racism.
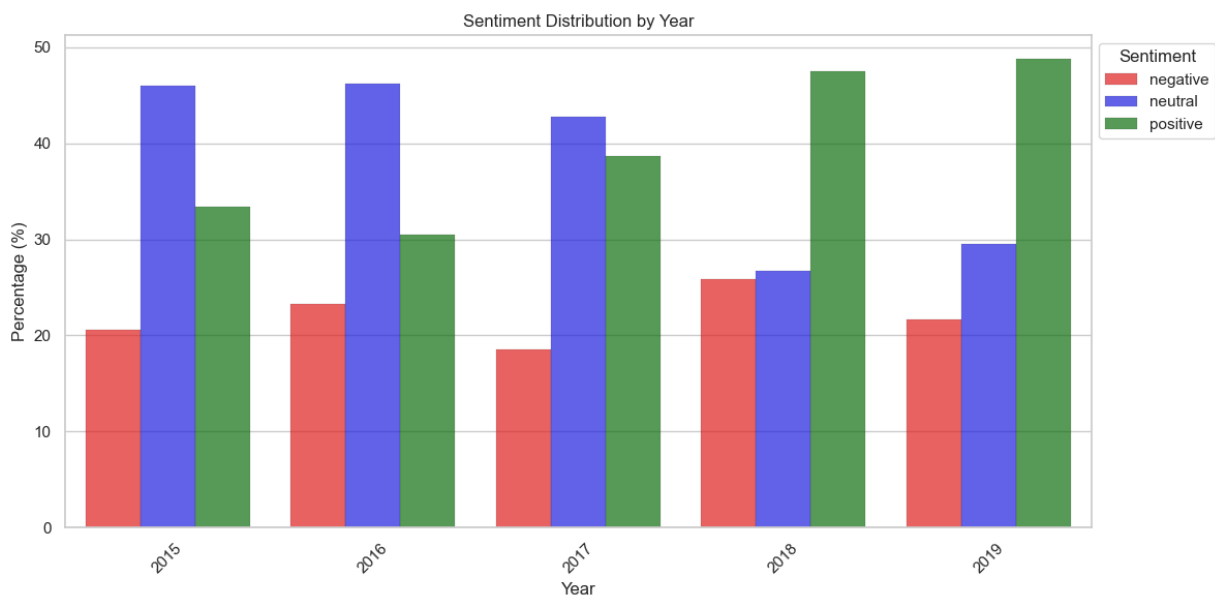
Figure 27. Sentiment distribution by year. Source: Author.

Over the course of five years, a noteworthy shift in sentiment becomes apparent, with 2017 serving as the pivotal starting point for this intriguing transformation. This period, following Dilma Rousseff's impeachment, was marked by a confluence of complex political and economic challenges that had been brewing since the early days of her second term and became increasingly pronounced after her removal from office. Additionally, a significant policy change occurred with the enactment of a new migration law in April 2017. While the initial overview of the entire dataset unveiled a prevalence of neutral sentiments, a more meticulous year-by-year analysis unveiled an evolving emotional landscape.

The transition from 2017 to 2018 stands out as particularly striking. Set against the backdrop of political uncertainty and an economy grappling with formidable challenges, there was a discernible surge in positive sentiments. During this period, neutral sentiments underwent a sharp decline, yielding to a substantial surge in positivity. Concurrently, negative sentiments experienced modest growth. This shift not only signifies a move towards emotional polarisation during the years 2018 and 2019 but also reflects a growing inclination toward more critical engagement with the subject matter. It might also suggest that, in tandem with the broader political landscape's polarization between right and left-wing ideologies, which intensified in 2018 and culminated in the election of Jair Bolsonaro, this polarization tendency extended to other societal issues within the Brazilian context.

Crucially, despite the backdrop of political and economic turbulence and the shift towards a more conservative government in Brazil under Bolsonaro's leadership, known for his alignment with Trump and staunch anti-migration stance, the prevalence of neutral sentiments did not predominantly shift towards negativity. This observation is intriguing, especially in a nation undergoing significant

political changes and grappling with formidable economic and political challenges. Additionally, as demonstrated in Figure 5, the years of 2018 and 2019 were the ones with the biggest number of asylum requests in Brazil.

It suggests that the sentiments expressed by the general public did not fully mirror those of their elected leader, highlighting the complexity of factors at play. In the context of the tweets analysed, and the prevailing discourse, it becomes evident that even with Bolsonaro's election, Twitter users expressed positivity toward migrants and refugee-related matters. Furthermore, the year 2018, which witnessed this pronounced shift, coincided with the commencement of Operation Welcome in Brazil—a humanitarian effort in the country's northern region. This synchronicity suggests widespread support, optimism, and a degree of empathy toward institutional actions aimed at addressing humanitarian challenges, even during a period of political transformation and nationalist rhetoric. It also suggests that even in times of political division and controversy, positive sentiments can prevail, reflecting the diversity of perspectives and the nuanced nature of public discourse surrounding migration and refugee issues in the context of this research.
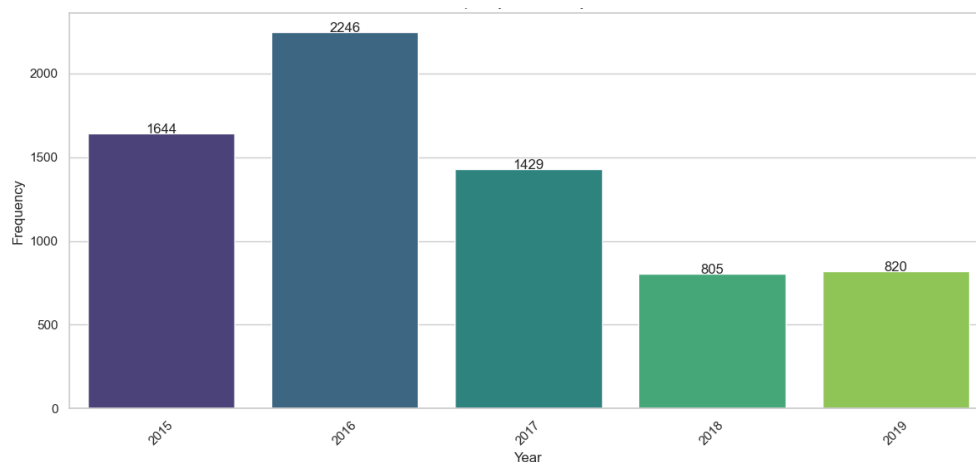
Figure 28. Number of tweets by year. Source: Author.
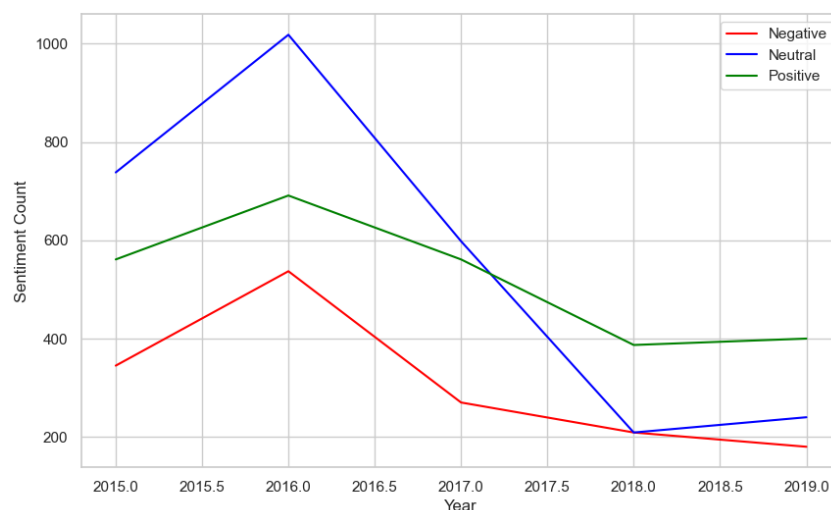
Figure 29. Overall sentiment distribution by year. Source: Author.

In the broader context of tweet distribution over these five years, a significant surge in engagement was witnessed from 2015 to 2016. However, this peak was followed by a gradual decline in interest and participation in the topic, starting in 2016 and continuing onwards. Analysing this trajectory alongside sentiment distribution reveals an intriguing pattern. As previously mentioned, the waning popularity of the subject coincided with a discernible departure from neutrality, giving way to a more pronounced presence of either positive or negative sentiments. This shift underscores a trend towards heightened emotional resonance and polarization within the discourse surrounding migration. The initial spike in engagement during 2015-2016 may be attributed to the notable increase in the flow of migrants from Venezuela to Brazil during that period, which evidently piqued public interest in the topic. A report from CONARE revealed that asylum requests from Venezuelans saw a substantial increase, rising from 829 in 2015 to 3.375 in 2016 (ACNUR 2018). This influx had several repercussions for the state of Roraima and the city of Pacaraima, situated at the frontier with Venezuela and serving as the primary entry point for Venezuelan refugees into Brazil.

In the context of global events, it is worth noting that between September and November of 2015, there is a notable upsurge in migration towards Europe, which coincides with the peak of the European Refugee Crisis. This international crisis, triggered by conflicts and instability in regions like Syria, Afghanistan, and Iraq, and Nigeria, led to a significant influx of refugees and migrants seeking safety in European countries. The images of overcrowded boats, perilous journeys, and refugee camps dominated international news and sparked widespread humanitarian concerns.

The European Refugee Crisis had a profound impact on the political, social, and economic landscapes of European nations. It prompted debates on immigration policies, border security, and the responsibilities of European countries in addressing this humanitarian challenge. The crisis also revealed divisions within the European Union regarding how to handle the influx of refugees, with some nations advocating for open borders and solidarity, while others advocated stricter immigration controls. This complex and dynamic situation also likely played a role in the increased attention and discussions on migration-related topics on social media platforms like Twitter during this period.

Taking a closer look at the year 2015, a significant surge in engagement with the topic becomes evident, particularly from the month of August onwards. Interestingly, this heightened interest aligns with a pivotal moment in the European Refugee Crisis. In August of that year, leaders from countries including Hungary, Poland, Czech Republic, and Slovakia publicly voiced their opposition to the European Commission's proposal for the redistribution of refugees across EU member states. Another compelling facet emerges in the month of April when negative sentiments outweighed positive ones. This alignment coincided with a series of measures introduced to address
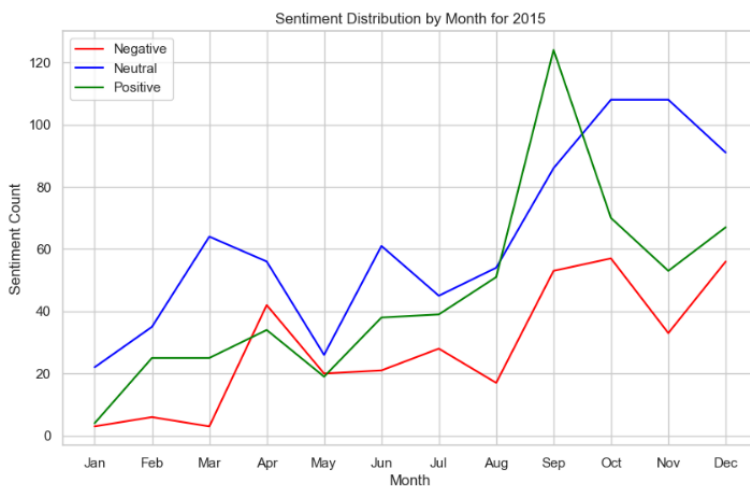
Figure 30. Sentiment distribution by month (2015). Source: Author.
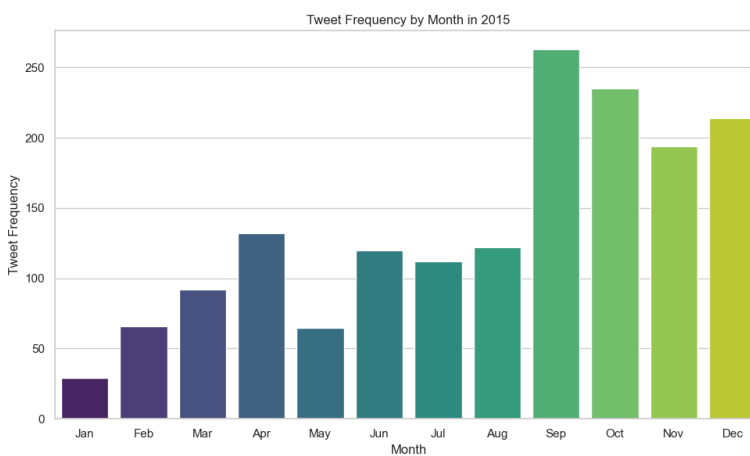
Figure 31. Frequency of tweets by month (2015). Source: Author.

the refugee crisis in Europe. These measures were prompted by tragic accidents that claimed the lives of hundreds of people in the Mediterranean (BBC News 2015). The prevalence of negative emotions during this period can be attributed to several factors, reactions evoked by news of these tragic incidents and some opinions against the measures taken by the EU (Carnegie Europe 2015).

Examining the sentiments expressed in tweets during this period, it's notable that all three sentiment categories—positive, negative, and neutral—experienced an increase. However, positive sentiments predominated among almost all the months in the tweets analysed during this critical period. This observation underscores the complexity of public opinion and highlights the coexistence of empathy, humanitarian concern, and dissenting viewpoints in the face of a multifaceted crisis.

The year 2016 commenced with a notable prevalence of negative sentiments concerning the topic of migration. In fact, this negativity extended into two specific months, March and April, which also coincided with heightened tweet activity. These two months stand out as periods when negative sentiments consistently overshadowed positive ones.

The month of March witnessed a surge in protests against the government of Dilma Rousseff, reflecting a charged political atmosphere in Brazil. Concurrently, the period saw the emergence of corruption allegations against former (and current) President Lula da Silva as part of the Operation Car Wash investigation. These significant events likely contributed to the prevalence of negative sentiments in tweets during this timeframe, underscoring the influence of political developments on public perceptions and discourse surrounding migration.

Internationally, as reported by *Médecins Sans Frontiers* (MSF), numerous rescue operations continued to occur in the Mediterranean and other locations (MSF 2016). The prevailing negative

emotions during this period can also be attributed to various factors, primarily the emotional reactions evoked by news of these tragic incidents.
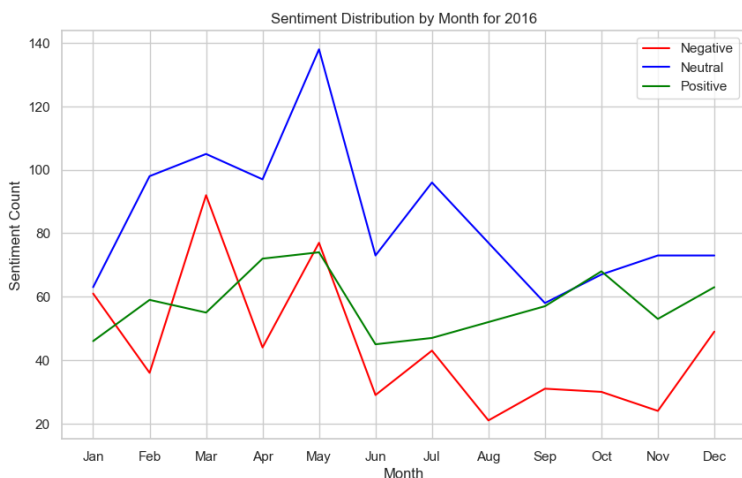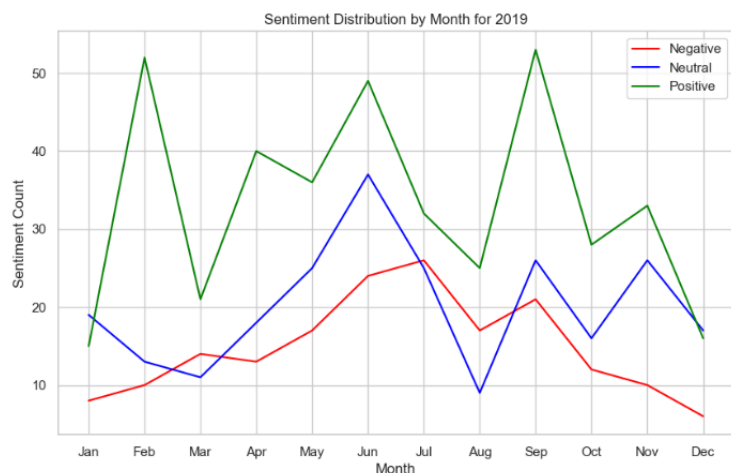

Figure 32. Distribution of sentiments by month (2016). Source: Author.
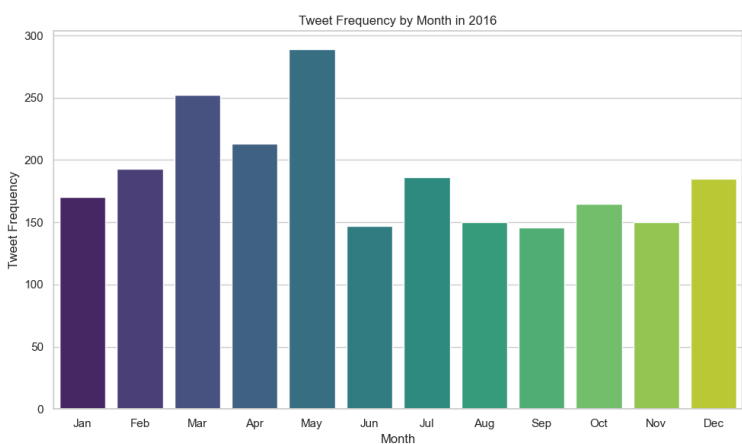
Figure 33. Frequency of tweets by month (2016). Source: Author.

Another intriguing aspect is that, amidst two months predominantly characterized by negative sentiments, April stands out as a period when positive sentiments exceeded the negative. This aligns with the previously mentioned statement released by FENAPEF in April of that year, which brought attention to the legal prohibition preventing foreign nationals from participating in political demonstrations in Brazil. Simultaneously, there was a decline in job prospects and a rise in unemployment rates. This shift towards positive sentiments during April suggests not only positive sentiment towards migrants and refugees but also engagement with the broader social and political landscape in opposition to the government's measures.

The notable decrease in negative sentiments in 2016 can be linked to the widespread positive support garnered by the first-ever refugee team competing in the Rio Olympics, which commenced in August 2016. The event might have served as a source of inspiration and unity, fostering a more positive outlook on the topic of migration and refugees.

In the early 2017, one of the initial actions taken by the newly elected President Donald Trump was to order the construction of a wall along the Mexico-United States border to regulate migratory flows into the United States. Despite the widespread attention and global repercussions of this decision, the sentiment chart reveals an initial increase in positive sentiments at the start of the year.
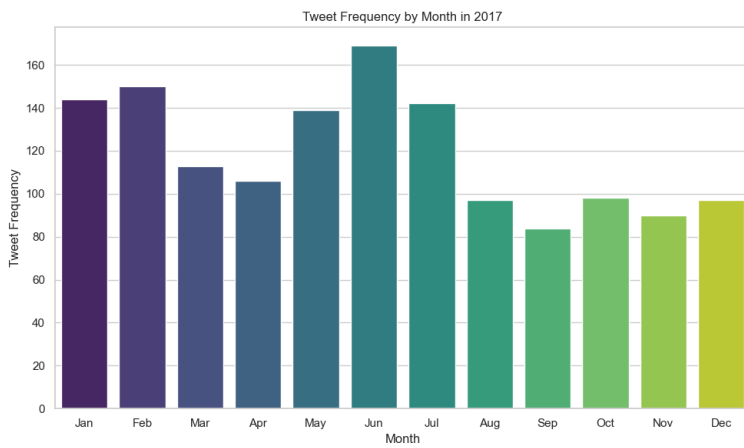
Figure 35. Frequency of tweets by month (2017). Source: Author.
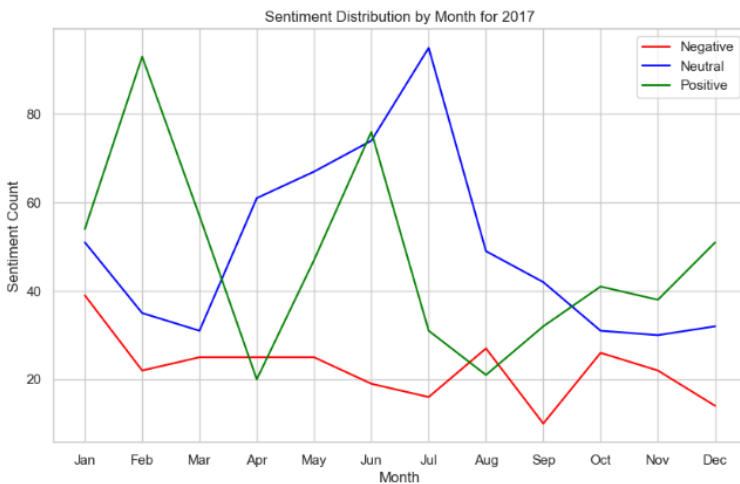
Figure 34. Distribution of sentiments by month (2017). Source: Author.

