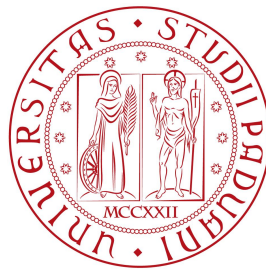


Università degli Studi di Padova  
Dipartimento di Scienze Statistiche  
Corso di Laurea triennale in

Statistica per l'Economia e l'Impresa



## **Power BI: dashboard interattive e analisi esplorativa delle vendite**

Relatore: dott. Erlis Ruli  
Dipartimento di Scienze Statistiche

Laureando: Davide Scalco  
Matricola n. 2016358

Anno Accademico 2023/2024

# Indice

<b>1</b>	<b>La rivoluzione digitale</b>	<b>4</b>
1.1	L'era dell'informazione . . . . .	4
1.2	Trasformazione digitale in Italia . . . . .	5
1.3	Le PMI italiane e il confronto con l'UE . . . . .	8
<b>2</b>	<b>Steelco S.p.A.</b>	<b>12</b>
2.1	Storia . . . . .	12
2.1.1	Lean production . . . . .	13
2.2	Mercati e prodotti . . . . .	14
2.3	Il gestionale . . . . .	16
<b>3</b>	<b>Business Intelligence</b>	<b>19</b>
3.1	Data visualization . . . . .	20
3.2	Power BI . . . . .	23
3.2.1	Power BI desktop . . . . .	24
3.2.2	Power BI service . . . . .	25
<b>4</b>	<b>Data Analysis</b>	<b>27</b>
4.1	Conoscenza dei dati . . . . .	27
4.2	CRISP-DM . . . . .	29
4.3	Analisi esplorativa . . . . .	32
4.3.1	Time series . . . . .	34
<b>5</b>	<b>Il Progetto</b>	<b>37</b>
5.1	Problema e obiettivi . . . . .	37
5.2	Dal gestionale al software . . . . .	39
5.3	Steelco nel mondo . . . . .	44
5.3.1	L'Italia . . . . .	47
5.4	Le vendite nel tempo . . . . .	48
5.4.1	AERs . . . . .	51
5.5	Panoramica di mercati e prodotti . . . . .	54

5.6	La pubblica amministrazione in Italia . . . . .	59
5.7	Conclusioni . . . . .	64

# Introduzione

Nello scenario economico moderno, gli input a disposizione per le imprese sono pressoché infiniti: una stimolante opportunità da un lato e una grave minaccia dall'altro.

Nell'epoca dei big data, riuscire a gestire questo quantitativo di informazioni a proprio favore è vitale per non sprofondare tra i competitors, sempre più rapidi e all'avanguardia. La costante innovazione, l'adozione di nuove tecnologie e solide infrastrutture ICT, alimentata dalla capacità di elaborare dati massicci, al giorno d'oggi rappresenta il requisito fondamentale per il successo delle organizzazioni.

L'intera economia mondiale, dopo l'avvento della rivoluzione digitale, ha virato verso l'innovazione tecnologica, riconoscendone le immense potenzialità. Una vera e propria corsa all'oro digitale, simile a quella americana, dove chi si è arricchito non ha estratto pepite dal terreno ma conoscenza dai dati.

La sfida imposta dai big data, non è solo una questione di quantità di informazioni da tenere sotto controllo, ma di capacità di analizzare dati complessi e riuscire ad interpretarli per trarne valore. Grazie a strumenti di data analysis e business intelligence le organizzazioni sono in grado di mettere ordine al caos informativo cui sono soggette, di cogliere opportunità nascoste e garantire supporto nelle scelte, potendo così attuare "decisioni informate".