This suggests that, even amid strong anti-migration rhetoric and policies, a significant level of positive sentiment persisted.

However, a notable shift occurred in April, marked by a substantial decline in positive sentiments on Twitter, with negative sentiments taking the lead. As previously observed in the analysis of hashtags, April 2017 witnessed the emergence of new hashtags expressing support for the vetoes applied by then-President Michel Temer in the New Migration Law, which was under consideration and subsequently approved in the following month. The debates surrounding which articles should be vetoed dominated the media, with some individuals expressing support for these vetoes and others opposing them. As previously mentioned, the law was ultimately approved in May 2017, albeit with significant articles vetoed. This dynamic reflects the complex and nuanced sentiments prevalent during this period of legal and political change.

In addition, another significant event captured public attention during that month. This event was the imprisonment of former Brazilian President Lula da Silva as part of the Operation Car Wash. Lula's arrest marked a pivotal moment in Brazilian politics and society, and also might have elicited a wide range of emotional responses, complementing the complex interplay of political, legal, and social factors shaping public discourse in Brazil.

The sentiment graph of 2018 reveals an intriguing pattern of emotional shifts. At the year's outset, there's a noticeable surge in positivity, possibly signaling support for Operation Welcome, which commenced early in the year. Notably, it also uncovers a period marked by a tendency toward polarization, especially notable in July and August. This fluctuation in sentiment appears to parallel the broader political landscape in Brazil during the eventful year of 2018.
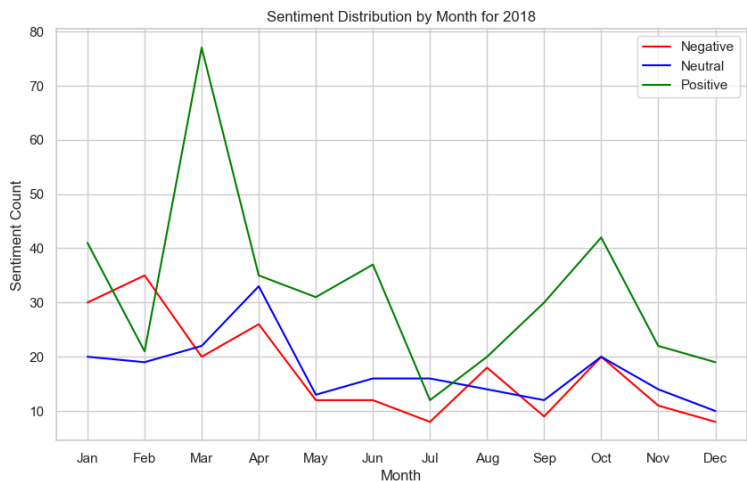
Figure 36. Distribution of sentiments by month (2018). Source: Author.

The year of 2018 was characterized by presidential campaigns and elections, during which Jair Bolsonaro aligned himself with global leaders advocating stringent anti-immigration measures. This political
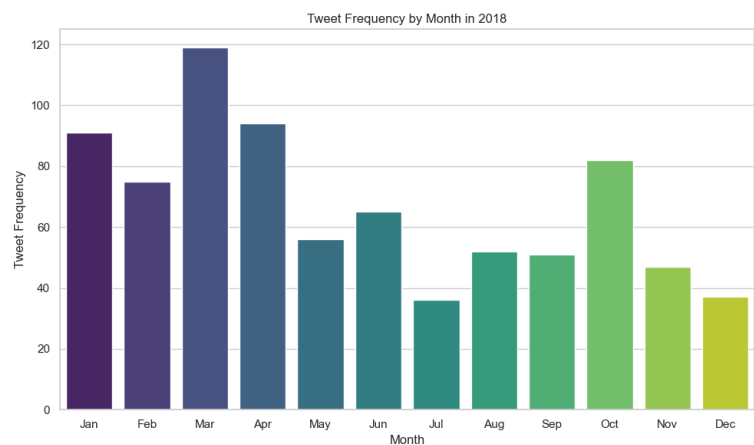
Figure 37. Frequency of tweets by month (2018). Source: Author.

period that started dragging the country towars polarity between lef and right-wing, and brought into light more conservative discourses, coincided with the deepening socio-economic crisis in Venezuela, leading to a substantial influx of Venezuelan migrants and refugees into Brazil. Conservative politicians pressed for the closure of Brazil's border with Venezuela, but this attempt was halted by the Federal Supreme Court (BBC News 2018a).

In February 2018, the federal government took steps to bolster security and assistance at the Venezuela-Brazil border. During this time, ministers announced a series of measures while visiting Boa Vista, including plans for the internalization of refugees, aimed at relocating them to other states for reception, along with commitments to provide additional support from both the Army and Federal Police.

Adding to the complexity of the situation, August 2018 witnessed protests in Pacaraima, where local residents demonstrated against the substantial influx of immigrants. These protesters voiced their concerns regarding security and healthcare issues. Additionally, an incident sparked by local residents through social media led to attacks on Venezuelan camps and the destruction of tents

and belongings. The escalating tensions prompted many Venezuelans to flee back across the border (BBC News 2018b).

Figure 38. Distribution of sentiments by month (2019). Source: Author.
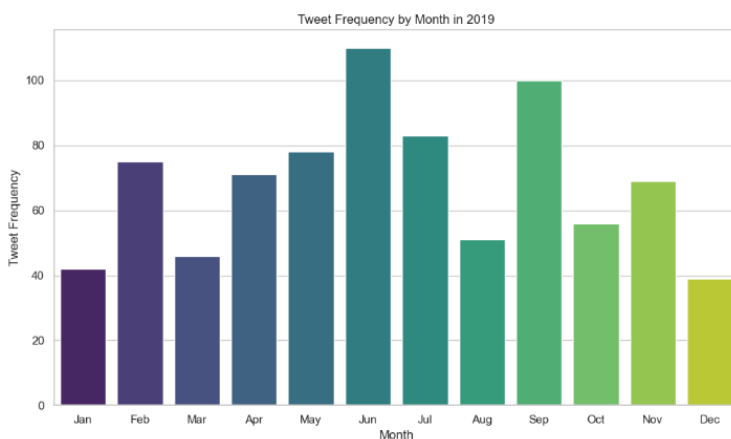
Figure 39. Frequency of tweets by month (2019). Source: Author.

The year of 2019 was marked by the inauguration of Jair Bolsonaro as President, and one of his early policy moves was to announce Brazil's withdrawal from the Global Pact of Migration, aligning the country with other conservative leaders globally. This signalled a shift towards treating migration as a domestic and security concern, departing from the previous advancements towards a more inclusive approach.

As Brazil's stance towards migration evolved, concerns emerged regarding inclusivity and human rights. Initially, the country implemented measures to encourage and accommodate Venezuelan migrants, reflecting a more open approach. However, subsequent steps reflected a more restricted posture. Notably, in July 2019, Decision n.666 was issued by the Ministry of Justice and Public Security, allowing for the summary deportation of migrants suspected of criminal involvement (Brasil 2019). Additionally, there were efforts to tighten immigration controls, and rhetoric labelling migrants as potentially dangerous gained traction. In relation to this, it is possible to consider that in the tendency to polarisation in the subsequent month might have something to do with the measures taken by the government.

It is worth noting that in December 2019, there is a noticeable decline in the prominence of the topic, with negative sentiment dropping significantly. This is particularly intriguing given the surge in xenophobic comments directed at the Chinese population in light of the onset of the COVID-19 pandemic. Further research of the subsequent year and focused in the COVID-19 topic can give a glimpse towards people's sentiment during the pandemic.

Despite these policy changes and shifting narratives, the sentiment graph for 2019 tells a different story. Positive sentiments prevailed throughout the year, suggesting a disconnect between

government opinions and public sentiment. This intriguing observation raises questions about the factors influencing public opinion regarding immigration during this period and highlights the complexity of the issue within Brazil's socio-political landscape.

Over the period spanning 2015 to 2019, Brazil underwent significant transformations in its migration landscape, strongly influenced by a blend of internal and external factors. The political and economic turmoil in Venezuela triggered a substantial influx of migrants and refugees into Brazil, with Roraima emerging as the primary gateway. Brazil's initial response was characterized by Operation Welcome, a humanitarian initiative involving the government, UNHCR, NGOs, and civil society, which underscored support and assistance for incoming Venezuelans.

Despite the initial shift toward a more conservative government, which began with Dilma's impeachment, the New Migration Law of 2017 marked a profound change in how Brazil approached migrants and refugees. The transition to a more open law, offering enhanced protection and rights, alongside the enduring welcoming legal approach to migration flows throughout this period, even amid a delicate political landscape, potentially reflects the robustness of democracy in the country. It also highlights the effective efforts of those responsible for safeguarding the rights of migrants and refugees.

An analysis of sentiment in Twitter data throughout these years reveals a dynamic emotional terrain. As the migrant influx grew, Brazil initially projected a welcoming posture, epitomized by Operation Welcome and the prevalence of positive sentiments. This phase might symbolize a compassionate response to a regional humanitarian crisis, with the upbeat sentiments mirroring public support for this approach.

The pivotal year 2018 emerged as a turning point in this narrative. The election of Jair Bolsonaro, renowned for his ultra-conservative and nationalist stances, heralded a shift in Brazil's migration approach. Aligning himself with leaders like former U.S. President Donald Trump, Bolsonaro's rhetoric underscored domestic security over global cooperation concerning migration policy, a marked departure from the more welcoming discourse of prior years.

Notably, despite these shifts, the sentiments expressed by the Brazilian population, as evident through sentiment analysis of Twitter data, exhibited a certain disconnect from the government's leader. Positive sentiments consistently held sway, indicating that, at least on social media platforms like Twitter, public sentiment toward migrant and refugee matters generally remained positive. This held true even during a period characterized by political polarization and policy alterations.

The predominance of positive sentiment in 2019 and throughout the overall period underscores the Brazilian population's overall favourable disposition toward this matter, shedding light on the intricate and multifaceted nature of migration in Brazil, even in the face of certain more restrictive governmental actions.

In conclusion, Brazil's journey during the 2015-2019 period mirrors the dynamic nature of migration and its intersection with politics, economics, and society. While government policies and political rhetoric evolved, the sentiments expressed by the Brazilian population remained largely optimistic, exemplifying the multifaceted character of migration in Brazil. This issue is not solely shaped by policy decisions but also by the compassionate perspective of the broader society.

Conclusion

The primary goal of this thesis was to explore the intricate nuances of Brazilian sentiment regarding the topic of migration and refugees during years of political instability, from 2015 to the end of 2019, with a focus on examining whether a range of international and national events played a role in shaping those sentiments.

This undertaking involved employing various qualitative methodologies, initially to understand the context, the analysed social media platform (Twitter), and the tools for data extraction and analysis. The results, however, encompass both quantitative and qualitative aspects, due to their involvement in textual analysis. This multifaceted approach not only offers a fresh perspective on the topic but also provides an alternative means of comprehending the sentiments expressed within the realm of social media.

In today's digitally interconnected world, the study of social media platforms holds paramount importance. These platforms have evolved into dynamic ecosystems that reflect and influence human behaviour, identity, and sentiment in unprecedented ways. People of all ages engage with social media on a daily basis, driven by diverse motivations. Marketing, social influencers, and news agencies continuously shape our perceptions of how we should behave, what we should wear, who we should believe, and what is considered good or bad. These influences contribute significantly to the formation of our values, identity, and understanding of the world. Neglecting this wealth of information about people's lives not only overlooks a substantial part of who they are but also neglects a crucial collective input that shapes our society as a whole.

Twitter, as a microcosm of a relevant part of our digital society as demonstrated in this thesis, stands as a critical tool for comprehending the complexities of human behaviour, sentiment, and the intricate web of interconnectedness that defines our world today. Its study transcends academic boundaries, offering a profound impact on policymaking and the pursuit of a more inclusive and informed society.

In the pursuit of achieving the aim of this research, it was set out to accomplish four key goals: (a) gain insights into public opinions regarding migrants and refugees; (b) investigate the correlation of international and national events on the nuances of sentiments expressed on Twitter during this specific period; (c) identify and quantify the specific terms used within the Twitter discourse concerning migrants and refugees; (d) ascertain whether discernible shifts in sentiment occurred, indicating a transition from predominantly positive to negative sentiment, or vice versa, particularly in response to changes in government leadership.

The first chapter aimed to provide a general and more focused overview of migration flows between 2010 and 2020. It was possible to observe the main events that drove people to migrate in this decade, with a greater emphasis on forced displacements, as they often spark debates and discussions and is a trending topic in the last decade. The emergence of new migration corridors, which led to the European crisis, brought various conflicts to the forefront that typically do not receive much attention in Western media and showed a hardening of migration and refugee policies in Europe.

More attention was given to migratory flows toward Brazil, particularly those related to events in Haiti and Venezuela. It was possible to observe an increase in asylum requests in Brazil between 2015 and 2020, the same period when Brazil was experiencing political turmoil. In this chapter, we also observed the development of the country's political and economic problems, a period of significant instability and polarization, as well as a shift from a more liberal government to a more conservative one. Changes in leadership also altered the discourse on migration, with some more restrictive measures being introduced.

However, this period marked an important milestone in Brazilian legislation on migrants and refugees, with the approval of the new migration law, even with some vetoes. This law brought a new outlook for migrants and refugees, with more rights and possibilities for inclusion. Lastly, we highlighted how Twitter played a significant role in Brazil's political movement during this period and the relevance of the platform for expression in the country. This entire chapter aimed not only to contextualize the debate but also to provide the necessary tools to understand the results of the data analysis.

The second chapter, in turn, aimed to emphasize the importance of Twitter, bringing to light the history of its development and its features, highlighting the significance the platform has gained over the years for our current society. Twitter is now a platform that provides information and serves as a space for social interaction, where both aspects interact organically. Due to its informative nature, there is a tendency for people to comment on or share this news, making news and information more likely to be topics of discussion (but not exclusively). Furthermore, its microblog format also encourages posts containing personal updates and opinions, making it a significant reservoir of public opinion.

The chapter also stressed the importance of this data for academic research and various opportunities and methods for analysing this data. Unfortunately, for society and academia, recent updates to Twitter have posed a significant obstacle to conducting research with the platform's data. Among popular social media networks like Facebook and Instagram, Twitter used to offer the most accessible data for research. Today, like the others, it requires a financial investment that severely limits the development of new research. More importantly, it raises a crucial debate about the control

of such data by private companies and how easily it can be erased. As mentioned earlier, social media is now a part of millions of people's lives worldwide, and many events affecting offline life happen there, such as the influence a tweet can have on stock market fluctuations or how a president's comment can impact a country's international relations. Therefore, it is of utmost importance that this debate continues, as it involves a substantial amount of historical data. Moreover, the chapter reinforced the choice of Twitter for this analysis, aligning with what was explained in the first chapter.

The third chapter provides brief information about the Python programming language and its libraries, particularly those chosen for data pre-processing and sentiment analysis. It also offers a brief explanation of sentiment analysis and the tool (VADER) used to perform it. It is important to note that the methodology, from choosing Python as the programming language to selecting libraries and the use of VADER, was the author's choice. This means that there are other ways to perform this analysis, and the choices largely depend on the desired objectives as well as the challenges encountered throughout the process. Lastly, the chapter offers a brief explanation of the process of gaining access to the API and data extraction.

Finally, the last chapter outlines all the necessary steps for data analysis. In this section, the codes needed for data cleaning and sentiment analysis are exemplified practically. However, each step was meticulously analysed and reviewed, resulting from numerous hours of coding, problem-solving, and studies. The results firstly show that analysing tweets in relation to geolocation poses a challenge, even though some studies used this information before (see Zagheni et al. 2014). Regarding this study, geolocation-based information would have been inconclusive, and the choice not to pursue this path aligned with the set objectives.

Regarding goal (a), we can see that the majority of sentiments are neutral, followed by positive sentiments, and lastly, negative sentiments. This indicates that, in contrast to previous mentioned research (Inuwa-Dutse, Liptrott, and Korkontzelos 2020; Yantseva 2020), in this context and within the scope of this study, negative sentiment did not prevail. Furthermore, it was possible to understand that over the five analysed years, there was a trend toward polarization that also mirrored the trend of more polarized political debates that started within those five years, particularly from 2017 and 2018. This trend underscores the complex interplay between political discourse and public sentiment, emphasizing the nuanced nature of Brazilian attitudes towards migration and refugees during the period. However, the shift from neutral to positive sentiment starting in 2017 suggests that, regardless of the change from a liberal to a more conservative government, the sentiment expressed by people on Twitter did not align with the changing from more welcoming to xenophobic or exclusionary discourses by the leaders in power (goal d).

Regarding goal (c), the terms people used to discuss migration can be seen both through the choice of hashtags and through the word cloud displaying terms and frequency. Overall, the terms

align with the sentiment of the tweets, being predominantly neutral or positive, with some instances of negative hashtags or terms, primarily related, as observed, to the new migration law, where the most frequently used hashtags are against the law, and the sentiments for that month show negativity. This observation highlights the influence of policy changes on public sentiment and the interplay between legislative actions and public discourse.

As for goal (b), it was possible to establish connections between national and international events and the nuances of people's sentiments on the platform. The analysis of hashtags and sentiment nuances by month showed that, in addition to demonstrating knowledge of the events, people also expressed opinions about them, indicating a relationship between the events and public sentiment. Finally, it was possible to identify some events and terms that demonstrated how the impact of international and national events can be discerned through people's tweets, showing that the method is capable of identifying public opinion and offering various insights into events and how they affect society.

As for the primary goal, it is evident that both national and international events had a discernible impact on how people expressed their sentiments about migrants and refugees on Twitter. The analysis revealed that, contrary to prevailing assumptions, negative sentiments did not dominate the discourse during this period of political instability. Instead, a majority of sentiments leaned towards neutrality, followed by positive expressions. The connection established between national and international events and the nuances of public sentiment on Twitter demonstrated a notable relationship between these events and the sentiments expressed by the Twitter community. People not only demonstrated awareness of these events but also used the platform to express their opinions about them. This underscores Twitter's role as both a reflector and influencer of public sentiment, offering valuable insights into how major events shape public discourse on migration and refugees.

This thesis has successfully delved into the intricate landscape of Brazilian sentiment concerning migration and refugees during years of political instability. It has shed light on the dynamic interplay between political events, policy changes, and public discourse, emphasizing the importance of understanding these dynamics for policymaking and fostering informed, inclusive societies. The insights gleaned from this research not only contribute to the literature on sentiment analysis but also provide valuable perspectives on the role of social media in shaping public opinion and discourse in a rapidly evolving digital age, championing the use of diverse methodologies in conjunction with traditional approaches. It also introduces new data and insights into the public opinion of a segment of Brazilian society concerning migration, which holds the potential to enhance our comprehension of societal dynamics and serve as an invaluable source of information. When combined with other research methods, these findings offer valuable perspectives for the development of more inclusive policies.

It is worth noting that this research operates within the constraints of a dataset encompassing over a million tweets, showcasing the myriad possibilities for further analysis within this extensive dataset. Furthermore, it is essential to emphasize that Twitter represents only a fraction of the population. Nevertheless, the fusion of this method with other analytical approaches can yield powerful insights for the social sciences.

The innovative aspect of conducting sentiment analysis within the context of tweets from a Portuguese-speaking country, as applied to migration studies, represents a significant contribution to both the field of digital discourse analysis and the broader realm of academic inquiry. While this endeavour has yielded important and insightful findings, it has also unveiled several challenges that warrant further exploration and development in the field. One notable limitation pertains to language-related constraints, which have imposed significant restrictions on the extent of data analysis feasible within the confines of this master's thesis.

As previously mentioned, the imperative need for tweet translation introduced a distinctive set of challenges. Translating the entire dataset, while desirable, would have necessitated a substantial financial investment, given the limitations inherent in online translation APIs that are not freely accessible. This underscores the evolving nature of sentiment analysis, offering a myriad of potential applications and opportunities for interdisciplinary collaboration.

In conclusion, this thesis has embarked on a comprehensive journey through the intricate landscape of Brazilian sentiment concerning migration and refugees amidst years of political turbulence. Leveraging a multifaceted approach and harnessing the power of social media analysis, we have gleaned invaluable insights into public opinions, the influence of national and international events, and the language employed in discussing migration. While this research has illuminated a specific segment of Brazilian society, it underscores the ever-evolving dynamics of digital discourse and its potential to inform policymaking and cultivate a more inclusive and informed society.