L'estrazione di conoscenza dai dati è un'attività tanto importante quanto complessa, composta di metodologie di analisi differenti, più o meno mirate, in grado di cogliere diversi aspetti dagli input. Ciò che è fondamentale nell'esecuzione di questo processo, è avere un obiettivo chiaro, dunque, sapere cosa si vuole conoscere, a chi comunicarlo e come. La rappresentazione visiva dei dati favorisce notevolmente l'elaborazione delle informazioni da parte degli utenti finali dei progetti di analisi, spesso manager d'azienda, i quali devono riuscire a cogliere le evidenze che si vuole comunicare, per riflettere sulle scelte strategiche da attuare. La business intelligence segue questo flusso da monte a valle: dall'acqui-

sizione dei dati, all'interpretazione mediante le analisi, fino alla comunicazione agli stakeholder.

Nell'elaborato seguente verrà illustrato un esempio di integrazione dei meccanismi di BI all'interno di un'azienda, nel dettaglio, Steelco S.p.A., impresa attiva nell'ambito medico-sanitario dove ho avuto l'opportunità di intraprendere un periodo di stage, nel quale sono stato reso partecipe delle dinamiche coinvolte in questo processo. Prima di parlare del progetto svolto in azienda, sono stati approfonditi gli argomenti più rilevanti sulla quale pone le basi l'elaborato, in particolare:

Nel primo capitolo viene illustrato il processo di trasformazione digitale delle imprese, che ha posto le radici più di mezzo secolo fa ed ora si ritrova ad essere all'apice della sua espansione. L'argomento viene poi approfondito per il panorama italiano negli ultimi anni, una realtà più vicina a noi in termini di spazio e tempo.

Nel secondo capitolo si introduce l'azienda nella quale il progetto è stato svolto, ripercorrendo la sua storia partendo dalle origini, ne saranno poi descritti i fattori che le hanno permesso di espandersi fino a diventare ciò che è oggi. A seguire, vengono illustrate le aree di mercato in cui è presente e le principali soluzioni che offre, infine, si descrivono le metodologie di gestione dei dati in attività.

Nel terzo capitolo si introduce la business intelligence, l'importanza di riuscire a comunicare le informazioni correttamente attraverso gli studi sulla data visualization e il software di BI utilizzato durante il progetto: Power BI.

Nel quarto capitolo, invece, vengono descritte alcune tecniche di data analysis, tra cui, il modello di estrazione di conoscenza dai dati CRISP-DM e il ruolo importante che l'analisi esplorativa rappresenta in una grande azienda. Di quest'ultima vengono descritti gli obiettivi principali e le tecniche utilizzate, in particolare, viene approfondita l'importanza delle analisi temporali nell'individuare tendenze e stagionalità.

Il quinto capitolo riguarda il progetto svolto, del quale si racconta il flusso percorso da monte a valle. Durante il suo svolgimento, verranno ripresi in causa i concetti e le tecniche descritti nei capitoli precedenti ed applicati ad un caso studio. Nello specifico, verrà attuata un'analisi esplorativa delle vendite, pescando i dati dal gestionale d'azienda e trasformandoli in rappresentazioni grafiche mediante il software Power BI, del quale si sfrutteranno le potenzialità di analisi e di comunicazione dei report generati. Per difendere la sicurezza aziendale, i dati sensibili verranno oscurati.

# Capitolo 1

## La rivoluzione digitale

### 1.1 L'era dell'informazione

La rivoluzione digitale sviluppatasi dagli anni 70-80 dello scorso secolo in poi ha dato origine all'epoca post-industriale, una realtà in cui il bene immateriale, come l'informazione, prevale rispetto all'industria.

Il mercato delle tecnologie dell'informazione e della comunicazione – ICT sta vivendo una crescita incontrollabile, con risvolti socio-economici che hanno portato ognuno di noi a convivere ogni giorno, nel tempo libero e nel lavoro.