As we look towards the future, it is imperative to persist in exploring innovative methodologies within the realm of social media analysis, recognizing their dynamic nature and their profound impact on our comprehension of the intricate web of interconnectedness that defines our modern world. This ongoing exploration promises to deepen our understanding of societal sentiments and pave the way for more informed decision-making, ultimately fostering a society that is better equipped to address the challenges and opportunities presented by migration and refugee issues in an increasingly interconnected global landscape.

List of Illustrations and Graphs

List of Tables

Bibliography

Abdullah, Nor Athiyah, Dai Nishioka, and Yuko Murayama. 2016. 'Questionnaire Testing: Identifying Twitter User's Information Sharing Behavior during Disasters'. *Journal of Information Processing* 24 (1): 20–28. https://doi.org/10.2197/ipsjjip.24.20.

ACNUR. 2018. 'Refúgio Em Números'. Brasil: Ministério da Justiça e Segurança Publica/CONARE. https://www.acnur.org/portugues/wp-content/uploads/2018/01/refugio-em-numeros-2010-2016.pdf.

Aday, Sean, Henry Farrell, Deen Freelon, Marc Lynch, John Sides, and Michael Dewar. 2013. 'Watching From Afar: Media Consumption Patterns Around the Arab Spring'. *American Behavioral Scientist* 57 (7): 899–919. https://doi.org/10.1177/0002764213479373.

Adi, Ana, Kristofer Erickson, and Darren G. Lilleker. 2014. 'Elite Tweets: Analyzing the Twitter Communication Patterns of Labour Party Peers in the House of Lords: Elite Tweets'. *Policy & Internet* 6 (1): 1–27. https://doi.org/10.1002/1944-2866.POI350.

Aguiar, Ruth Amaral. 2020. 'O Governo de Jair Bolsonaro e o Impacto Sobre Refugiados No Brasil'. *Conjuntura Internacional* (blog). 27 October 2020. /web/20230714114455/https://pucminasconjuntura.wordpress.com/2020/10/27/o-governo-de-jair-bolsonaro-e-o-impacto-sobre-refugiados-no-brasil/.

Agustiningsih, Kartikasari Kusuma, Ema Utami, and Hanif Al Fatta. 2021. 'Sentiment Analysis of COVID-19 Vaccine on Twitter Social Media: Systematic Literature Review'. *2021 IEEE 5th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, November. https://doi.org/10.1109/icitisee53823.2021.9655960.

Al Jazeera. 2019. 'Brazil's Lula Convicted to Keep Him from 2018 Election: Report'. *Al Jazeera* (blog). 10 June 2019. /web/20230714101942/https://www.aljazeera.com/news/2019/6/10/brazils-lula-convicted-to-keep-him-from-2018-election-report.

Alabrah, Amerah, Husam M. Alawadh, Ofonime Dominic Okon, Talha Meraj, and Hafiz Tayyab Rauf. 2022. 'Gulf Countries' Citizens' Acceptance of COVID-19 Vaccines—A Machine Learning Approach'. *Mathematics* 10 (3): 467–467. https://doi.org/10.3390/math10030467.

Almeida, Maria Hermínia Tavares de. 2014. 'A Política Social No Governo Lula'. *Novos Estudos - CEBRAP* 1 (November). http://socialsciences.scielo.org/scielo.php?script=sci_arttext&pid=S0101-33002005000100002#nota01.

Alromema, Waseem. 2022. 'Sentiment Analysis Applications during COVID-19 Pandemics: An Exploratory Review'. *International Journal of Recent Technology and Engineering* 10 (6): 114–18. https://doi.org/10.35940/ijrte.f6855.0310622.

AlSayyad, Nezar, and Muna Guvenc. 2015. 'Virtual Uprisings: On the Interaction of New Social Media, Traditional Media Coverage and Urban Space during the "Arab Spring"'. *Urban Studies* 52 (11): 2018–34. https://doi.org/10.1177/0042098013505881.

Amaral, Ana Paula Martins, and Luiz Rosado Costa. 2017. 'A (Não) Criminalização Das Migrações e Políticas Migratórias No Brasil: Do Estatuto Do Estrangeiro à Nova Lei de Migração |

Migration Policy and Migrations' (Non) Criminalization in Brazil: From the Foreigners Statute to the New Migration Law' 31 (2): 208–28. https://doi.org/10.5335/rjd.v31i2.7147.

Amaral, Marcelo Santos, and José Antonio Gomes De Pinho. 2018. 'Eleições Parlamentares No Brasil: O Uso Do Twitter Na Busca Por Votos'. *Revista de Administração Contemporânea* 22 (4): 466–86. https://doi.org/10.1590/1982-7849rac2018170269.

Ansari, Tarique Jamal, and Naseem Ahmad Khan. 2021. 'Worldwide COVID-19 Vaccines Sentiment Analysis Through Twitter Content' 18 (6). https://doi.org/10.29333/ejgm/11316.

Aragón, Pablo, Karolin Eva Kappler, Andreas Kaltenbrunner, David Laniado, and Yana Volkovich. 2013. 'Communication Dynamics in Twitter during Political Campaigns: The Case of the 2011 Spanish National Election: Communication Dynamics in Twitter During Political Campaigns'. *Policy & Internet* 5 (2): 183–206. https://doi.org/10.1002/1944-2866.POI327.

Azevedo, Rita. 2015. 'Setembro de 2015: Bolsonaro Chama Refugiados de "Escória Do Mundo"'. *Exame* (blog). 22 September 2015. https://web.archive.org/web/20230710165416/https://exame.com/favicon.ico.

Baeninger, Rosana, and Roberta Guimarães Peres. 2017. 'Migração de Crise: A Migração Haitiana Para o Brasil'. *Revista Brasileira de Estudos de População* 34 (1): 119–43. https://doi.org/10.20947/s0102-3098a0017.

Baptista, Érica Anita, Gabriella Hauber, and Maiara Orlandini. 2022. 'Despolitização e Populismo'. *Media & Jornalismo* 22 (40): 105–19. https://doi.org/10.14195/2183-5462_40_5.

Barba, Mariana della, and Marina Wentzel. 2016. 'Discurso de Bolsonaro Deixa Ativistas "Estarrecidos" e Leva OAB a Pedir Sua Cassação'. *BBC News* (blog). 19 April 2016. /web/20230714100936/https://www.bbc.com/portuguese/noticias/2016/04/160415_bolsonaro_ongs_oab_mdb.

Bauman, Zygmunt. 2016. *Strangers at Our Door*. 1st ed. Malden, MA: Polity.

BBC News. 2015. 'UE Anuncia Pacote de Medidas Para Crise de Imigração No Mediterrâneo'. *BBC News Brasil* (blog). 20 April 2015. /web/20230920133048/https://www.bbc.com/portuguese/noticias/2015/04/150420_regras_ue_imigrantes_rb.

———. 2016. 'Brazil Petrobras Scandal: Former President Lula Questioned'. *BBC News* (blog). 4 March 2016. /web/20230714101250/https://www.bbc.com/news/world-latin-america-35725415.

———. 2018a. 'A Cronologia Da Crise Migratória Em Pacaraima, Na Fronteira Entre Brasil e Venezuela'. /Web/20230920131843/https://www.bbc.com/portuguese/brasil-45242682. *BBC News* (blog). 20 August 2018.

———. 2018b. '"Bota Fogo!": O Ataque de Brasileiros a Imigrantes Venezuelanos Em Pacaraima'. *BBC News* (blog). 20 August 2018. /web/20230920111147/https://www.bbc.com/portuguese/brasil-45245644.

———. 2019. 'Twitter's Retweet Inventor Says Idea Was "Loaded Weapon"'. *BBC News* (blog). 25 July 2019. https://web.archive.org/web/20230713090105/https://www.bbc.com/news/technology-49113278.

———. 2021. 'Lula: Brazil Ex-President's Corruption Convictions Annulled'. *BBC News* (blog). 9 March 2021. /web/20230714102303/https://www.bbc.com/news/world-latin-america-56326389.

Bello, Valeria, and Sarah Léonard. 2023. *The Spiralling of the Securitisation of Migration in the European Union*. Research in Ethnic and Migration Studies. Abingdon: Routledge.

Bianchi, Daniella. 2020. 'What Can Tweets Tell Us about This Historic Moment?' *Twitter Blog* (blog). 2020. /web/20230714111018/https://business.twitter.com/en/blog/twitter-conversation-report-black-lives-matter.html.

Böhme, Marcus H., André Gröger, and Tobias Stöhr. 2020. 'Searching for a Better Life: Predicting International Migration with Online Search Keywords'. *Journal of Development Economics* 142 (January): 102347. https://doi.org/10.1016/j.jdeveco.2019.04.002.

Boulianne, Shelley. 2015. 'Social Media Use and Participation: A Meta-Analysis of Current Research'. *Information, Communication & Society* 18 (5): 524–38. https://doi.org/10.1080/1369118X.2015.1008542.

Bouvier, Gwen. 2015. 'What Is a Discourse Approach to Twitter, Facebook, YouTube and Other Social Media: Connecting with Other Academic Fields?' *Journal of Multicultural Discourses* 10 (2): 149–62. https://doi.org/10.1080/17447143.2015.1042381.

Boyd, D., S. Golder, and G. Lotan. 2010. 'Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter'. In *2010 43rd Hawaii International Conference on System Sciences*, 1–10. Honolulu, HI: IEEE. https://doi.org/10.1109/HICSS.2010.412.

Brasil, Governo do. 1980. *Estatuto do Estrangeiro*. http://www.planalto.gov.br/ccivil_03/leis/l6815.htm.

———. 1997. *Define mecanismos para a implementação do Estatuto dos Refugiados de 1951, e determina outras providências*. http://www.planalto.gov.br/ccivil_03/leis/l9474.htm.

———. 2012. *LEI Nº 12.663, DE 5 DE JUNHO DE 2012*. https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/l12663.htm.

———. 2017. *Lei de Migração*.

———. 2019. *PORTARIA Nº 666, DE 25 DE JULHO DE 2019: Dispõe sobre o impedimento de ingresso, a repatriação e a deportação sumária de pessoa perigosa ou que tenha praticado ato contrário aos princípios e objetivos dispostos na Constituição Federal*. Vol. 143. https://www.in.gov.br/web/dou/-/portaria-n-666-de-25-de-julho-de-2019-207244569.

Brettell, Caroline, and James Frank Hollifield, eds. 2023. *Migration Theory: Talking across Disciplines*. Fourth edition. New York, NY: Routledge.

Britannica, Editors of Encyclopaedia. 2023. 'Twitter'. In *Encyclopedia Britannica*. https://www.britannica.com/topic/Twitter.

Broersma, Marcel, and Todd Graham. 2013. 'TWITTER AS A NEWS SOURCE: How Dutch and British Newspapers Used Tweets in Their News Coverage, 2007–2011'. *Journalism Practice* 7 (4): 446–64. https://doi.org/10.1080/17512786.2013.802481.

Brum, Stéphany, Maria Isabel Santos Lima, and Naiara Sandi de Almeida Alcantara. 2021. 'Da Onda Rosa à Ascenção Do Conservadorismo: Análise Longitudinal Do Governo de Dilma Rousseff Ao de Jair Bolsonaro'. *Desigualdade e Ciência Política*, December. https://doi.org/10.5380/2dcp2021.artcomp12p323-353.

Brumat, Leiza, and Andrew Geddes. 2023. 'Refugee Recognition in Brazil under Bolsonaro: The Domestic Impact of International Norms and Standards'. *Third World Quarterly* 44 (3): 478–95. https://doi.org/10.1080/01436597.2022.2153664.

Bruns, Axel, and Stefan Stieglitz. 2012. 'Quantitative Approaches to Comparing Communication Patterns on Twitter'. *Journal of Technology in Human Services* 30 (3–4): 160–85. https://doi.org/10.1080/15228835.2012.744249.

Burgess, Jean Elizabeth, and Nancy K. Baym. 2020. *Twitter: A Biography*. New York: New York University Press.

Carnegie Europe. 2015. 'Understanding Central Europe's Opposition to Refugees'. 25 September 2015. /web/20230920132657/https://carnegieeurope.eu/strategiceurope/61404.

Carta Capital. 2022. 'TRF-2 Extingue Ação Contra Dilma Por Supostas Pedaladas Fiscais'. *Carta Capital* (blog). 28 March 2022. /web/20230714102748/https://www.cartacapital.com.br/cartaexpressa/trf-2-extingue-acao-contra-dilma-por-supostas-pedaladas-fiscais/.

Castells, Manuel. 2011. 'A Network Theory of Power'. *International Journal of Communication* 5 (April): 773–87.

Cavalcanti, Leonardo, T. Oliveira, and M. Macedo. 2020. 'Imigração e Refúgio No Brasil. Relatório Anual 2020.' Observatório das Migrações Internacionais; Ministério da Justiça e Segurança Pública/ Conselho Nacional de Imigração e Coordenação Geral de Imigração Laboral. https://portaldeimigracao.mj.gov.br/images/dados/relatorio-anual/2020/OBMigra_RELAT%C3%93RIO_ANUAL_2020.pdf.

Center for an Informed Public - University of Washington. 2023. 'Twitter's API Access Changes Could Mark "End of an Era" in Academic Research on the Platform'. Center for an Informed Public - University of Washington. 2 February 2023. https://web.archive.org/web/20230713123725/https://www.cip.uw.edu/2023/02/02/twitters-api-access-changes-academic-research/.

Chang, Hsia-Ching, and Hemalata Iyer. 2012. 'Trends in Twitter Hashtag Applications: Design Features for Value-Added Dimensions to Future Library Catalogues'. *Library Trends* 61 (1): 248–58. https://doi.org/10.1353/lib.2012.0024.

Chen, Adrian. 2021. 'Building a Data Stream to Assist with COVID-19 Research'. *Twitter Blog* (blog). 16 March 2021. https://web.archive.org/web/20230713131043/https://blog.twitter.com/engineering/en_us/topics/insights/2021/how-we-built-a-data-stream-to-assist-with-covid-19-research.

Chen, Yiyi, Harald Sack, and Mehwish Alam. 2021. 'MigrationsKB: A Knowledge Base of Public Attitudes towards Migrations and Their Driving Factors'. *ArXiv*. https://doi.org/10.48550/arXiv.2108.07593.

———. 2022. 'Analyzing Social Media for Measuring Public Attitudes toward Controversies and Their Driving Factors: A Case Study of Migration'. *Social Network Analysis and Mining* 12 (1). https://doi.org/10.1007/s13278-022-00915-7.

Clayton, Jonathan, and Hereward Holland. 2015. 'Over One Million Sea Arrivals Reach Europe in 2015'. *UNHCR News* (blog). 10 December 2015. /web/20230713152345/https://www.unhcr.org/news/stories/over-one-million-sea-arrivals-reach-europe-2015.

Conectas. 2019. 'Você Sabe o Caminho Que a Lei de Migração Percorreu Para Ser Aprovada?' *Conectas - Direitos Humanos* (blog). 18 September 2019. /web/20230714114112/https://www.conectas.org/noticias/voce-sabe-o-caminho-que-a-lei-de-migracao-percorreu-para-ser-aprovada/.

Cruz, Edgar Gómez, and Ramaswami Harindranath. 2020. 'WhatsApp as "Technology of Life": Reframing Research Agendas'. *First Monday* 25 (January). https://firstmonday.org/ojs/index.php/fm/article/download/10405/8318.

Data Reportal. 2023a. 'Global Social Media Statistics'. *Data Reportal* (blog). 2023. /web/20230714111547/https://datareportal.com/social-media-users.

———. 2023b. 'Twitter Users, Stats and Trends'. *Data Reportal* (blog). 11 May 2023. /web/20230714111818/https://datareportal.com/essential-twitter-stats.

De Melo, Tiago, and Carlos M.S. Figueiredo. 2020. 'A First Public Dataset from Brazilian Twitter and News on COVID-19 in Portuguese'. *Data in Brief* 32 (October): 106179. https://doi.org/10.1016/j.dib.2020.106179.

Dijck, José Van. 2011. 'Tracing Twitter: The Rise of a Microblogging Platform'. *International Journal of Media & Cultural Politics* 7 (3): 333–48. https://doi.org/10.1386/macp.7.3.333_1.

Dijk, Teun A. van. 2017. 'How Globo Media Manipulated the Impeachment of Brazilian President Dilma Rousseff'. *Discourse & Communication* 11 (2): 199–229. https://doi.org/10.1177/1750481317691838.

Dixon, S. 2022. 'Number of Twitter Users Worldwide from 2019 to 2024 (in Millions)'. Statista. 14 December 2022. https://www.statista.com/statistics/303681/twitter-users-worldwide/.

Earl, Jennifer, Heather McKee Hurwitz, Analicia Mejia Mesinas, Margaret Tolan, and Ashley Arlotti. 2013. 'THIS PROTEST WILL BE TWEETED: Twitter and Protest Policing during the Pittsburgh G20'. *Information, Communication & Society* 16 (4): 459–78. https://doi.org/10.1080/1369118X.2013.777756.

Edwards, Richard, and Robin Usher. 2007. *Globalisation & Pedagogy*. 0 ed. Routledge. https://doi.org/10.4324/9780203945001.

Exame. 2018. 'Bolsonaro Quer Campo de Refugiados Para Venezuelanos Em Roraima'. *Exame* (blog). 14 March 2018. https://web.archive.org/web/20230713151149/https://exame.com/brasil/bolsonaro-quer-campo-de-refugiados-em-roraima/.

Eysenbach, Gunther. 2011. 'Can Tweets Predict Citations? Metrics of Social Impact Based on Twitter and Correlation with Traditional Metrics of Scientific Impact'. *Journal of Medical Internet Research* 13 (4): e123. https://doi.org/10.2196/jmir.2012.

Fang, Xing, and Justin Zhan. 2015. 'Sentiment Analysis Using Product Review Data'. *Journal of Big Data* 2 (1): 5. https://doi.org/10.1186/s40537-015-0015-2.

Ferawati, Kiki, Kongmeng Liew, Eiji Aramaki, and Shoko Wakamiya. 2022. 'Monitoring Mentions of COVID-19 Vaccine Side Effects from Japanese and Indonesian Twitter: Infodemiological Study (Preprint)'. *JMIR Infodemiology* 2 (2): e39504–e39504. https://doi.org/10.2196/39504.

Firdaniza, Firdaniza, Budi Ruchjana, Diah Chaerani, and Jaziar Radianti. 2021. 'Information Diffusion Model in Twitter: A Systematic Literature Review'. *Information* 13 (1): 13. https://doi.org/10.3390/info13010013.

Fischer, Eileen, and A. Rebecca Reuber. 2011. 'Social Interaction via New Social Media: (How) Can Interactions on Twitter Affect Effectual Thinking and Behavior?' *Journal of Business Venturing* 26 (1): 1–18. https://doi.org/10.1016/j.jbusvent.2010.09.002.

Flew, Terry, Axel Bruns, Jean Burgess, Kate Crawford, and Frances Shaw. 2014. 'Social Media and Its Impact on Crisis Communication: Case Studies of Twitter Use in Emergency Management in Australia and New Zealand'. In . https://eprints.qut.edu.au/63707/.

Freire-Vidal, Yerka, Eduardo Graells-Garrido, and Francisco Rowe. 2021. 'A Framework to Understand Attitudes towards Immigration through Twitter'. *Applied Sciences* 11 (20): 9689. https://doi.org/10.3390/app11209689.

Freitas Castro, Emília Lana de. 2020. 'The Legal Framework of Migration in Brazil and in MERCOSUL', January, 67–125. https://doi.org/10.1007/978-3-030-46608-4_4.

Galvão, Thiago Gehre, Bruna de Paula Miranda Pereira, and Maurício Kenyatta Barros da Costa. 2021. 'The Future of Brazil-Venezuela Partnership under a Migration Crisis Context: Confidence Building Measures by Border Understandings'. *Monções* 10 (20): 51–79. https://doi.org/10.30612/rmufgd.v10i20.14706.

Georgi, Fabian. 2019. 'The Role of Racism in the European "Migration Crisis": A Historical Materialist Perspective.' In *Racism After Apartheid: Challenges for Marxism and Anti-Racism*, edited by Vishwas Satgar, 96–117. Wits University Press. https://doi.org/10.18772/22019033061.

Golbeck, Jennifer, Justin M. Grimes, and Anthony Rogers. 2010. 'Twitter Use by the U.S. Congress'. *Journal of the American Society for Information Science and Technology*, n/a-n/a. https://doi.org/10.1002/asi.21344.

Goldstein, Ariel Alejandro. 2016. 'The Contribution of the Liberal-Conservative Press to the Crisis of Dilma Rousseff's Second Term'. Edited by Claudia Alvares. *Cogent Social Sciences* 2 (1): 1253202. https://doi.org/10.1080/23311886.2016.1253202.

Gomes Pessanha, Gabriel Rodrigo, Taylor Oliveira Fidelis, Clarissa Dourado Freire, and Eduardo Almeida Soares. 2020. '#FIQUEEMCASA: ANÁLISE DE SENTIMENTO DOS USUÁRIOS DO TWITTER EM RELAÇÃO AO COVID19'. *HOLOS* 5 (December): 1–20. https://doi.org/10.15628/holos.2020.11147.

Gottschalk, Thilo, and Francesca Pichierri. 2022. 'About Migration Flows and Sentiment Analysis on Twitter Data: Building the Bridge between Technical and Legal Approaches to Data Protection'. *LEGAL*.

Gupta, Prasoon, Sanjay Kumar, Sanjay Kumar, Sanjay Kumar, Rajiv Ranjan Suman, R. R. Suman, Vinay Kumar, Vinod Kumar, and Vinay Kumar. 2021. 'Sentiment Analysis of Lockdown in India During COVID-19: A Case Study on Twitter'. *IEEE Transactions on Computational Social Systems* 8 (4): 992–1002. https://doi.org/10.1109/tcss.2020.3042446.

Haas, Hein de, Stephen Castles, and Mark J. Miller. 2022. *The Age of Migration: International Population Movements in the Modern World*. Sixth edition, Reprinted by Bloomsbury Academic. London: Bloomsbury Academic.

Handerson, Joseph. 2015. 'DIASPORA. AS DINÂMICAS DA MOBILIDADE HAITIANA NO BRASIL, NO SURINAME E NA GUIANA FRANCESA'. Rio de Janeiro: Universidade Federal do Rio de Janeiro. https://www.migrante.org.br/wp-content/uploads/2015/05/Diaspora_Haitiana_tese_Joseph_Handerson.pdf.

Hardeniya, Nitin. 2015. *NLTK Essentials: Build Cool NLP and Machine Learning Applications Using NLTK and Other Python Libraries*. Packt Open Source. Birmingham: Packt Publ.

Hartwig, Samuel. 2020. 'Frontex: From Coordinating Controls to Combating Crime'. *Eucrim - The European Criminal Law Associations' Forum*. https://doi.org/10.30709/eucrim-2020-010.

Hashtag Org. 2012. 'What Is a (#) Hashtag?' 24 June 2012. https://www.hashtags.org/.

Haustein, Stefanie, Timothy D. Bowman, Kim Holmberg, Andrew Tsou, Cassidy R. Sugimoto, and Vincent Larivière. 2016. 'Tweets as Impact Indicators: Examining the Implications of Automated "Bot" Accounts on Twitter: Tweets as Impact Indicators: Examining the Implications of Automated "Bot" Accounts on Twitter'. *Journal of the Association for Information Science and Technology* 67 (1): 232–38. https://doi.org/10.1002/asi.23456.

Hermida, Alfred. 2010. 'TWITTERING THE NEWS: The Emergence of Ambient Journalism'. *Journalism Practice* 4 (3): 297–308. https://doi.org/10.1080/17512781003640703.

Holmberg, Kim, and Mike Thelwall. 2014. 'Disciplinary Differences in Twitter Scholarly Communication'. *Scientometrics* 101 (2): 1027–42. https://doi.org/10.1007/s11192-014-1229-3.

Horstmann, Cay S., and Rance D. Necaise. 2016. *Python for Everyone*. 2nd edition. Hoboken, NJ: John Wiley & Sons.

Hovy, Dirk. 2020. *Text Analysis in Python for Social Scientists: Discovery and Exploration*. Cambridge Elements Elements in Quantitative and Computational Methods for the Social Sciences. Cambridge New York Port Melbourne, Australia New Delhi, India Singapore: Cambridge University Press.

Huang, Binxuan, and Kathleen M. Carley. 2019. 'A Large-Scale Empirical Study of Geotagging Behavior on Twitter'. In *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 365–73. Vancouver British Columbia Canada: ACM. https://doi.org/10.1145/3341161.3342870.

Hughes, Amanda Lee, and Leysia Palen. 2009. 'Twitter Adoption and Use in Mass Convergence and Emergency Events'. *International Journal of Emergency Management* 6 (3/4): 248. https://doi.org/10.1504/IJEM.2009.031564.

Hutto, C., and Eric Gilbert. 2014. 'VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text'. *Proceedings of the International AAAI Conference on Web and Social Media* 8 (1): 216–25. https://doi.org/10.1609/icwsm.v8i1.14550.

Iasulaitis, Sylvia, and Isabella Vicari. 2021. 'The Salience of Traditional Moral Values: Bolsonaro's Electoral Competition Strategy on Twitter'. *International Journal of Social Science Studies* 9 (5): 153–65. https://doi.org/10.11114/ijsss.v9i5.5313.

Inuwa-Dutse, Isa, Mark Liptrott, and Ioannis Korkontzelos. 2020. 'Migration and Refugee Crisis: A Critical Analysis of Online Public Perception.' *ArXiv*, July. https://doi.org/10.48550/arXiv.2007.09834.

Jagdale, Rajkumar S., Vishal S. Shirsat, and Sachin N. Deshmukh. 2019. 'Sentiment Analysis on Product Reviews Using Machine Learning Techniques'. In *Cognitive Informatics and Soft Computing*, edited by Pradeep Kumar Mallick, Valentina Emilia Balas, Akash Kumar Bhoi, and Ahmed F. Zobaa, 768:639–47. Advances in Intelligent Systems and Computing. Singapore: Springer Singapore. https://doi.org/10.1007/978-981-13-0617-4_61.

Jubilut, Liliana Lyra. 2006. 'Refugee Law and Protection in Brazil: A Model in South America?' *Journal of Refugee Studies* 19 (1): 22–44. https://doi.org/10.1093/jrs/fej006.

Jubilut, Liliana Lyra, and Gabriel Gualano de Godoy, eds. 2017. *Refúgio No Brasil: Comentários à Lei 9.474/97*. Quartier Latin/ACNUR. Sao Paulo. https://www.acnur.org/portugues/wp-content/uploads/2018/02/Ref%C3%BAgio-no-Brasil-Coment%C3%A1rios-%C3%A0-lei-9.474-97-2017.pdf.

Kantrowitz, Alex. 2019. 'The Man Who Built The Retweet: "We Handed A Loaded Weapon To 4-Year-Olds"'. *BuzzFeed News* (blog). 23 July 2019. https://web.archive.org/web/20230713122101/https://www.buzzfeednews.com/article/alexkantrowitz/how-the-retweet-ruined-the-internet.

Kapoteli, Evridiki, Vasiliki Chouliara, Paraskevas Koukaras, and Christos Tjortjis. 2022. 'Social Media Sentiment Analysis Related to COVID-19 Vaccinations'. *Intelligent Systems Reference Library*, September, 47–69. https://doi.org/10.1007/978-3-031-11170-9_3.

Kingsley, Patrick. 2015. 'Refugee Crisis: Apart from Syrians, Who Is Travelling to Europe?' 10 September 2015. https://web.archive.org/web/20230713152128/https://www.theguardian.com/world/2015/sep/10/refugee-crisis-apart-from-syrians-who-else-is-travelling-to-europe.

Klar, Samara, Yanna Krupnikov, John Barry Ryan, Kathleen Searles, and Yotam Shmargad. 2020. 'Using Social Media to Promote Academic Research: Identifying the Benefits of Twitter for Sharing Academic Work'. Edited by Jarosław Jankowski. *PLOS ONE* 15 (4): e0229446. https://doi.org/10.1371/journal.pone.0229446.

Kumar, Pritish. 2011. 'Twitter Now with Integrated Photo-Sharing Service and Completely New Twitter Search'. 1 June 2011. https://web.archive.org/web/20110630083417/http://techshrimp.com/2011/06/01/twitter-now-with-integrated-photo-sharing-service-and-completely-new-twitter-search/.

Kwak, Haewoon, Changhyun Lee, Hosung Park, and Sue Moon. 2010. 'What Is Twitter, a Social Network or a News Media?' In *Proceedings of the 19th International Conference on World Wide Web*, 591–600. Raleigh North Carolina USA: ACM. https://doi.org/10.1145/1772690.1772751.

Langlois, Jill. 2016. 'The Syrians Starting over in Brazil'. *UNHCR News* (blog). 26 March 2016. /web/20230714113517/https://www.unhcr.org/news/stories/syrians-starting-over-brazil.

Lasorsa, Dominic L., Seth C. Lewis, and Avery E. Holton. 2012. 'NORMALIZING TWITTER: Journalism Practice in an Emerging Communication Space'. *Journalism Studies* 13 (1): 19–36. https://doi.org/10.1080/1461670X.2011.571825.

Luo, Zhunchen, Miles Osborne, and Ting Wang. 2015. 'An Effective Approach to Tweets Opinion Retrieval'. *World Wide Web* 18 (3): 545–66. https://doi.org/10.1007/s11280-013-0268-7.

M. Botelho, Mariane Motta De Campos, Mayra Regina Coimbra, and Deborah Luísa Vieira Dos Santos. 2022. 'Discursos e Redes Anticorrupção No Twitter: "Bolsonaro Incorruptível"'. *Liinc Em Revista*. https://doi.org/10.18617/liinc.v18i2.6003.

Manfrinato, Helena de Morais, and Mirian Alves de Souza. 2019. 'Refugees of the Syrian Conflict and the Struggle for Housing in Brazil'. *Refugee Outreach and Research Network* (blog). 23 March 2019.

Marcec, Robert, and Robert Likić. 2021. 'Using Twitter for Sentiment Analysis towards AstraZeneca/Oxford, Pfizer/BioNTech and Moderna COVID-19 Vaccines.' *Postgraduate Medical Journal*, August. https://doi.org/10.1136/postgradmedj-2021-140685.

Marinho, Helder. 2010. 'Brazil's Lula Leaves Office With 83% Approval Rating, Folha Says'. *Bloomberg* (blog). 19 December 2010. /web/20230714094121/https://www.bloomberg.com/news/articles/2010-12-19/brazil-s-lula-leaves-office-with-83-approval-rating-folha-says.

Martino, Andressa A., and Julia Bertino Moreira. 2020. 'A Política Migratória Brasileira Para Venezuelanos: Do "Rótulo" Da Autorização de Residência Temporária Ao Do Refúgio (2017-2019)' 28 (60): 151–66. https://doi.org/10.1590/1980-85852503880006009.

Marwick, Alice, and Danah Boyd. 2011. 'To See and Be Seen: Celebrity Practice on Twitter'. *Convergence: The International Journal of Research into New Media Technologies* 17 (2): 139–58. https://doi.org/10.1177/1354856510394539.

Massey, Douglas S., Joaquin Arango, Graeme Hugo, Ali Kouaouci, Adela Pellegrino, and J. Edward Taylor. 1993. 'Theories of International Migration: A Review and Appraisal'. *Population and Development Review* 19 (3): 431. https://doi.org/10.2307/2938462.

Masuda, Hisashi, Spring H. Han, and Jungwoo Lee. 2022. 'Impacts of Influencer Attributes on Purchase Intentions in Social Media Influencer Marketing: Mediating Roles of Characterizations'. *Technological Forecasting and Social Change* 174 (January): 121246. https://doi.org/10.1016/j.techfore.2021.121246.

Maynard, Diana, and Mark A. Greenwood. 2014. *Who Cares about Sarcastic Tweets? Investigating the Impact of Sarcasm on Sentiment Analysis*. Erscheinungsort nicht ermittelbar: European Language Resources Association (ELRA). https://www.researchgate.net/publication/325193763_Who_cares_about_sarcastic_tweets_Investigating_the_impact_of_sarcasm_on_sentiment_analysis.

McClain, Colleen, Regina Widjaya, Gonzalo Rivero, and Aaron Smith. 2021. 'The Behaviors and Attitudes of U.S. Adults on Twitter: Comparing Highly Active and Less Active Tweeters COLLEEN MCCLAIN, REGINA WIDJAYA, GONZALO RIVERO AND AARON SMITH'. *Pwe Research Center* (blog). 15 November 2021.

https://web.archive.org/web/20230713095223/https://www.pewresearch.org/internet/2021/1
1/15/2-comparing-highly-active-and-less-active-tweeters/.

Mello, Anna Carolina Raposo de. 2019. 'Brazilian Foreign Policy on Twitter: Digital Expression of
Attitudes in the Early Months of Bolsonaro\'s Administration', December.
https://doi.org/10.11606/d.101.2020.tde-14022020-104934.

Mendes, José Sacchetta Ramos, and Fábio Bensabath Bezerra de Menezes. 2019. 'POLÍTICA
MIGRATÓRIA NO BRASIL DE JAIR BOLSONARO: "PERIGO ESTRANGEIRO" E
RETORNO À IDEOLOGIA DE SEGURANÇA NACIONAL', no. 247 (December): 302–21.
https://doi.org/10.25247/2447-861x.2019.n247.p302-321.

Messina, Chris. 2007. 'Groups for Twitter; or A Proposal for Twitter Tag Channels'. 25 August 2007.
https://web.archive.org/web/20230713081056/https://factoryjoe.com/2007/08/25/groups-
for-twitter-or-a-proposal-for-twitter-tag-channels/.

Milesi, Rosita. 2016. 'Haitianos No Brasil: Dados Estatísticos, Informações e Uma Recomendação'.
Instituto Migrações e Direitos Humanos (blog). 30 December 2016.
/web/20230714091221/https://www.migrante.org.br/migracoes/migracao-haitiana/haitianos-
no-brasil-dados-estatisticos-informacoes-e-uma-recomendacao/.

Mohammadi, Ehsan, Mike Thelwall, Mary Kwasny, and Kristi L. Holmes. 2018. 'Academic
Information on Twitter: A User Survey'. Edited by Sergi Lozano. *PLOS ONE* 13 (5):
e0197265. https://doi.org/10.1371/journal.pone.0197265.

Moraes, Marcela, John Gountas, Sandra Gountas, and Piyush Sharma. 2019. 'Celebrity Influences on
Consumer Decision Making: New Insights and Research Directions'. *Journal of Marketing
Management* 35 (13–14): 1159–92. https://doi.org/10.1080/0267257X.2019.1632373.

MSF. 2016. 'EU Migration Crisis Update - May 2016'. *Medecins Sans Frontieres* (blog). 6 May 2016.
/web/20230920132344/https://www.msf.org/eu-migration-crisis-update-may-2016.

Müller-Prove, Matthias. 2008. 'Modell und Anwendungsperspektive des Social Tagging'.
https://doi.org/10.25656/01:12757.

Naseem, Usman, Muhammad Imran Razzak, Imran Razzak, Imran Razzak, Matloob Khushi, Peter
W. Eklund, Jinman Kim, and Jinman Kim. 2021. 'COVIDSenti: A Large-Scale Benchmark
Twitter Data Set for COVID-19 Sentiment Analysis'. *IEEE Transactions on Computational
Social Systems* 8 (4): 1003–15. https://doi.org/10.1109/tcss.2021.3051189.

Niño, Edgar Andrés Londoño, and Edgar Andrés Londoño Niño. 2018. 'Questão de Segurança Ou
de Direitos Humanos? A Imigração Venezuelana e as Mudanças Na Política Externa
Brasileira | Security or Human Rights Issue? Venezuelan Immigration and Changes of
Brazilian Foreign Policy' 9 (1): 122–37. https://doi.org/10.12957/rmi.2018.36049.

Nunes, Felipe, and Carlos Ranulfo Melo. 2017. 'Impeachment, Political Crisis and Democracy in
Brazil'. *Revista de Ciencia Politica* 37 (2). https://www.scielo.cl/pdf/revcipol/v37n2/0718-
090X-revcipol-37-02-0281.pdf.

OAS. 1984. *Cartagena Declaration on Refugees*.
https://www.oas.org/dil/1984_cartagena_declaration_on_refugees.pdf.

Oliveira, Danielle Sant'Ana de Albuquerque Oliveira. 2018. 'OS AVANÇOS E RETROCESSOS
DA POLÍTICA MIGRATÓRIA BRASILEIRA NA CRIAÇÃO DA LEI Nº 13.455/2017 E

DO SEU DECRETO REGULAMENTADOR Nº 9.199/2017'. Universidade de Brasilia. https://bdm.unb.br/bitstream/10483/23464/1/2018_DanielleSantanaDeAlbuquerqueOliveira_tcc.pdf.

Osava, Mario. 2018. 'Wars, Crises and Catastrophes Drive Immigration to Brazil'. *InterPress Service* (blog). 2 February 2018. /web/20230714113847/http://www.ipsnews.net/2018/02/wars-crises-catastrophes-drive-immigration-brazil/.

Otte, Evelien, and Ronald Rousseau. 2002. 'Social Network Analysis: A Powerful Strategy, Also for the Information Sciences'. *Journal of Information Science* 28 (6): 441–53. https://doi.org/10.1177/016555150202800601.

Paiva, André Luiz De, André Spuri Garcia, and Valderí De Castro Alcântara. 2017. 'Disputas Discursivas Sobre Corrupção No Brasil: Uma Análise Discursivo-Crítica No Twitter'. *Revista de Administração Contemporânea* 21 (5): 627–47. https://doi.org/10.1590/1982-7849rac2017160163.

Panko, Ben. 2017. 'A Decade Ago, the Hashtag Reshaped the Internet'. *Smithsonian Magazine* (blog). 23 August 2017. https://web.archive.org/web/20230713083335/https://www.smithsonianmag.com/smart-news/decade-ago-hashtag-reshaped-internet-180964605/.

Panzarino, Matthew. 2013. 'The Three Reasons Twitter Didn't Sell To Facebook'. 4 November 2013. https://web.archive.org/web/20230712163545/https://techcrunch.com/web/20230712163545/https://techcrunch.com/2013/11/04/the-three-reasons-twitter-didnt-sell-to-facebook/.

Pariser, Eli. 2012. *O filtro invisível: O que a internet está escondendo de você*. 1st ed. Zahar.

Patodkar, Vaibhavi N and Sheikh I.R. 2016. 'Twitter as a Corpus for Sentiment Analysis and Opinion Mining'. *International Journal of Advanced Research in Computer and Communication Engineering* 5 (12): 320–22. https://doi.org/10.17148/ijarcce.2016.51274.

Recuero, Raquel. 2015. *Análise de Redes Para Mídia Social*. Porto Alegre, RS: Editora Sulina.

Reenlsober, Danilo, and Issaaf Karhawi. 2022. 'ETHOS NAS REDES: OS SEGUIDORES DE JAIR BOLSONARO NO TWITTER'. *Revista GEMInIS* 13 (1): 166–87. https://doi.org/10.53450/2179-1465.rg.2022v13i1p166-187.

Reis, Rossana Rocha. 2011. 'A Política Do Brasil Para as Migrações Internacionais'. *Contexto Internacional* 33 (1): 47–69. https://doi.org/10.1590/S0102-85292011000100003.

Reuters. 2021. 'Brazil Supreme Court Confirms Ruling That Judge Was Biased against Lula'. *Reuters* (blog). 24 June 2021. /web/20230714102420/https://www.reuters.com/world/americas/brazil-supreme-court-confirms-ruling-that-judge-was-biased-against-lula-2021-06-23/.

Ribeiro, Filipe N., Matheus Araújo, Pollyanna Gonçalves, Marcos André Gonçalves, and Fabrício Benevenuto. 2016. 'SentiBench - a Benchmark Comparison of State-of-the-Practice Sentiment Analysis Methods'. *EPJ Data Science* 5 (1): 23. https://doi.org/10.1140/epjds/s13688-016-0085-1.

Rodrigues, Theófilo Codeço Machado, and Theófilo Codeço Machado Rodrigues. 2018. 'O Papel Da Mídia Nos Processos de Impeachment de Dilma Rousseff (2016) e Michel Temer (2017)'. *Revista Contracampo* 37 (2). https://doi.org/10.22409/contracampo.v0i0.1108.

Ross, C., M. Terras, C. Warwick, and A. Welsh. 2011. 'Enabled Backchannel: Conference Twitter Use by Digital Humanists'. *Journal of Documentation* 67 (2): 214–37. https://doi.org/10.1108/00220411111109449.

Rowe, Francisco, Michael Mahony, Eduardo Graells-Garrido, Marzia Rango, and Niklas Sievers. 2021. 'Using Twitter Data to Monitor Immigration Sentiment', March. https://doi.org/10.17605/osf.io/x47qa.

Santos, Daiana Kathrin Santana. 2023. 'Análise de sentimentos dos usuários do Twitter sobre a eleição presidencial de 2022 no Brasil'. Sao Paulo: Universidade Federal de Sao Paulo. https://repositorio.unifesp.br/11600/67118.

Santos, Fabiano, José Szwako, Universidade Estadual do Rio de Janeiro, Brazil, and Universidade Estadual do Rio de Janeiro, Brazil. 2016. 'Da Ruptura à Reconstrução Democrática No Brasil'. *Saúde Em Debate* 40 (spe): 114–21. https://doi.org/10.1590/0103-11042016s10.