L'interconnessione permette alle informazioni di viaggiare, su scala globale, a tempo di un battito di ciglia e di innescare processi automatizzati. Basti pensare, banalmente, all'acquisto di un libro tramite e-commerce: non c'è la necessità di recarsi presso il negozio, né di sapere dove si trovi o pagare con denaro fisico. È intuibile da questo semplice esempio, come la rivoluzione digitale abbia modificato il modo di fare impresa e la necessità di queste ultime di adattarsi per sopravvivere nel mercato. Le innovazioni digitali alimentano ogni fase del business, dalla sponsorizzazione alla vendita, mediante la connessione tra prodotti, processi, macchine e persone che porta alla generazione di dati, i quali vengono definiti "il petrolio dell'era digitale".

Si dice che il 90% dei dati mondiali sia stato creato negli ultimi due anni, un dato evocativo, attraverso il quale viene descritta l'immensità del fenomeno e della sua crescita, della quale però, non tutte le aziende sono consapevoli. Quante riusciranno a cavalcare lo tsunami dei big data e quante invece verranno travolte?

Celebre è l'esempio di Blockbuster, azienda che arrivò a contare più di novemila store in tutto il mondo ma non riuscì a mantenere il passo con i tempi: quando la concorrenza si spostò dal noleggio di film offline

allo streaming online precipitò senza più rialzarsi, sovrastata dai nuovi colossi come Netflix.

Come Blockbuster si possono citare moltissimi altri esempi di imprese soffocate da un modo di fare business sempre più pretenzioso, nate o cresciute in un mondo in cui “velocità” è la parola d’ordine. Velocità come rapidità di adattamento alle esigenze del cliente o del mercato, velocità di comunicazione, per non finire nell’ombra del cumulo di notizie a cui siamo costantemente soggetti e velocità in termini di efficienza, come riduzione del lead time dei processi aziendali dall’ideazione di un prodotto alla sua spedizione.

Per definizione, l’era dell’informazione nacque quando la velocità di trasferimento delle informazioni divenne maggiore di quella del normale trasferimento fisico nel mondo reale. (Wikipedia, 2023a)

Capiamo, dunque, come nel mercato moderno l’uomo abbia bisogno di equipaggiarsi dei giusti strumenti per controllare e sfruttare l’onda di dati che si sta abbattendo su di noi, per cavalcarla, così da vincere tra i competitors, consapevole che da solo non potrebbe farcela e ne verrebbe travolto.

## 1.2 Trasformazione digitale in Italia

In seguito al primo censimento permanente delle imprese nel 2019, condotto dall’Istituto nazionale di statistica – Istat, venne pubblicato un comunicato nel quale è stata approfondita la situazione che le imprese italiane stavano vivendo in merito alla trasformazione digitale.

Questa analisi è stata condotta sfruttando indicatori statistici riferiti per lo più all’adozione e all’utilizzo di tecnologie ICT, evidenziando il processo di evoluzione delle organizzazioni che mirano al raggiungimento della “digital maturity”, intesa come condizione per ottimizzare i flussi informativi all’interno dell’impresa, con effetti positivi in termini di efficienza e competitività.

Sono stati monitorati, nel dettaglio, investimenti in tecnologie digitali di tipo infrastrutturale come connessione a Internet o acquisto di servizi cloud e investimenti specializzati quali: Big Data, applicazioni di IoT, stampa 3D, robotica, ecc.

In primis, è stato possibile definire 4 profili d’impresa, definiti sulla base della composizione degli strumenti tecnologici di portafoglio. Da questa traspare come, tendenzialmente, di pari passo alla crescita del business

cresca la necessità di investire negli strumenti del futuro, infatti, tra le imprese con 10-19 addetti ad effettuare almeno un investimento tra le tecnologie sopra citate è il 73,2%, che sale al 97,1% per le grandi imprese con più di 500 dipendenti.