Santos, Lucas L., Reinaldo A. C. Bianchi, and Anna H. R. Costa. 2023. 'FinBERT-PT-BR: Análise de Sentimentos de Textos Em Português Do Mercado Financeiro'. In *Anais Do II Brazilian Workshop on Artificial Intelligence in Finance (BWAIF 2023)*, 144–55. Brasil: Sociedade Brasileira de Computação. https://doi.org/10.5753/bwaif.2023.231151.

Sassen, Saskia. 2007. 'A Sociology of Globalization', January.

Sato, Mia. 2022. 'Buying Twitter "Is Not a Way to Make Money," Says Musk in TED Interview'. *The Verge* (blog). 14 April 2022. https://web.archive.org/web/20220414173256/https://www.theverge.com/2022/4/14/230253 43/elon-musk-twitter-takeover-ted-talk-quote-stock-buyout.

Scott, John. 2012. *What Is Social Network Analysis?* Online-Ausg. London: Bloomsbury Academic.

Scott, Kate. 2015. 'The Pragmatics of Hashtags: Inference and Conversational Style on Twitter'. *Journal of Pragmatics* 81 (May): 8–20. https://doi.org/10.1016/j.pragma.2015.03.015.

Sharma, Dipti, Munish Sabharwal, Vinay Goyal, and Mohit Vij. 2020. 'Sentiment Analysis Techniques for Social Media Data: A Review'. In *First International Conference on Sustainable Technologies for Computational Intelligence*, edited by Ashish Kumar Luhach, Janos Arpad Kosa, Ramesh Chandra Poonia, Xiao-Zhi Gao, and Dharm Singh, 1045:75–90. Advances in Intelligent Systems and Computing. Singapore: Springer Singapore. https://doi.org/10.1007/978-981-15-0029-9_7.

Shi, Juan, Kin Keung Lai, Ping Hu, and Gang Chen. 2017. 'Understanding and Predicting Individual Retweeting Behavior: Receiver Perspectives'. *Applied Soft Computing* 60 (November): 844–57. https://doi.org/10.1016/j.asoc.2017.08.044.

Shuai, Xin, Alberto Pepe, and Johan Bollen. 2012. 'How the Scientific Community Reacts to Newly Submitted Preprints: Article Downloads, Twitter Mentions, and Citations'. Edited by Christos A. Ouzounis. *PLoS ONE* 7 (11): e47523. https://doi.org/10.1371/journal.pone.0047523.

Siapera, Eugenia, Moses Boudourides, Sergios Lenis, and Jane Suiter. 2018. 'Refugees and Network Publics on Twitter: Networked Framing, Affect, and Capture'. *Social Media + Society* 4 (1): 205630511876443. https://doi.org/10.1177/2056305118764437.

Simões, A., J. Hallak Neto, L. Cavalcanti, T. Oliveira, and M. Macedo. 2020. *Relatório Conjuntural: Tendências Da Imigração e Refúgio No Brasil, 1º Quadrimestre/2020*. Brasília, DF:

Observatório das Migrações Internacionais; Ministério da Justiça e Segurança Pública/ Coordenação Geral de Imigração Laboral. https://portaldeimigracao.mj.gov.br/pt/dados.

Souza, Jessé. 2016. *A Radiografia Do Golpe: Entenda Como e Por Que Você Foi Enganado*. São Paulo, SP: LeYa.

Spyer, Juliano. 2017. *Social Media in Emergent Brazil: How the Internet Affects Social Mobility*. UCL Press. https://doi.org/10.2307/j.ctt1wc7rdn.

Ta'amneh, Issam Mostafa, and Abeer Al-Ghazo. 2021. 'The Importance of Using Hashtags on Raising Awareness about Social Issues'. *International Journal of Learning and Development* 11 (4): 10. https://doi.org/10.5296/ijld.v11i4.19139.

Tatagiba, Luciana. 2018. 'Entre as Ruas e as Instituições: Os Protestos e o Impeachment de Dilma Rousseff'. *Lusotopie* 17 (1): 112–35. https://doi.org/10.1163/17683084-12341714.

TechTarget. n.d. 'What Is an API Endpoint?' *TechTarget* (blog). Accessed 2 June 2023. https://www.techtarget.com/searchapparchitecture/definition/API-endpoint#:~:text=An%20API%20endpoint%20is%20a,server%20and%20receiving%20a%20response.

Terdiman, Daniel. 2007. 'To Twitter or Dodgeball at SXSW?' CNET. 10 March 2007. https://web.archive.org/web/20131203224143/http://news.cnet.com/8301-17939_109-9696264-2.html.

Theocharis, Yannis, Will Lowe, Jan W. Van Deth, and Gema García-Albacete. 2015. 'Using Twitter to Mobilize Protest Action: Online Mobilization Patterns and Action Repertoires in the Occupy Wall Street, Indignados, and Aganaktismenoi Movements'. *Information, Communication & Society* 18 (2): 202–20. https://doi.org/10.1080/1369118X.2014.948035.

Toh, Michelle, and Juliana Liu. 2023. 'Elon Musk Says He's Cut about 80% of Twitter's Staff'. *CNN* (blog). 12 April 2023. https://web.archive.org/web/20230712164656/https://edition.cnn.com/2023/04/12/tech/elon-musk-bbc-interview-twitter-intl-hnk/index.html.

Tonhati, Tânia, Leonardo Cavalcanti, and Antônio Tadeu De Oliveira. 2022. 'A Decade of Growth in Migration in Brazil (2010–2020) and the Impact of the COVID-19 Pandemic'. In *Migration in South America*, edited by Gioconda Herrera and Carmen Gómez, 99–120. IMISCOE Research Series. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-11061-0_5.

Tornes, Adam. 2020. 'Enabling Study of the Public Conversation in a Time of Crisis'. *Twitter Developer Platform Blog* (blog). 24 April 2020. https://web.archive.org/web/20230713131421/https://blog.twitter.com/developer/en_us/topics/tools/2020/covid19_public_conversation_data.

———. 2021. 'Enabling the Future of Academic Research with the Twitter API'. *Twitter Developer Platform* (blog). 26 January 2021.

Trevizan, Karina. 2017. 'Brasil Enfrenta Pior Crise Já Registrada Poucos Anos Após Um Boom Econômico'. *G1* (blog). 7 March 2017. /web/20230714085645/https://g1.globo.com/economia/noticia/brasil-enfrenta-pior-crise-ja-registrada-poucos-anos-apos-um-boom-economico.ghtml.

Twitter Help Centre. n.d. 'About Replies and Mentions'. Twitter Help Centre. Accessed 10 June 2023a. https://help.twitter.com/en/using-twitter/mentions-and-replies#:~:text=A%20mention%20is%20a%20Tweet,Tweet%20in%20their%20Notifications%20tab.

———. n.d. 'How to Retweet'. Twitter Help Centre. Accessed 10 June 2023b. https://help.twitter.com/en/using-twitter/how-to-retweet.

Twitter News. 2022. 'How Many People Come to Twitter for News? As It Turns out, a LOT'. *Twitter News* (blog). 12 September 2022. https://web.archive.org/web/20230712172650/https://blog.twitter.com/en_us/topics/insights/2022/how-many-people-come-twitter-for-news.

Uebel, Roberto Rodolfo Georg. 2018. 'Política externa migratória brasileira : das migrações de perspectiva à hiperdinamização das migrações durante os governos Lula da Silva e Dilma Rousseff'. Universidade Federal do Rio Grande do Sul. https://lume.ufrgs.br/handle/10183/188410.

Uebel, Roberto Rodolfo Georg, and Sonia Ranincheski. 2017. 'Pontes Ou Muros? As Diferentes Ações Dos Governos de Lula Da Silva, Dilma Rousseff e Michel Temer Em Relação Às Migrações Internacionais Para o Território Brasileiro' 16 (2).

UNHCR. 2022. 'Syria Refugee Crisis Explained'. *UNHCR News* (blog). 8 July 2022. /web/20230714112044/https://www.unrefugees.org/news/syria-refugee-crisis-explained/.

United Nations. 2018. *World Migration Report*. Geneva: UNITED NATIONS PUBNS.

———. 2022. 'UN Marks Anniversary of Devastating 2010 Haiti Earthquake'. *UN News* (blog). 12 January 2022. /web/20230714090604/https://news.un.org/en/story/2022/01/1109632.

Van Dijck, José, and Thomas Poell. 2013. 'Understanding Social Media Logic'. *Media and Communication* 1 (1): 2–14. https://doi.org/10.17645/mac.v1i1.70.

Vigevani, Tullo, and Gabriel Cepaluni. 2007. 'Lula's Foreign Policy and the Quest for Autonomy through Diversification'. *Third World Quarterly* 28 (7): 1309–26. https://doi.org/10.1080/01436590701547095.

Waters, Richard D., and Jensen M. Williams. 2011. 'Squawking, Tweeting, Cooing, and Hooting: Analyzing the Communication Patterns of Government Agencies on Twitter: Models of Public Relations on Twitter'. *Journal of Public Affairs* 11 (4): 353–63. https://doi.org/10.1002/pa.385.

Webb, Helena, Marina Jirotka, Bernd Carsten Stahl, William Housley, Adam Edwards, Matthew Williams, Rob Procter, Omer Rana, and Pete Burnap. 2017. 'The Ethical Challenges of Publishing Twitter Data for Research Dissemination'. In *Proceedings of the 2017 ACM on Web Science Conference*, 339–48. Troy New York USA: ACM. https://doi.org/10.1145/3091478.3091489.

Williams, Shirley A., Melissa M. Terras, and Claire Warwick. 2013. 'What Do People Study When They Study Twitter? Classifying Twitter Related Academic Papers'. *Journal of Documentation* 69 (3): 384–410. https://doi.org/10.1108/JD-03-2012-0027.

World Migration Report. 2020. Geneva: International Organization for Migration (IOM).

———. 2022. International Organization for Migration (IOM).

Wu, Shaomei, Jake M. Hofman, Winter A. Mason, and Duncan J. Watts. 2011. 'Who Says What to Whom on Twitter'. In *Proceedings of the 20th International Conference on World Wide Web*, 705–14. Hyderabad India: ACM. https://doi.org/10.1145/1963405.1963504.

Xiong, Ying, Moonhee Cho, and Brandon Boatwright. 2019. 'Hashtag Activism and Message Frames among Social Movement Organizations: Semantic Network Analysis and Thematic Analysis of Twitter during the #MeToo Movement'. *Public Relations Review* 45 (1): 10–23. https://doi.org/10.1016/j.pubrev.2018.10.014.

Yang, Guobin. 2016. 'Narrative Agency in Hashtag Activism: The Case of #BlackLivesMatter'. *Media and Communication* 4 (4): 13–17. https://doi.org/10.17645/mac.v4i4.692.

Yang, Jiang, and Scott Counts. 2010. 'Predicting the Speed, Scale, and Range of Information Diffusion in Twitter'. *Proceedings of the International AAAI Conference on Web and Social Media* 4 (1): 355–58. https://doi.org/10.1609/icwsm.v4i1.14039.

Yantseva, Victoria. 2020. 'Migration Discourse in Sweden: Frames and Sentiments in Mainstream and Social Media': *Social Media and Society* 6 (4): 205630512098105. https://doi.org/10.1177/2056305120981059.

Yardi, Sarita, and Danah Boyd. 2010. 'Dynamic Debates: An Analysis of Group Polarization Over Time on Twitter'. *Bulletin of Science, Technology & Society* 30 (5): 316–27. https://doi.org/10.1177/0270467610380011.

Yuan, Shupei, and Chen Lou. 2020. 'How Social Media Influencers Foster Relationships with Followers: The Roles of Source Credibility and Fairness in Parasocial Relationship and Product Interest'. *Journal of Interactive Advertising* 20 (2): 133–47. https://doi.org/10.1080/15252019.2020.1769514.

Zagheni, Emilio, Venkata Rama Kiran Garimella, Ingmar Weber, and Bogdan State. 2014. 'Inferring International and Internal Migration Patterns from Twitter Data'. In *Proceedings of the 23rd International Conference on World Wide Web*, 439–44. Seoul Korea: ACM. https://doi.org/10.1145/2567948.2576930.

Zhao, Dejin, and Mary Beth Rosson. 2009. 'How and Why People Twitter: The Role That Micro-Blogging Plays in Informal Communication at Work'. In *Proceedings of the 2009 ACM International Conference on Supporting Group Work*, 243–52. Sanibel Island Florida USA: ACM. https://doi.org/10.1145/1531674.1531710.

Zong, Chengqing, Rui Xia, and Jiajun Zhang. 2021. *Text Data Mining*. Singapore: Springer Singapore. https://doi.org/10.1007/978-981-16-0100-2.

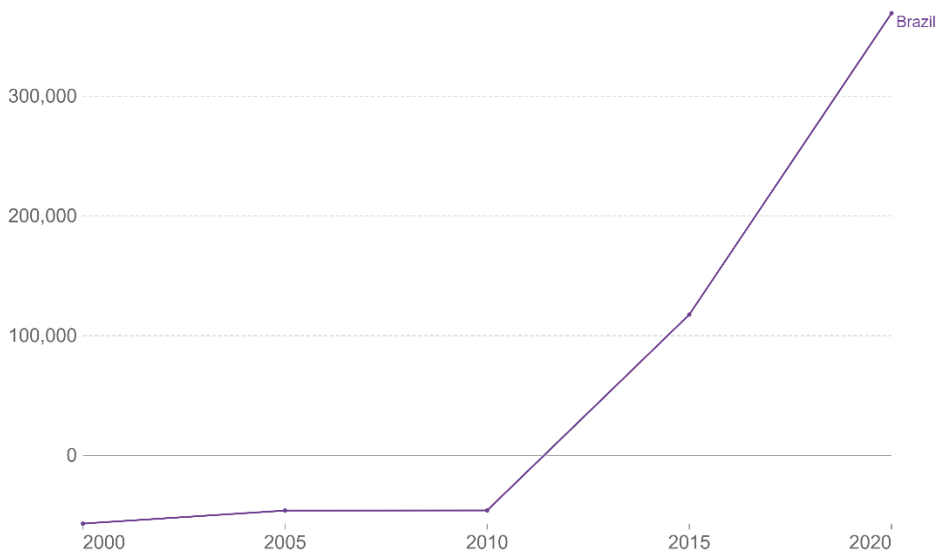# Annexes
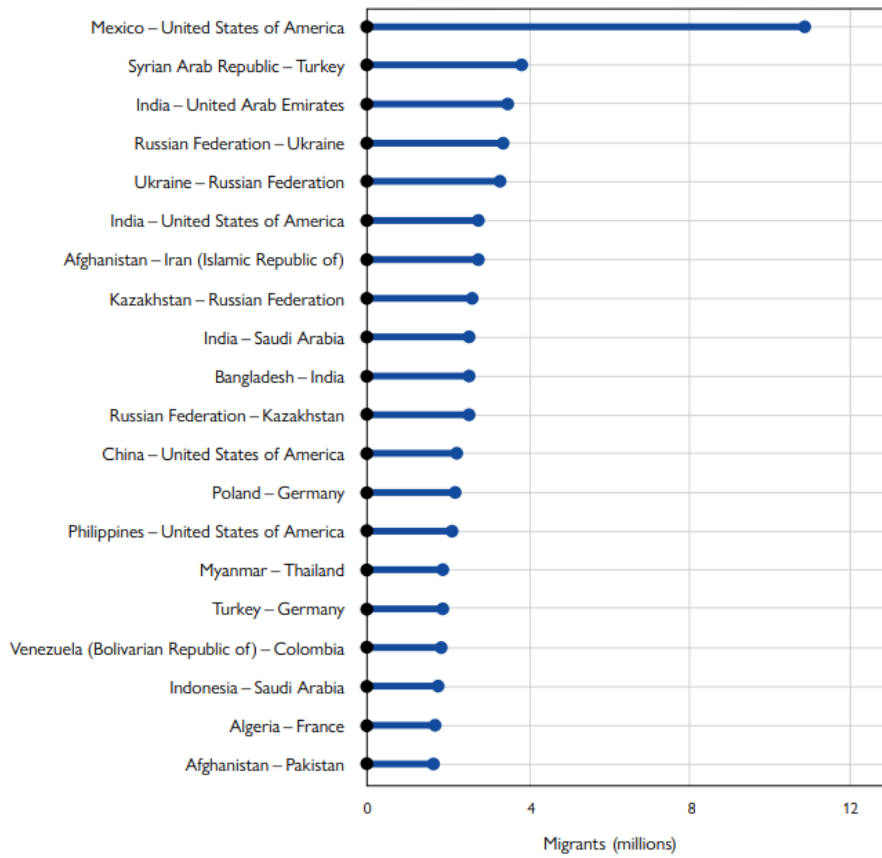
Figure 1

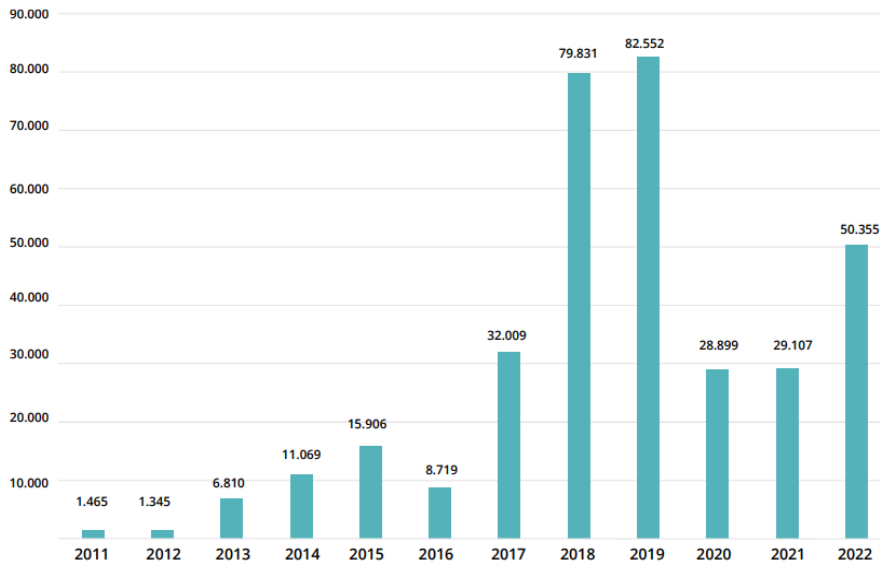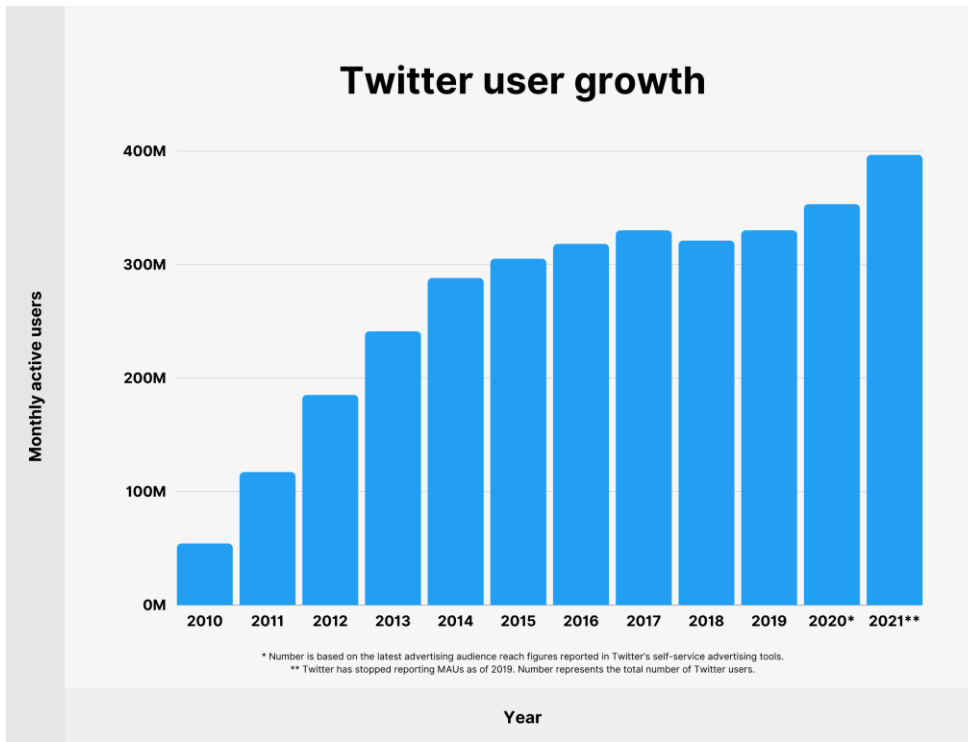Figure 2
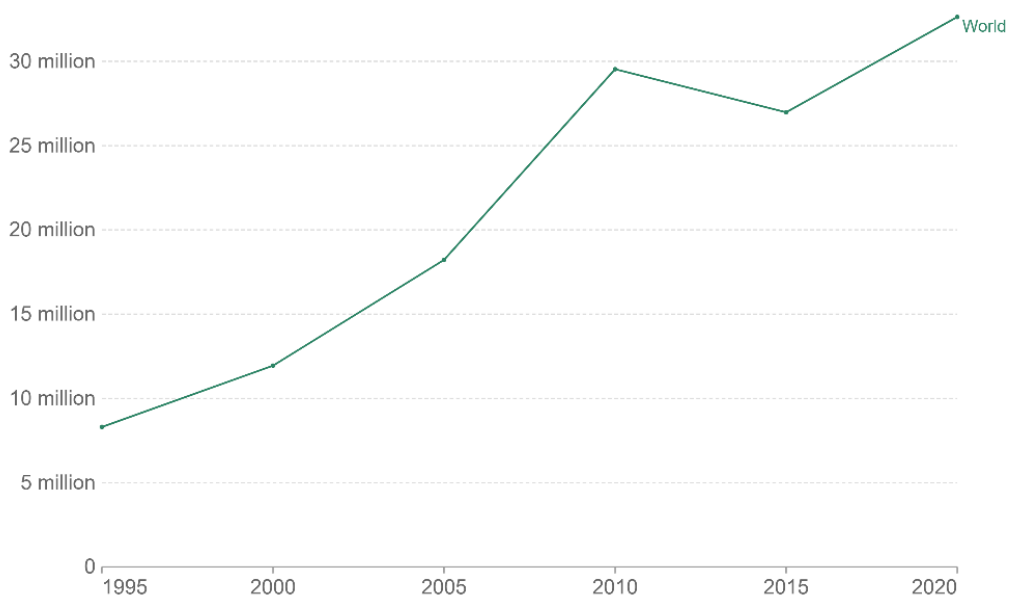
Figure 3
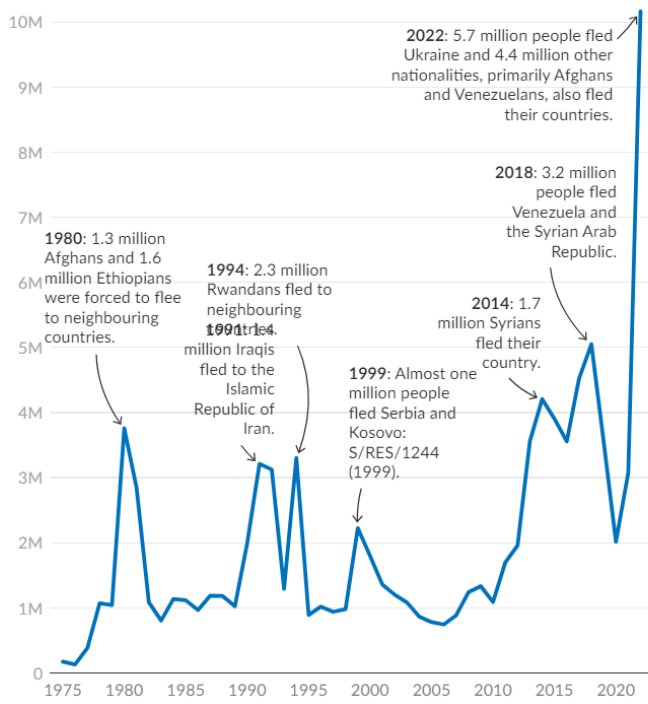
Figure 4

Figure 5

Figure 6

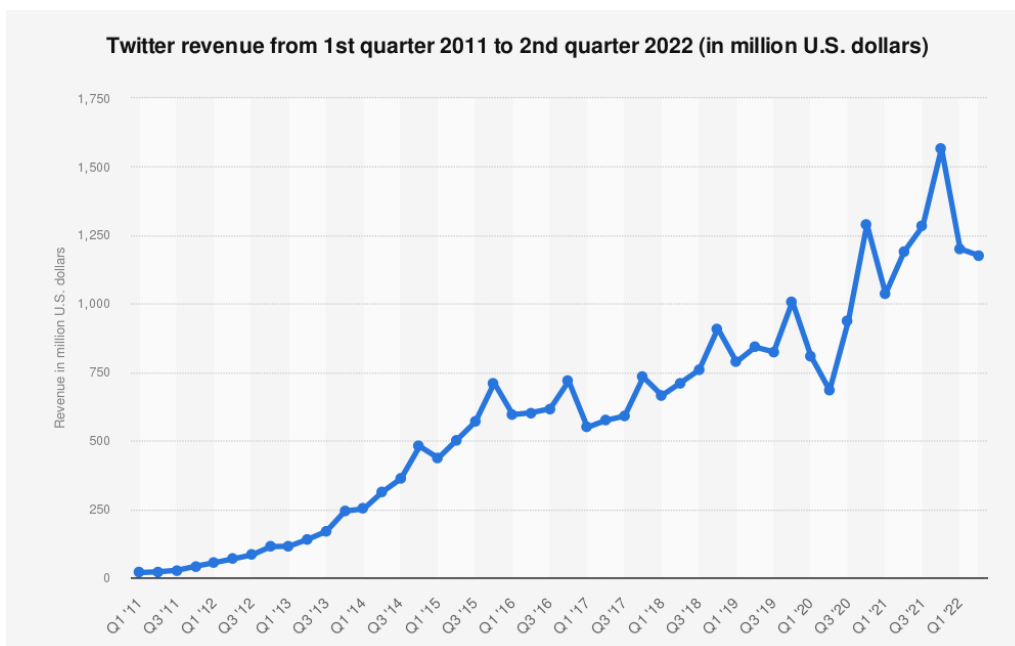Figure 7

Figure 8

135

Figure 9
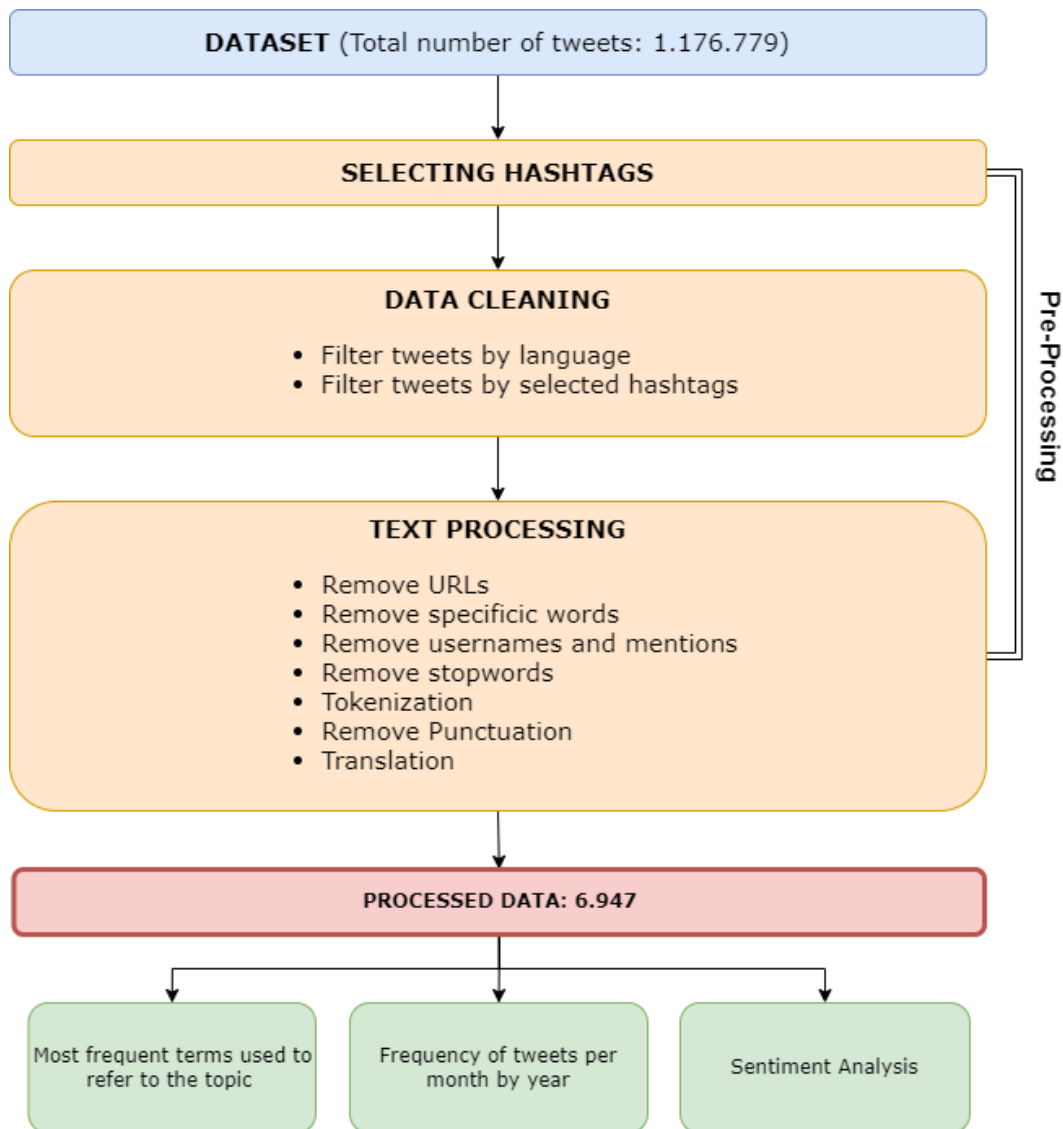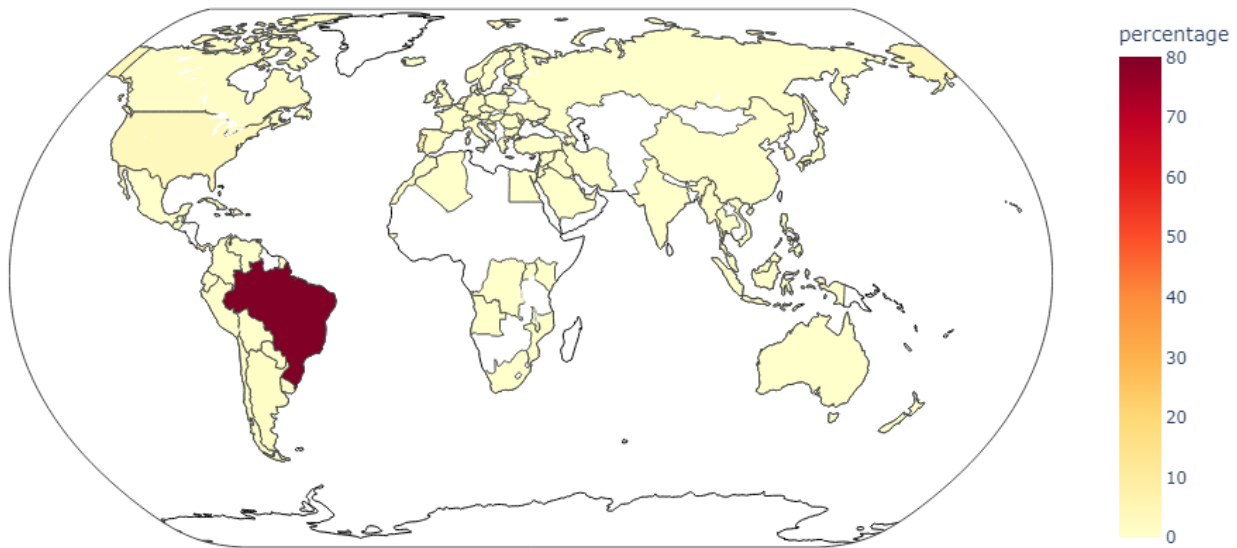

Figure 10

Figure 11

Figure 12

```python
specific_words = [
    'trânsito', 'rodovia', 'bitcoin', 'tráfego', 'AnchietaImigrantes',
    'ecovias', 'Anchieta', 'Radial', 'radial', 'dengue', 'copa', 'Leste',
    'anchieta', 'TIM', 'tim', 'Kaysar', 'novela', 'Bandeirantes',
    'operacaobetalab', 'bbb', 'transito', 'trânsito', 'orcars', 'oscar',
    'grammy', 'grammys', 'kaysar', 'timbetalab', 'masterchefbr', 'enem',
    'futebol', 'bundesliga', 'corinthians'
]
```

Figure 13

Figure 14

Monthly Frequency of Tweets by Year

Figure 15

```python
# Define a regular expression pattern to match URLs
url_pattern = r'http[s]?://(?:[a-zA-Z]|[0-9]|[$-_@.&+]|[!*\\(\\),]|(?:%[0-9a-fA-F][0-9a-fA-F]))+'

# Remove URLs from the 'text' column using regular expressions
df_filtered['text'] = df_filtered['text'].str.replace(url_pattern, '', regex=True).copy()
```

Figure 16

```python
# List of specific words to remove
specific_words_to_remove = ['RT', 'q' , 'vc', 'htt...', 'http...', 'http', 'https', 'htt', 'htt', 'p', 'pq', 'd', 'in',
                            'tá', 'km', 'n', 'aí', 'ai', 'R', 'h', 'etc', 'la', 'lá', 'at', 'e']

# Regular expression pattern to match the specific words
pattern = r'\b(?:' + '|'.join(specific_words_to_remove) + r')\b'

# Remove the specific words from the 'text' column using regular expressions
df_filtered['text'] = df_filtered['text'].str.replace(pattern, '', flags=re.IGNORECASE, regex=True).copy()
```

Figure 17

```python
# Define a regular expression pattern to match mentions (usernames)
mention_pattern = r'@[\w\d]+'

# Remove mentions from the 'text' column using regular expressions
df_filtered['text'] = df_filtered['text'].str.replace(mention_pattern, '', flags=re.IGNORECASE, regex=True).copy()
```

Figure 18

```python
# Tokenize the text in the 'text' column using NLTK
df_filtered['tokens'] = df_filtered['text'].apply(word_tokenize)
```

Figure 19

```
['a', 'à', 'ao', 'aos', 'aquela', 'aquelas', 'aquele', 'aqueles', 'aquilo', 'as', 'às', 'até', 'com', 'como', 'da', 'das', 'de', 'dela', 'delas', 'dele', 'deles', 'depois', 'do', 'dos', 'e', 'é', 'ela', 'elas', 'ele', 'eles', 'em', 'entre', 'era', 'eram', 'éramos', 'essa', 'essas', 'esse', 'esses', 'esta', 'está', 'estamos', 'estão', 'estar', 'estas', 'estava', 'estavam', 'estávamos', 'este', 'esteja', 'estejam', 'estejamos', 'estes', 'esteve', 'estive', 'estivemos', 'estiver', 'estivera', 'estiveram', 'estiverem', 'estivermos', 'estivesse', 'estivessem', 'estivéssemos', 'estou', 'eu', 'foi', 'fomos', 'for', 'fora', 'foram', 'fôramos', 'forem', 'formos', 'fosse', 'fossem', 'fôssemos', 'fui', 'há', 'haja', 'hajam', 'hajamos', 'hão', 'havemos', 'haver', 'hei', 'houve', 'houvemos', 'houver', 'houvera', 'houverá', 'houveram', 'houvéramos', 'houverão', 'houverei', 'houverem', 'houveremos', 'houveria', 'houveriam', 'houveríamos', 'houvermos', 'houvesse', 'houvessem', 'houvéssemos', 'isso', 'isto', 'já', 'lhe', 'lhes', 'mais', 'mas', 'me', 'mesmo', 'meu', 'meus', 'minha', 'minhas', 'muito', 'na', 'não', 'nas', 'nem', 'no', 'nos', 'nós', 'nossa', 'nossas', 'nosso', 'nossos', 'num', 'numa', 'o', 'os', 'ou', 'para', 'pela', 'pelas', 'pelo', 'pelos', 'por', 'qual', 'quando', 'que', 'quem', 'são', 'se', 'seja', 'sejam', 'sejamos', 'sem', 'ser', 'será', 'serão', 'serei', 'seremos', 'seria', 'seriam', 'seríamos', 'seu', 'seus', 'só', 'somos', 'sou', 'sua', 'suas', 'também', 'te', 'tem', 'tém', 'temos', 'tenha', 'tenham', 'tenhamos', 'tenho', 'terá', 'terão', 'terei', 'teremos', 'teria', 'teriam', 'teríamos', 'teu', 'teus', 'teve', 'tinha', 'tinham', 'tínhamos', 'tive', 'tivemos', 'tiver', 'tivera', 'tiveram', 'tivéramos', 'tiverem', 'tivermos', 'tivesse', 'tivessem', 'tivéssemos', 'tu', 'tua', 'tuas', 'um', 'uma', 'você', 'vocês', 'vos']
```

Figure 20

```python
additional_stopwords = ['a', 'à', 'agora', 'ainda', 'além', 'algo', 'algumas', 'alguns', 'ali', 'ano', 'anos', 'antes', 'ao',
    'aos', 'apenas', 'apoio', 'após', 'aquela', 'aquelas','aqui', 'aquele', 'aqueles', 'aquilo', 'área', 'as', 'às', 'assim',
    'até', 'atrás', 'através', 'baixo', 'bastante', 'bem', 'boa', 'boas', 'bom', 'bons', 'breve', 'cá', 'cada', 'catorze',
    'cedo', 'cento','certamente', 'certeza', 'cinco', 'coisa', 'coisas', 'com', 'como', 'conselho', 'contra', 'contudo',
    'corrente','cuja', 'cujas', 'cujo', 'cujos', 'da', 'dá', 'dão', 'daquela', 'daquelas', 'daquele', 'daqueles', 'daqui',
    'daquilo','das', 'de', 'debaixo', 'dela', 'delas', 'dele', 'deles', 'dêm', 'demais', 'dentro', 'depois', 'desde', 'dessa',
    'dessas', 'desse', 'desses', 'desta', 'destas', 'deste', 'destes', 'deve', 'deverá', 'dez', 'dezenove', 'dezesseis',
    'dezassete', 'dezoito', 'dia', 'diante', 'diz', 'dizem', 'do', 'dona', 'donas', 'dono', 'donos', 'dos', 'doze', 'duas',
    'dúvida', 'e', 'ela', 'elas', 'ele', 'eles', 'em', 'embora', 'entre', 'então', 'entanto', 'era', 'eram', 'éramos',
    'é', 'essa', 'essas', 'esse', 'esses', 'esta', 'estamos', 'está', 'estão', 'estar', 'estará', 'estas', 'estás', 'estava',
    'estavam', 'este', 'estes', 'esteve', 'estive', 'estivemos', 'estiveram', 'estiveste', 'estivestes', 'estou', 'eu',
    'exemplo', 'falta', 'fará', 'favor', 'faz', 'fazeis', 'fazem', 'fazemos', 'fazer', 'fazes', 'fez', 'fim', 'final', 'foi',
    'fomos', 'for', 'fora', 'foram', 'formos', 'fosse', 'fossem', 'foste', 'fostes', 'fui', 'geral', 'grande', 'grandes',
    'grupo', 'hoje', 'hora', 'horas', 'ir', 'irá', 'isso', 'isto', 'já', 'lá', 'lado', 'ligado', 'local', 'logo', 'longe',
    'lugar', 'maior', 'maioria', 'maiorias', 'mais', 'mal', 'mas', 'máximo', 'me', 'meio', 'menor', 'menos', 'mês',
    'meses', 'meu', 'meus', 'mil', 'minha', 'minhas', 'momento', 'muito', 'muitos', 'na', 'nada', 'naquela', 'naquelas',
    'naquele', 'naqueles', 'nas', 'nem', 'nenhuma', 'nessa', 'nessas', 'nesse', 'nesses', 'nesta', 'nestas', 'neste', 'nestes',
    'ninguém', 'no', 'nos', 'nós', 'nossa', 'nossas', 'nosso', 'nossos', 'nova', 'novas', 'nove', 'novo', 'novos', 'num',
    'numa', 'número', 'nunca', 'nuns', 'o', 'obra', 'obrigada', 'obrigado', 'oitava', 'oitavo', 'oito', 'onde', 'ontem',
    'onze', 'os', 'ou', 'outra', 'outras', 'outro', 'outros', 'para', 'pra','parece', 'parte', 'partir', 'paucas', 'pela','pelas'
    'pelo', 'pelos', 'perto', 'pode', 'pôde', 'podem', 'poderá', 'podia', 'pois', 'ponto', 'pontos', 'por', 'porque', 'porquê',
    'pouca', 'pouco', 'poucos', 'primeira', 'primeiras', 'primeiro', 'primeiros', 'própria', 'próprias', 'próprio', 'próprios',
    'quáis', 'qual', 'qualquer', 'quando', 'quanto', 'quarta', 'quarto', 'quatro', 'que', 'quem', 'quer', 'quê', 'quinta',
    'quinto', 'quinze', 'relação', 'sabe', 'são', 'se', 'segunda', 'segundo', 'sei', 'seis', 'seja', 'sejam', 'sempre', 'sendo',
    'ser', 'será', 'seu', 'seus', 'sexta', 'sexto', 'sim', 'sistema', 'sob', 'sobre', 'sois', 'somos', 'sou', 'sua', 'suas',
    'tal', 'talvez', 'também', 'tanta', 'tantas', 'tanto', 'tantos', 'te', 'tem', 'têm', 'temos', 'tendes', 'tenho', 'tens',
    'ter', 'terá', 'terão', 'terceira', 'terceiro', 'teu', 'teus', 'teve', 'ti', 'tido', 'tinha', 'tinham', 'tive', 'tivemos',
    'tiveram', 'tiveste', 'tivestes', 'toda', 'todas', 'todo', 'todos', 'trabalho', 'três', 'treze', 'tu', 'tua', 'tuas',
    'tudo', 'última', 'últimas', 'último', 'últimos', 'um', 'uma', 'umas', 'uns', 'usa', 'usar', 'vai', 'vais', 'valor',
    'veja', 'vem', 'vens', 'ver', 'verdade', 'verdadeiro', 'vez', 'vezes', 'viagem', 'vindo', 'vinte', 'você', 'vocês',
    'vos', 'vós', 'vossa', 'vossas', 'vosso', 'vossos', 'zero', 'via']

# Combine both lists of stopwords
portuguese_stopwords = stop + additional_stopwords

df_filtered['filtered_tokens'] =
    df_filtered['tokens'].apply(lambda tokens:[word for word in tokens if word.lower() not in portuguese_stopwords])
```