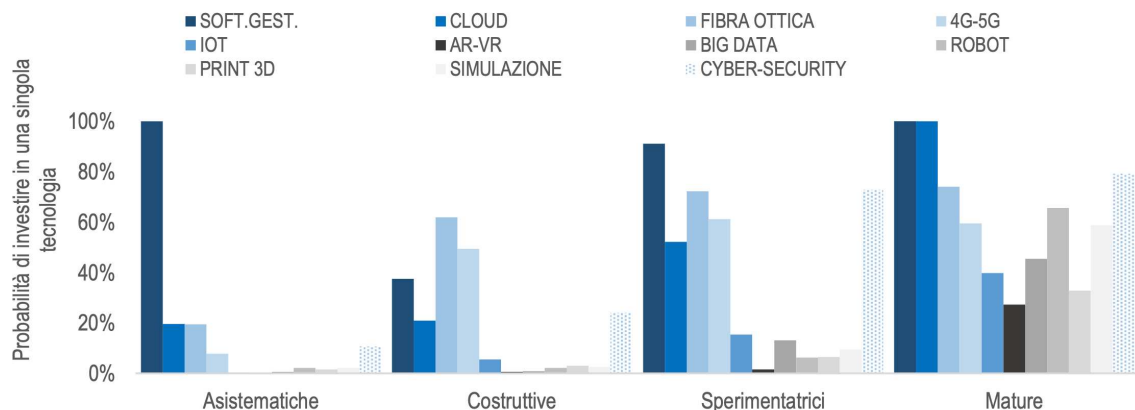


Figura 1.1: Imprese con almeno 10 addetti in base alla probabilità di investire in tecnologie digitali e al loro grado di maturità digitale.

Il primo gruppo comprende le imprese definite “asistematiche”, caratterizzate per l’adozione di almeno un software gestionale e per la presenza limitata di investimenti in tecnologie strutturali, dovuti ad una dimensione aziendale ridotta oppure ad una predisposizione del settore a non investire.

Nel secondo insieme, circa il 45% del totale, troviamo realtà “costruttive”, per le quali notiamo l’interesse di usufruire di connessione a Internet in mobilità (4G-5G e fibra ottica), potendo così, sfruttare le varie tecnologie di sensoristica in remoto. È interessante notare come, già questo gruppo di imprese, reputi essenziale l’investimento in sicurezza: in generale, al crescere del livello di maturità digitale cresce nelle imprese anche l’esigenza di mettere in sicurezza i propri apparati.

Nel terzo gruppo ci sono le imprese “sperimentatrici”, ovvero la maggioranza delle realtà con più di 100 dipendenti, le quali, appunto, sperimentano diverse combinazioni di soluzioni informatiche, giungendo ormai alla soglia della digital maturity. Qui vediamo i primi significativi sforzi nella valorizzazione dei flussi informativi attraverso big data, simulazione e robotica.

Infine, il quarto insieme comprende le imprese digitalmente “mature” in quanto realizzano investimenti in ogni categoria di tecnologie digitali, rappresentano il modello di innovazioni che il mercato imprenditoriale



italiano dovrebbe seguire, pur essendo solo il 3,8% del totale.

In seguito, invece, viene svolto un interessante approfondimento per cogliere dalle protagoniste una considerazione sui risvolti ottenuti usufruendo di piattaforme digitali, un quadro che ogni dirigente di impresa dovrebbe tenere in memoria per comprendere le potenzialità ottenute, in generale, cavalcando l'onda delle informazioni.