Figure 21

```python
def join_tokens_into_phrase(tokens):
    return ' '.join(tokens)

# Define a pattern to remove punctuation
punctuation_pattern = r'[^\w\s]'

# Define a list of specific punctuation signs to remove
specific_punctuation = [':', ';', '.', ',', '?', '&']

# Remove the specified punctuation from filtered tokens
def remove_specific_punctuation(tokens):
    return [re.sub(punctuation_pattern, '', word) for word in tokens if word not in specific_punctuation]

df_filtered['cleaned_tokens'] = df_filtered['filtered_tokens'].apply(remove_specific_punctuation)

# Combine cleaned tokens into a list
df_filtered['cleaned_text'] = df_filtered['cleaned_tokens'].apply(join_tokens_into_phrase)

# Remove hashtags from the 'cleaned_text' column
df_filtered['cleaned_text'] = df_filtered['cleaned_text'].str.replace(r'#\w+', '', regex=True)
```

Figure 22

```python
def reorganize_dataframe(df):
    # Reset the index
    df.reset_index(drop=True, inplace=True)

    # Rearrange columns in the desired order
    new_column_order = ['id', 'text', 'created_at', 'author_id', 'lang', 'in_reply_to_user_id', 'tokens',
                        'filtered_tokens', 'cleaned_tokens', 'cleaned_text']
    df = df[new_column_order]

    return df

# Call the reorganize_dataframe function on the pre-processed dataframe
df_filtered_org = reorganize_dataframe(df_filtered)
```

Figure 23

Figure 24

```python
import nltk
from nltk.sentiment.vader import SentimentIntensityAnalyzer
nltk.download('vader_lexicon')
sia = SentimentIntensityAnalyzer()
```

Figure 25

```python
def classify_sentiment(compound_score):
    if compound_score >= 0.05:
        return 'positive'
    elif compound_score <= -0.05:
        return 'negative'
    else:
        return 'neutral'
```
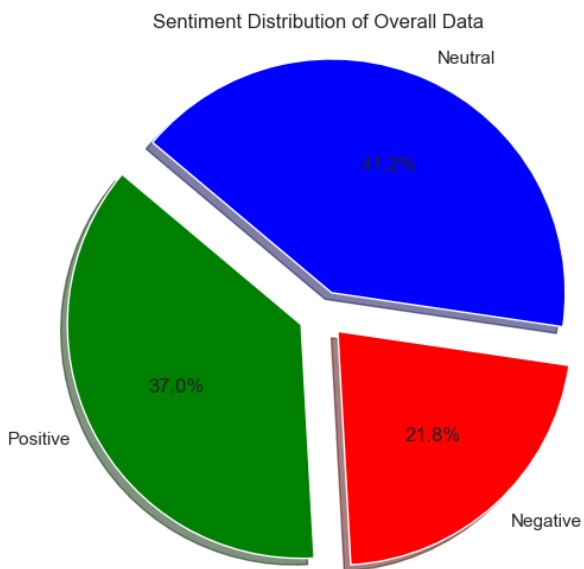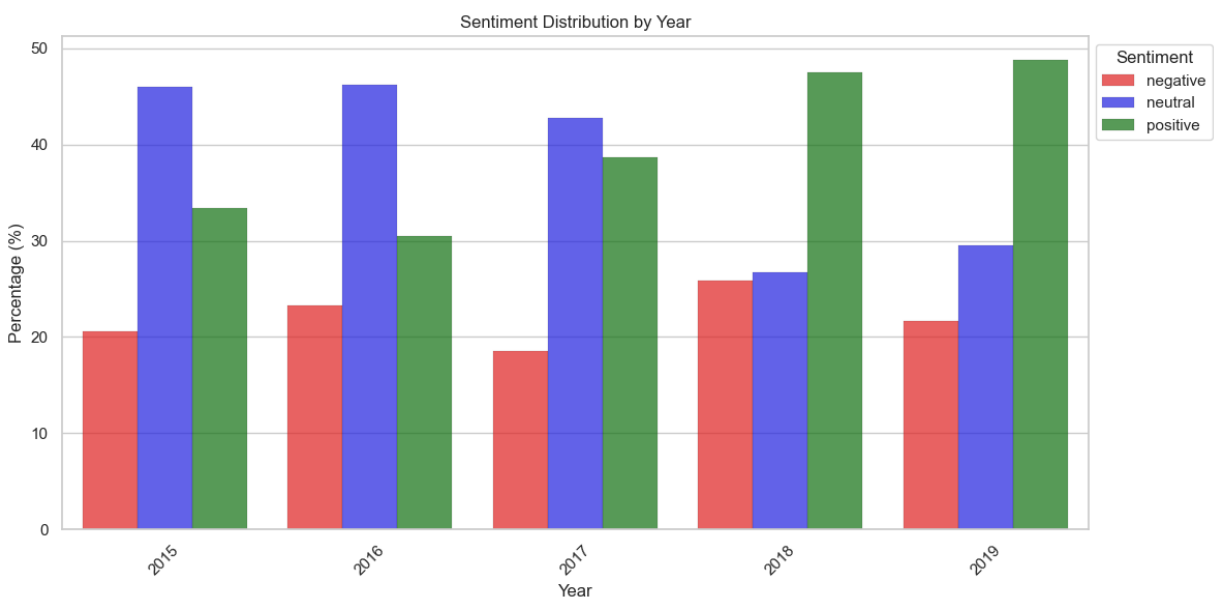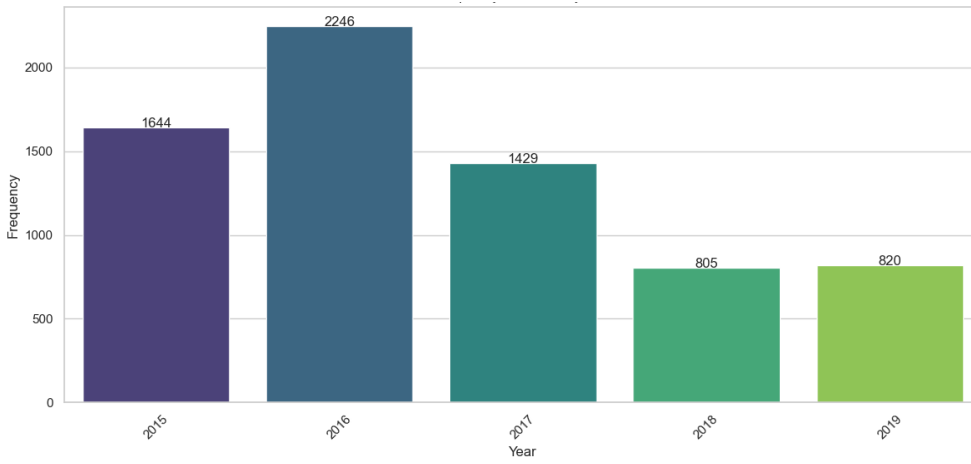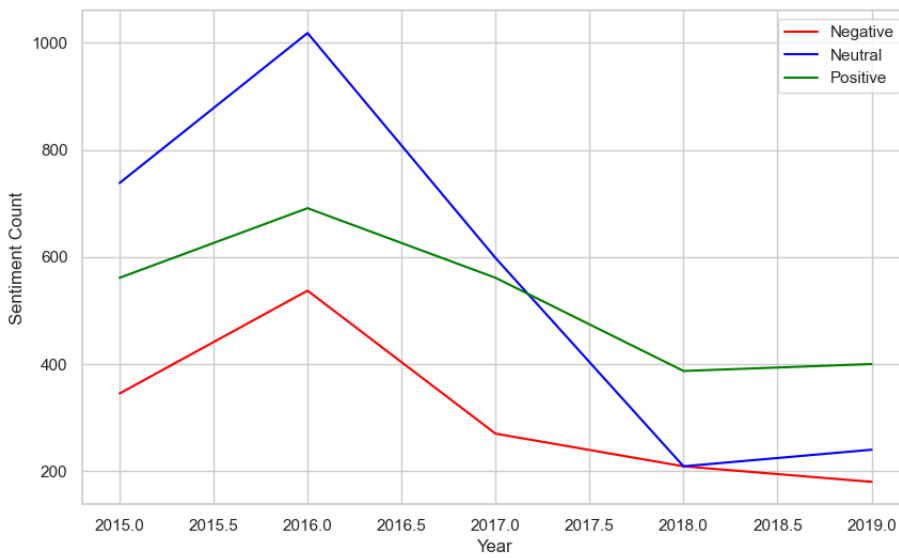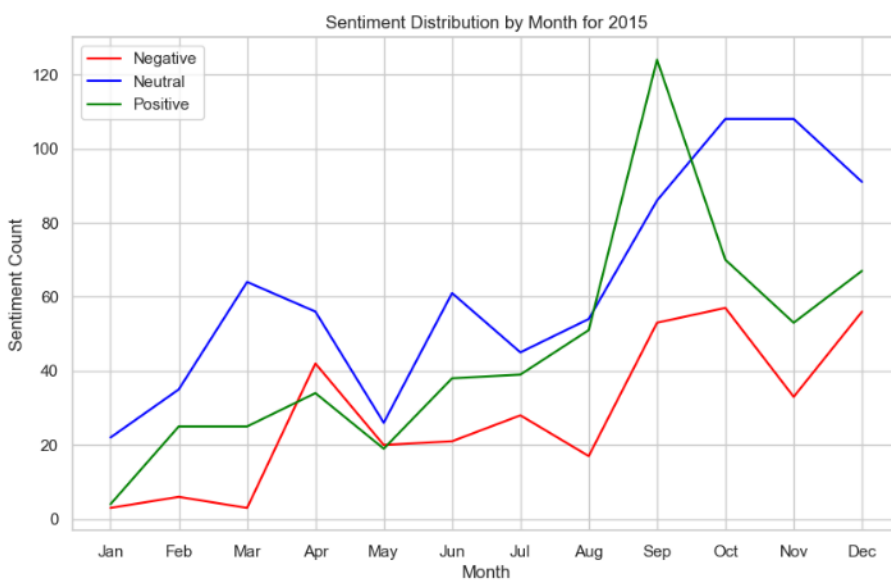
Figure 26

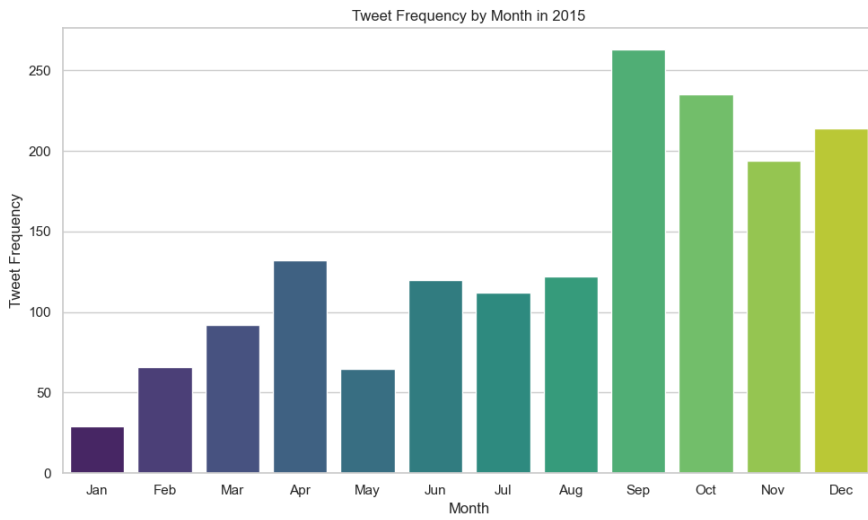Figure 27

Figure 28

Figure 29

Figure 30

Sentiment Distribution by Month for 2015

Figure 31

Tweet Frequency by Month in 2015

Figure 32

Sentiment Distribution by Month for 2016

Figure 33

Tweet Frequency by Month in 2016

Figure 34

Tweet Frequency by Month in 2017

Figure 35

Sentiment Distribution by Month for 2017

Figure 36

Sentiment Distribution by Month for 2018

Figure 37

Tweet Frequency by Month in 2018

Figure 38

Sentiment Distribution by Month for 2019
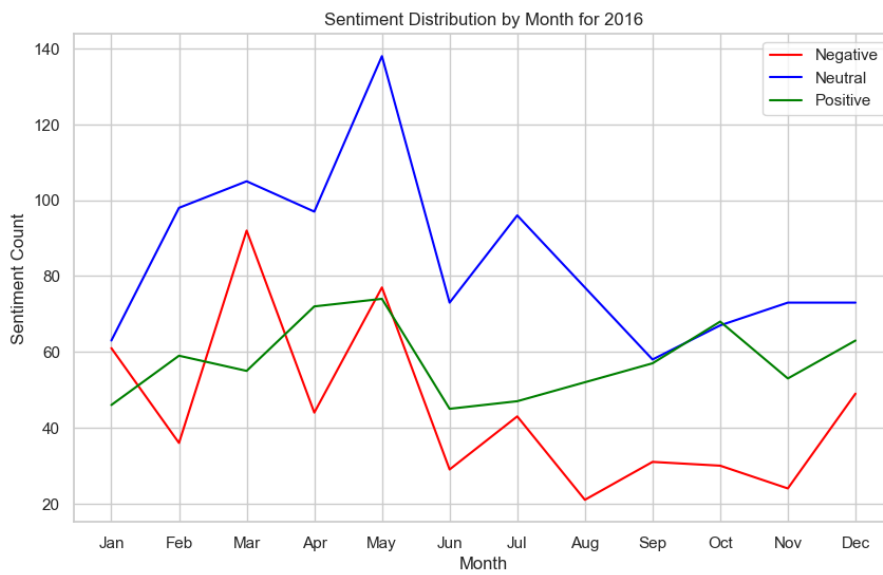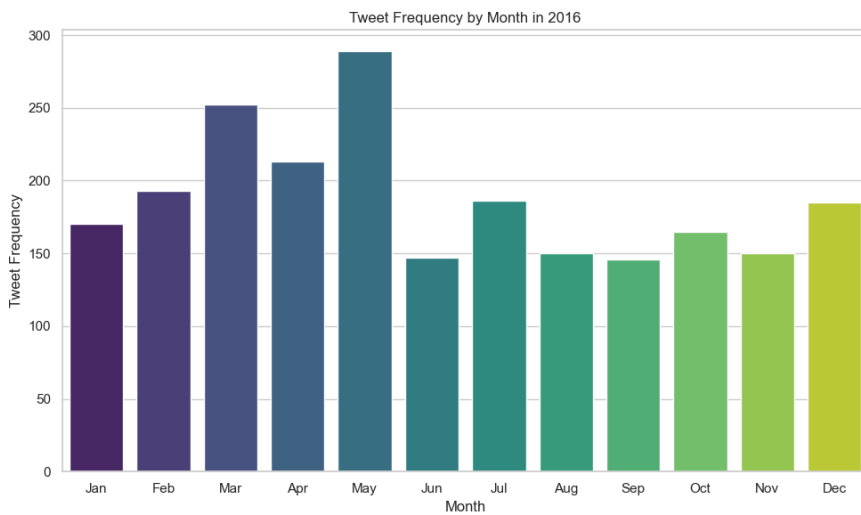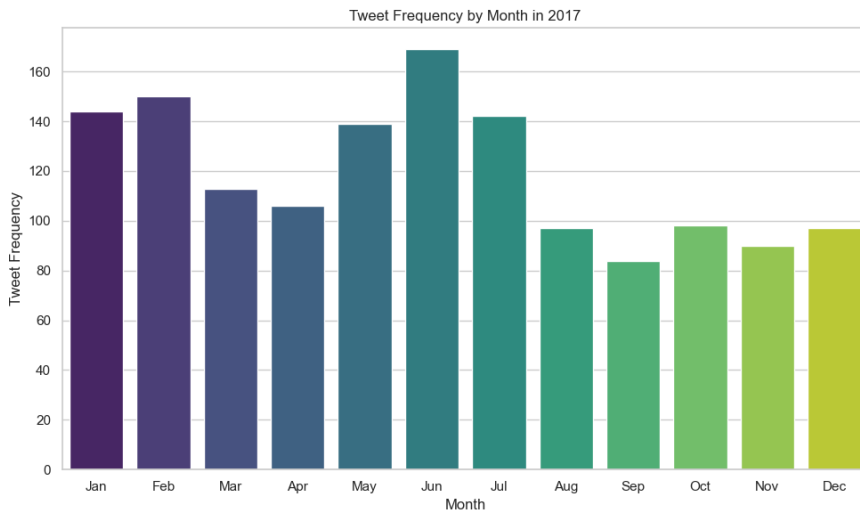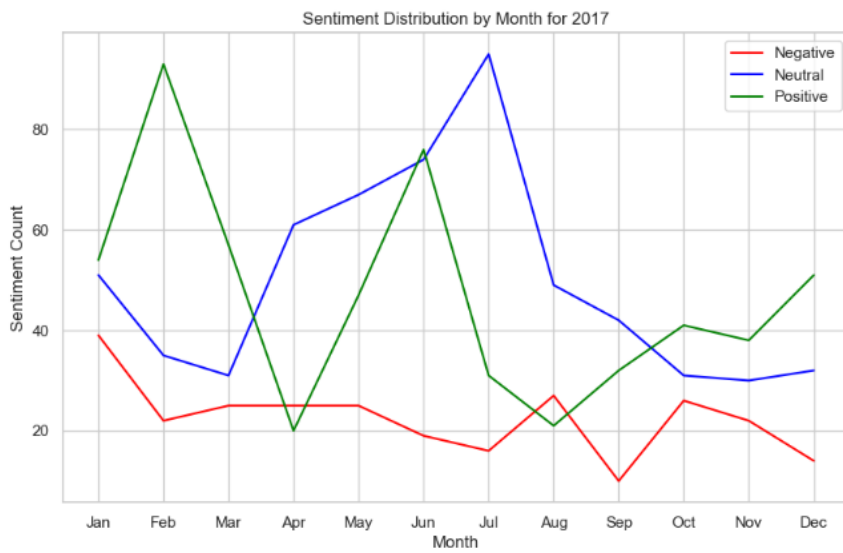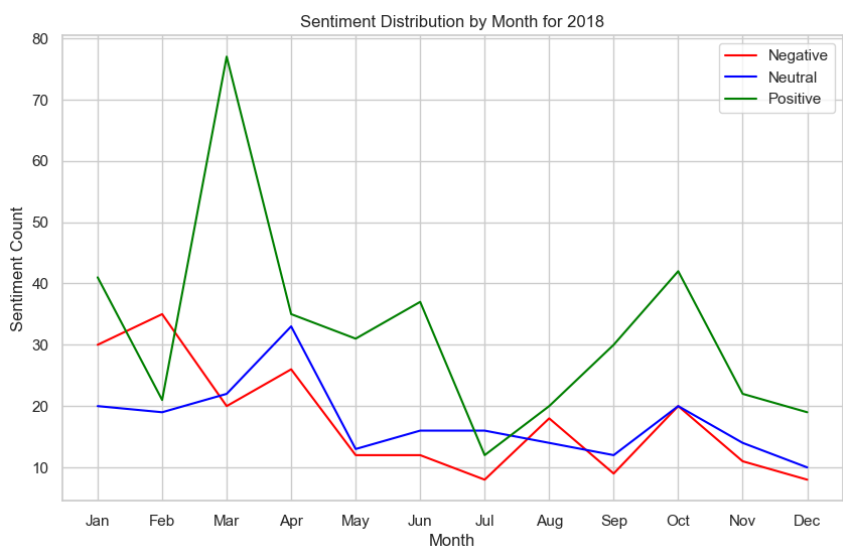
Figure 39
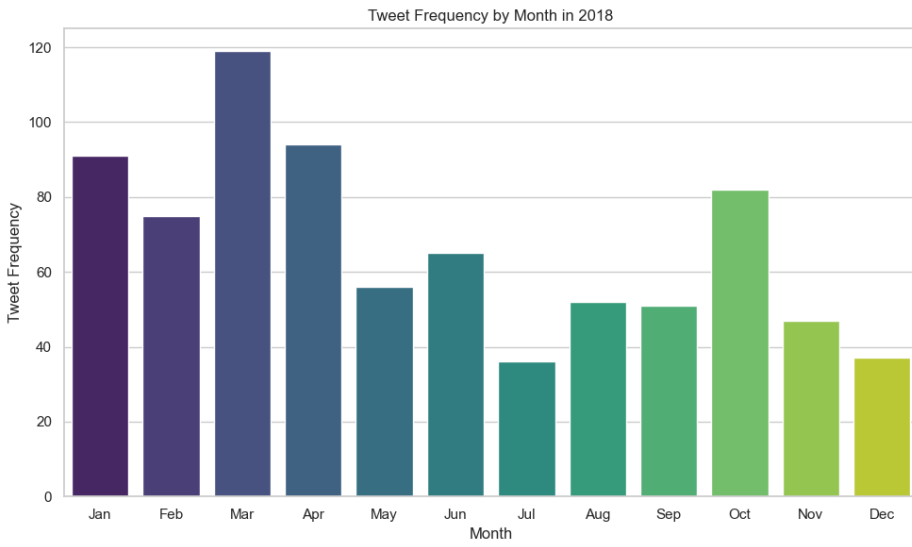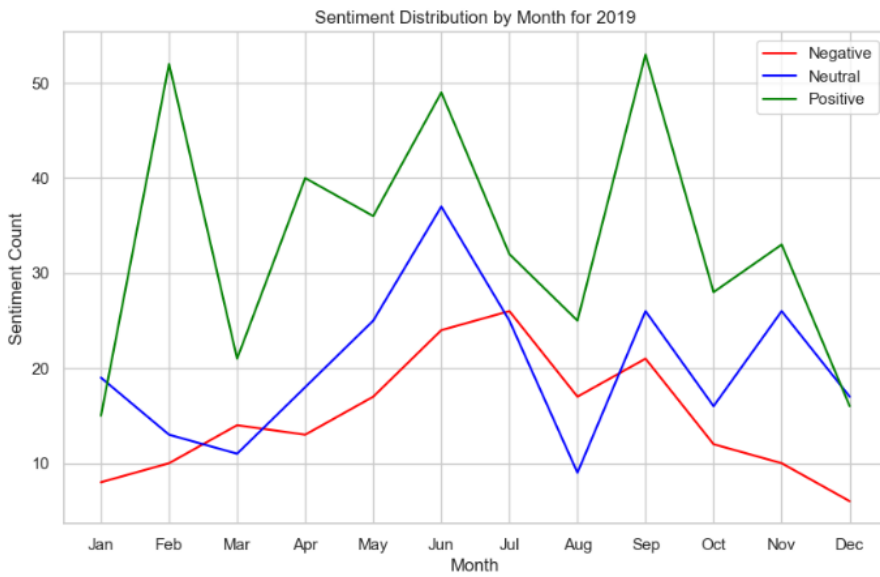
Tweet Frequency by Month in 2019