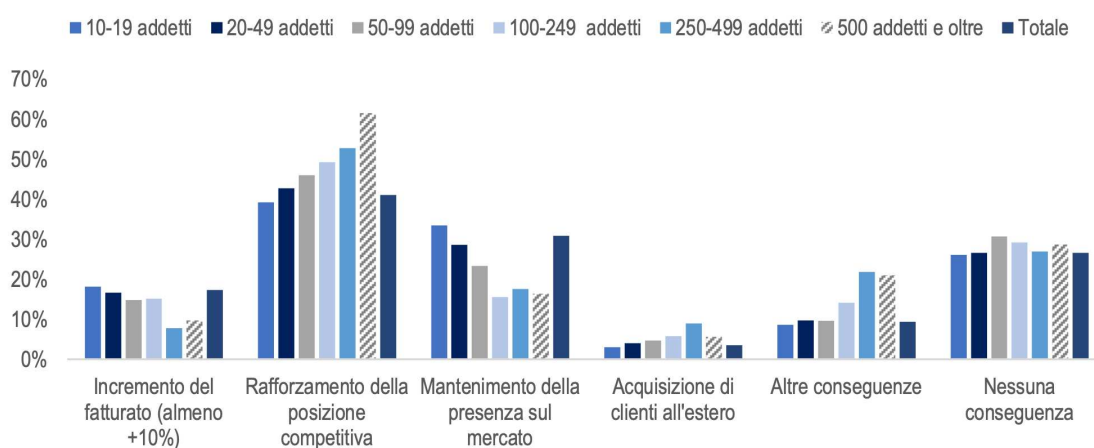


Figura 1.2: Consequenze per le imprese con almeno 10 addetti dovute all'utilizzo di piattaforme digitali, percentuale sul totale delle imprese per classe di addetti.

Da questo studio risultano le seguenti evidenze in merito alle opportunità create dalle piattaforme digitali:

Il 17,3% delle imprese con 10+ addetti risponde in merito ad un incremento del fatturato totale conseguito maggiore del 10%.

Il 41,1% delle imprese evidenzia un rafforzamento della propria posizione competitiva (con un picco del 61,5% per le imprese 500+), ottenuto banalmente dalla presenza sul sito web di una piattaforma digitale. Questa si può considerare una dinamica di mantenimento per il 30,8% delle piccole imprese che vogliono risultare stabili nel mercato.

Si scorge, inoltre, un dato negativo sulla potenzialità delle piattaforme di ampliarsi a livello globale, sfruttata solamente dal 3,5% delle imprese. (Istat, 2019)

## 1.3 Le PMI italiane e il confronto con l'UE

Se dal primo comunicato si evince come la digitalizzazione sia un processo fruttuoso ma non ancora radicato e in continua evoluzione, a distanza di tre anni ne abbiamo la conferma.

Con il censimento permanente delle imprese concluso nel 2023, l'Istat ha potuto elaborare un breve report, sulla falsa riga del precedente, evidenziando come la corsa alla digitalizzazione in Italia non stia seguendo la retta via, pur avendo il piano nazionale di ripresa e resilienza – PNRR più cospicuo dell'intera Unione europea (191,5 miliardi di EUR, di cui il 25,1 % destinato alla transizione digitale nel 2022).

In particolare, si vuole enfatizzare come le attività di innovazione delle imprese siano in calo rispetto al triennio precedente e rispetto alla media UE, in particolare per le piccole e medie imprese – PMI, che risultano di conseguenza meno produttive.

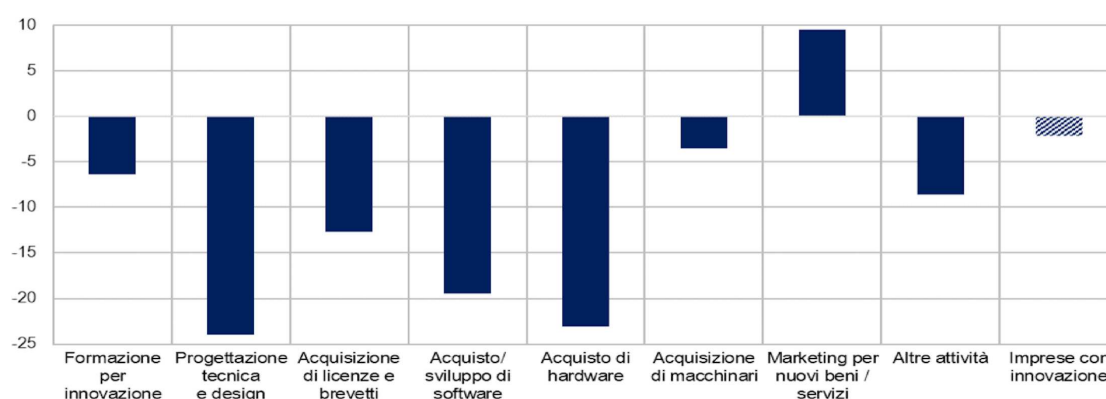


Figura 1.3: Variazione percentuale delle attività di innovazione per il totale delle imprese.

Dalla figura riassuntiva seguente è evidente come, sebbene la percentuale complessiva di imprese che hanno svolto almeno un'attività di innovazione sia diminuita solo del 2,1% nel 2022 rispetto al 2018, le attività più orientate al futuro, quali investimenti software e hardware, siano state abbandonate. (Istat, 2023a)

Il progetto viene presentato dal titolo seguente: “Nel 2023 piccole e medie imprese ancora penalizzate in attività specialistiche di digitalizzazione”, un focus intenzionale sul divario di rivoluzione digitale presente tra le grandi imprese (>250 addetti) e le PMI (10-250 addetti), che assume particolare rilievo poiché queste ultime rappresentano il 20,7% delle

attività ed occupano il 42,6% degli occupati in Italia.

Attraverso un indicatore composto, il Digital Intensity Index – DII, viene valutato il comportamento delle aziende italiane in merito a 12 caratteristiche chiave per definire il grado di digital maturity.

Ne risulta, come sopra citato, un delta significativo a scapito delle PMI nelle attività che richiedono maggiore competenza specialistica come per l'analisi di dati (25,7% le PMI e 74,1% le grandi imprese) e in quelle più legate alla complessità organizzativa e dimensionale come per l'utilizzo di software gestionali ERP e CRM (rispettivamente 41,4% contro 85,0% e 18,5% contro 53,4%), interessante è anche il dato sull'utilizzo di almeno due social media (28,0% e 55,0%). Si nota, inoltre, una lacuna nell'economia generale per quanto riguarda l'utilizzo di e-commerce per vendite online e l'integrazione di sistemi di intelligenza artificiale – AI. (Istat, 2023b)

Nel dettaglio:

<b>Indicatori della transazione digitale</b>	<b>10+ Addetti</b>	<b>250+ Addetti</b>
<i>1. Addetti connessi &gt;50%</i>	47,1%	58,3%
<i>2. Utilizzo di AI</i>	5,0%	24,1%
<i>3. Velocità BL fissa in download &gt;= 30 Mbit/s</i>	84,8%	96,9%
<i>4. Analisi dei dati effettuata all'interno o all'esterno dell'impresa</i>	26,6%	74,1%
<i>5. Acquisto di servizi di cloud computing</i>	61,4%	85,7%
<i>6. Acquisto di servizi di cloud computing sofisticati o intermedi</i>	55,1%	80,1%
<i>7. Utilizzo di social media</i>	57,3%	81,4%
<i>8. Utilizzo di software ERP</i>	42,2%	85,0%
<i>9. Utilizzo di software CRM</i>	19,2%	53,4%
<i>10. Utilizzo di almeno due social media</i>	28,5%	55,0%
<i>11. Valore vendite online &gt;=1% dei ricavi totali</i>	13,5%	37,5%
<i>12. Vendite web &gt;1% dei ricavi totali e B2C &gt;10% dei ricavi web</i>	8,3%	13,7%

Figura 1.4: 12 indicatori analizzati, per le imprese a forza lavoro 10+ e 250+.

È noto come, man mano che lo stivale Italiano viene percorso dal tacco verso l'alto si evidenzia un notevole progresso economico e tecnologico a scapito, dunque, del sud. Basti pensare che nel 2023, il 51,8% della

totalità delle imprese è attiva al nord Italia. Il confine di questo trend di crescita, tuttavia, non si ferma sulle Alpi, ma continua espandendosi nell'Europa intera.

La Commissione Europea si occupa ogni anno di monitorare gli stati membri dell'UE rispetto a vari indicatori, di nostro interesse è il DESI – Digital Economy and Society Index, e ciò che ne risulta dall'ultimo comunicato (2022) è allarmante.