Table 1

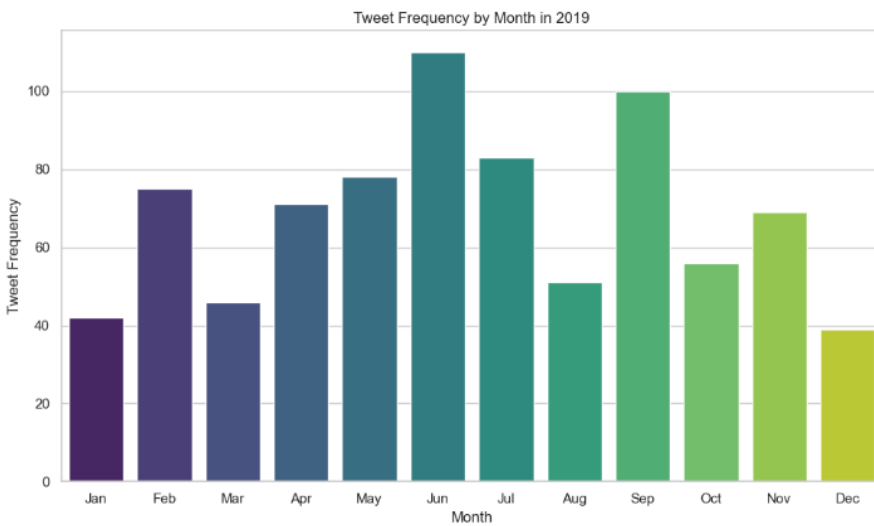| YEAR	MONTH	HASHTAGS AND FREQUENCY
2015	January	acnur (16), brasil (13), imigrantes (12), imigração (6), refugiados (4), eumigrante (4), expressentry (3), sírios (2), imigrante (2), imigracao (2), somostodosmigrantes (2).
	February	refugiados (34), eumigrante (20), onu (12), acnur (11), migrante (11), imigração (10), imigrantes (9), opentosyria (9), maismédicos (5), migración (4), imigrante (3).
	March	refugiados (75), acnur (38), syria (25), onu (23), imigração (22), imigrantes (21), imigracao (8), imigrante (7), eupertenço (7), frontex (5), ibelong (4).
	April	refugiados (69), imigrantes (55), imigração (23), acnur (19), onu (13), migration (10), imigrante (6), xenophobia (6), migrants (5), migrante (5), migrantes (4).
	May	refugiados (46), imigrantes (45), onu (19), imigração (10), imigrante (8), opentosyria (6), acnur (5).
	June	refugiados (72), diamundialdorefugiado (26), onu (23), imigração (18), imigrantes (18), refugiado (14), xenofobianão (13), todossomosrefugiados (12), migrantes (5), diamundialdosrefugiados (5), worldrefugeeday (4), acnur (4), refugees (4).
	July	refugiados (88), síria (29), onu (26), imigração (26), imigrantes (17), refugiado (13), migrante (7).
	August	refugiados (70), líbia (40), diamundialhumanitário (35), imigração (26), imigrantes (26), sharehumanity (25), onu (14), refugiado (8), haitianos (7), sírios (7), migrantes (5), compartilhehumanidade (5), migrante (4).
	September	compartilhehumanidade (220), refugiados (163), refugeeswelcome (78), refugeecrisis (13), imigrantes (11), refugees (9), refugiado (7), acnur (6), refugeescrisis (5), welcomerefugees (5), somostodosimigrantes (4), imigrante (4).
	October	refugiados (186), imigração (27), imigrantes (22), refugeeswelcome (14), compartilhehumanidade (8), refugees (7), crisederefugiados (7), imigracao (6), migración (6), refugiado (5), migrante (5), migrantes (5).
	November	refugiados (165), refugeecrisis (33), refugeeswelcome (25), imigrantes (20), openeuborders (13), imigraçãogerainovação (13), fechaasfronteirasportugal (12), ibelong (8), imigração (7), openborders (6), crisederefugiados (5), estamosxlosrefugiados (5), haitianosficamnobrasil (5), migrantes (4), somostodosimigrantes (4).
	December	refugiados (200), imigrantes (27), refugees (17), refugiado (17), compartilhehumanidade (15), elrefugiado (10), soumigrante (8), somosmigrantes (8), derechoshumanos (7), imigrante (6), acnur (6).
2016	January	refugiados (163), imigrantes (16), migrantes (9), naufrágio (9), acnur (7), refugees (7), imigração (6), migrante (6), deportaci (6), refugeeswelcome (5), crisedosrefugiados (5), refugeecrisis (5).
	February	refugiados (164), refugees (21), imigrante (18), migrantes (11), imigrantes (10), migração (9), welcomerefugees (9), safepassage (8), imigração (7), acnur (7), corridoiumanitari (7), expressentry (6), sirios (6), refugeecrisis (5).
	March	refugiados (372), refugeeswelcome (46), operativomigrante (33), imigrantes (17), withsyria (16), refugioderechoshumanos (12), refugees (10), racismo (10), prouracisme (10), ueturkey (9), uemata (8), migrações (7), imigração (7), refugeecrisis (7), ueturquia (7), crisis (6), refugiado (6), migrantes (5).
	April	refugiados (186), imigrantes (24), refugiado (21), refugiats (19), yorefugio (16), overthefortress (13), migrante (11), imigração (10), refugeeswelcome (9), refugeecameras (9), acnur (8), imigrante (7), openeuborders (7), refugees (6), migrantes (6), crisisrefugiados (6), minutorefugiados (5).

	May	refugiados (235), cambioclimático (44), imigrantes (41), refugeeswelcome (27), acnur (22), jantarhumanitario (13), unhcr (13), eusourefugio (13), imigracao (13), jeansforrefugees (8), refugiadosnobrasil (7), refugees (7), migrações (7), imigração (5), migrante (5).
	June	refugiados (127), comosrefugiados (46), imigrantes (43), acnur (34), teamrefugees (29), withrefugees (26), imigração (15), imigracao (12), worldrefugeeday (12), diamundialdorefugiado (11), sírios (6), nobordersnonations (6), iamanimmigrant (5), refugees (5), diamundialdosrefugiados (5), mediterrâneo (4).
	July	refugiados (144), comosrefugiados (63), acnur (32), teamrefugees (30), imigrantes (21), refugiado (20), imigração (15), opentheborders (13), refúgio (10), refugiats (9), imigracao (8), refugiadas (6), migrante (6), prayforsyria (6), refugiada (5), refugees (5), migrantes (5), moçambicanos (5).
	August	comosrefugiados (400), teamrefugees (285), refugiados (106), refugeeolympicteam (45), imigrantes (27), imigração (25), teamrefugee (19), imigracao (13), migrantes (6), atletasrefugiados (6), crisisderefugiados (6), refugiado (5).
	September	comosrefugiados (164), refugiados (117), imigração (27), teamrefugess (24), imigrantes (23), acnur (22), imigracao (17), onu (16), withrefugees (12), un4refugeesmigrants (11), migrantes (9), teamrefugees (5).
	October	refugiados (119), imigrantes (35), imigração (29), imigracao (24), migrantes (20), welcomerefugees (15), euacolho (14), acnur (12), migrante (12), caritasrj (11), refugees (6), un4refugeesmigrants (5), inmigración (5)
	November	refugiados (135), comosrefugiados (43), imigração (34), imigrantes (27), imigracao (23), withrefugees (18), migrantes (11), refugiado (9), migrante (8), imigrante (6), immigration (6), rifugiato (5), migrarédireito (5).
	December	refugiados (132), euaceitorefugiados (55), nãoaceitamosrefugiados (51), meuamigorefugiado (48), imigração (47), imigracao (29), imigrantes (21), naoaceitamosrefugiados (20), migrantes (18), refugiado (12), conlosrefugiados (10), migramundo (9), comosrefugiados (6), migrarédireito (6), imigracaoitaliana (6), migrante (6), prayforaleppo (5), refugees (4), acolhidas (4), imigracaonononordeste (4), migrantsday (4).
2017	January	refugiados (113), imigrantes (46), imigração (36), refugees (19), migrante (17), imigracao (16), imigracaoitaliana (7), migrantes (7), euaceitorefugiados (6), imigracaonononordeste (5), muslimban (5), comosrefugiados (4), refugeeswelcome (4), acnur (4).
	February	refugiados (59), volemacollir (59), imigração (34), imigrantes (23), imigracao (18), refugees (8), refugiats (7), vistos (6), acnur (5), refugeeswelcome (5).
	March	refugiados (91), imigração (47), imigrantes (24), migramundo (14), onu (14), migranti (13), imigracao (13), amorsemfronteiras (9), refugiadas (7), acnur (7), direitoshumanos (6), syriacrisis (5), withrefugees (5), migrantes (4), refugees (4).
	April	vetatemer (334), leidamigraçãonão (49), refugiados (42), migracaonao (39), imigração (38), rejeitemigracao (26), aloysiomigracaoagoranao (25), leidemigracaonao (25), imigrantes (22), vistos (13), acnur (13), rejeiteleidemigracao (13), vetamigracaotemer (12), migramundo (11), vetatemerleiimigracao (5), imigracao (4), unhcr (4), aloysiomigraçãoagoranao (4), migraçãonão (4), refugiado (3), refugees (3), migrações (3), leidemigraçaonao (3), migrarédireito (3).
	May	vetatemer (260), migracaoveta (73), refugiados (45), imigrantes (42), terroristaleimigracao (41), imigração (33), vetatemermigração (28), leidemigracaonao (21), iwelcome (9), euacolho (9), migrações (9), acnur (7), comosrefugiados (7), vetaimigracaotemer (6), migramundo (6), naoaleidemigracao (4).

	June	refugiados (102), imigração (65), imigrantes (61), refugeeday (26), comosrefugiados (20), diamundialdorefugiado (17), acnur (12), withrefugees (11), imigracao (8), díamundialdorefuxiado (7).
	July	imigrantes (97), refugiados (87), exodus (56), imigração (47), imigrante (6), acnur (6), migramundo (5), migrarédireito (5).
	August	imigrantes (61), refugiados (50), acnur (33), imigração (29), racismo (12), exodus (10), refugees (8), migramundo (7), migrantes (5), imigracao (4)
	September	imigrantes (71), refugiados (56), onu (28), imigração (22), acnur (18), refugiadosvavel (13), refugees (8), migrantes (6), migramundo (6).
	October	acnur (109), imigrantes (85), refugiados (52), imigração (32), migramundo (12), migrante (6), migrantes (6), refugees (6), refugiado (6).
	November	imigrantes (63), refugiados (46), imigração (35), imigrante (20), acnur (7), migramundo (6), humanflow (5), migrante (5), migranti (5), migrantes (4).
	December	refugiados (257), diadelosderechoshumanos (204), imigrantes (65), imigração (22), migrantes (21), acnur (16), conlosrefugiados (9), acogida (5), refugees (4), corredoreshumanitarios (4), imigrante (3), direitoshumanos (3).
2018	January	refugiados (74), imigrantes (65), imigração (24), acnur (19), migrantes (17), refugiado (13), migrante (11), refugiadoslgtb (6), accoglienza (5), rifugiato (5), refugees (5)
	February	refugiados (61), imigrantes (41), imigração (8), migrantes (8), migramundo (4), migrante (3), comosrefugiados (3), acnur (3), pactoglobalsobrerefugiados (3).
	March	refugiados (80), imigrantes (60), migrantes (47), acnur (8), semfronteiras (8), operaçãoacolhida (7), humanrights (5), imigrante (5), imigração (4), artemigrante (4), crisehumanitária (3), direitoshumanos (3), derechoshumanos (3), fraternidadesemfronteiras (3).
	April	refugiados (58), freeopenarms (49), operaçãoacolhida (44), imigrantes (32), imigração (20), direitoshumanos (10), migramundo (5), unhcr (4), acnur (4), imigrantebrasileiro (3), migrantes (3), refúgioemnúmeros (3).
	May	imigrantes (51), refugiados (40), imigração (19), imigrante (9), respeitaosrefugiados (9), acnur (7), operacaoacolhida (6), migrantes (5), venezuelanos (5), migraciones (4), fronterasur (4), migramundo (3), direitoshumanos (3),migrante (3).
	June	refugiados (51), imigrantes (44), imigração (22), refugeeswelcome (11), acnur (11), migrantes (9), direitoshumanos (8), meuamigorefugiadocopa (7), imigrantesdevolta (6), refugeeswelcomegfbpa (6), xenofobia (5), racismo (5), stopracismo (5), bienvenidos (5), worldrefugeeday2018 (5), refugiats (5), comosrefugiados (4).
	July	refugiados (29), imigração (22), imigrantes (10), criminiimmigrati (8), refugeeswelcomegfbpa (6), acnur (5), oim (5), migração (5), migrantes (5).
	August	refugiados (32), imigrantes (20), criminiimmigrati (17), imigração (14), acnur (9), venezuelanos (7), venezolanos (4).
	September	refugiados (52), imigração (16), xenofobia (9), imigrantes (9), migrantes (8), venezuelanos (6), direitoshumanos (4), criminiimmigrati (3), refugiado (3), acnur (3), deportação (3).
	October	refugiados (54), caravanamigrante (20), acnur (18), imigrantes (16), imigração (11), migrantes (6), venezuelanos (6), refugiado (5), direitoshumanos (4), migrante (3), imigrante (3).
	November	acnur (30), imigração (19), imigrantes (19), venezuela (19), refugiados (18), caravanamigrante (7), imigrante (6), migrantes (4), criminiimmigrati (4), fronteira (4), imigracao (3), venezuelanos (3), orepelosrefugiados (2), direitoshumanos (2)
	December	refugiados (19), imigrantes (16), direitoshumanos (10), imigração (9), acnur (7), sonbienvenidos (5), migrantes (4), migrações (3), pactoglobalimigracao (2), imigracao (2), refugee (2), diainternacionaldosmigrantes (2), reassentados (2)

2019	January	operaçãoacolhida (186), imigrantes (28), refugiados (23), acnur (16), imigração (7), refuxiados (6), migrantes (5), venezuelanos (5), criminiimmigrati (3), caravanamigrante (3), direitoshumanos (2), imigracao (2), conlosrefugiados (2), refugees (2), venezuelanas (2), refugiadas (2).
	February	refugiado (28), refugiados (27), imigrantes (21), acnur (18), operaçãoacolhida (17), imigração (11), migrantes (6), migration (5), migrante (5), direitoshumanos (4), imigracao (4), refuxiados (3), humanity (2), refugees (2).
	March	operaçãoacolhida (43), refugiados (35), imigrantes (10), migrantes (9), acnur (7), comosrefugiados (6), imigração (5), venezuelanos (4), migrante (3), imigracao (3), standwithiraqirefugees (3), migração (2), formigration (2), refuxiados (2), refugees (2), crisenavenezuela (2), imigraçãoitaliana (2)
	April	operaçãoacolhida (56), acnur (39), refugiados (25), imigrantes (17), migrantes (12), resettlement4iranian (12), imigração (10), imigrante (5), migrante (5), migração (4), cáritasbrasileira (3), eumigrante (3), muros (2)
	May	refugiados (45), xenofobianão (39), imigrantes (31), operaçãoacolhida (22), migrantes (14), imigração (12), acnur (12), migrante (8), refugiado (8), imigrante (6), xenofobia (4), visto (4), direitoshumanos (4), criminiimmigrati (3).
	June	refugiados (47), imigrantes (31), migrantes (23), imigração (21), comosrefugiados (16), operaçãoacolhida (14), diamundialdorefugiado (7), operacaoacolhida (6), visto (6), semfronteiras (6), imigrante (6), festadoimigrante (5), acnur (5), imigracao (5), venezuelanos (5), refugees (5), diadorefugiado (5), migrante (4), direitoshumanos (3), estrangeiros (3), semananacionaldosrefugiados (3).
	July	refugiados (41), imigrantes (26), imigração (18), migrantes (17), operacaoacolhida (13), migrante (9), refugiado (8), cáritas (4), venezuelanos (3), acnur (3), seawatch (2), migração (2).
	August	refugiados (42), acnur (16), imigrantes (13), migrantes (13), operaçãoacolhida (12), imigração (9), diamundialhumanitario (5), migrante (4), integração (2), xenofobia (2), imigracao (2).
	September	refugiados (63), migrantes (14), imigrantes (11),imigração (11), operaçãoacolhida (11), acnur (7), naosetrataapenasdemigrantes (6), refuxiadas (5), imigrante (4), venezuelanos (3), refugees (3), opacolhida (3), refugee (3), refugeeswelcome (2), migrante (2), xenofobia (2), aldeiadosimigrantes (2)
	October	operaçãoacolhida (76), refugiados (19), opacolhida (12), imigrantes (9), migrantes (8), operacaoacolhida (7), imigração (6), migrante (6), refugiado (5), acnur (4), caritas (4), refugiadoscristaos (3), oim (2), xenófobas (2), discriminação (2), imigrante (2).
	November	refugiados (24), assistência (17), imigração (16), imigrantes (9),opacolhida (9), operaçãoacolhida (8), acnur (8), imigracao (7), migrante (6), migrantes (5), refugeeforum (3), direitos (3), diversidadcultural (3).
	December	refugiados (23), venezuela (22), operaçãoacolhida (11), refugeeforum (10), acnur (8), imigrantes (8), migrantes (7), migrante (7), imigracao (3), onu (3), imigração (2).