L'Italia si colloca al 18° posto tra i 27 stati membri dell'unione europea.

	Italia	UE
	posizione in classifica	punteggio
<b>DESI 2022</b>	<b>18</b>	<b>49,3</b>
		<b>52,3</b>

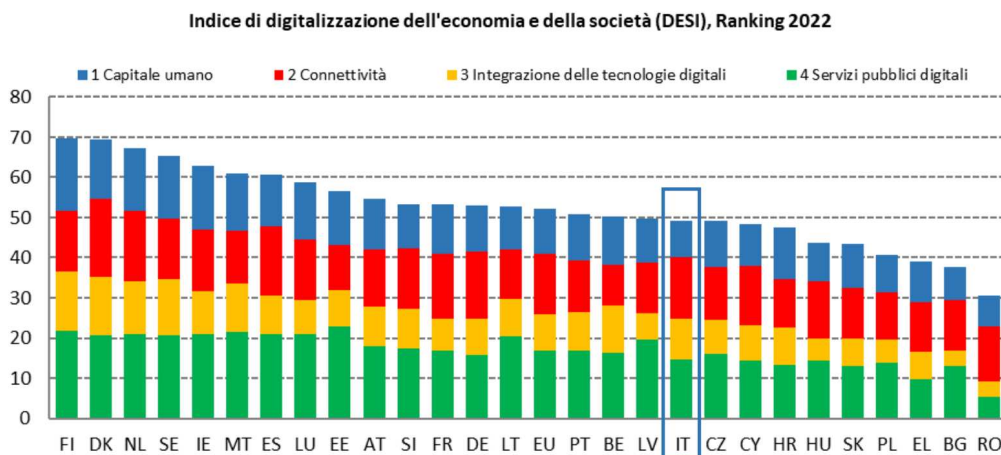


Figura 1.5: Punteggio DESI dei paesi UE, contributo percentuale suddiviso per categoria di digitalizzazione.

La scomposizione suggerita dallo studio eseguito, mette in rilievo una forte sofferenza da parte del capitale umano, in primis, rispetto alle diverse categorie, seguito dall'istituzione di servizi pubblici digitali.

Come analizzato precedentemente, le imprese italiane hanno ancora grandi passi avanti da fare in termini di integrazione di tecnologie digitali, soprattutto specializzate, nonostante ciò, il livello base di digitalizzazione delle PMI italiane è sopra la media europea (60% delle imprese italiane contro il 55% europeo) e lo sono di gran lunga anche le integrazioni di servizi cloud e la fatturazione elettronica. Anche la connettività è un punto forte, in questo campo l'Italia si colloca al 7° posto tra gli stati membri dell'UE.

Vale la pena rimarcare la necessità di formazione delle persone per colmare il vuoto che presenta il nostro paese in termini di capitale umano, qui l'Italia si colloca al 25° posto su 27, un dato agghiacciante che ci serve come spunto di riflessione. La maggior parte delle iniziative messe in piedi e pianificate per i prossimi anni verte proprio su questo fattore di istruzione e formazione della popolazione in generale. (CE, 2022)

La situazione attuale si può riassumere secondo alcuni indicatori interessanti:

<b>Misure per il capitale umano</b>	<b>ITALIA</b>	<b>MEDIA UE</b>
<i>1. Individui in possesso di competenze digitali di base</i>	46%	54%
<i>2. Individui in possesso di competenze digitali superiori</i>	23%	26%
<i>3. Individui in possesso di competenze di base per la creazione di contenuti digitali</i>	58%	66%
<i>4. Specialisti ICT sul totale degli occupati di fascia 15-74 anni</i>	3,8%	4,5%
<i>5. Imprese che forniscono formazione in materia di ICT</i>	15%	20%
<i>6. Laureati nel settore ICT</i>	1,4%	3,9%

Figura 1.6: Indicatori percentuali di misura della digitalizzazione del capitale umano, in Italia e in Europa.

Da questa semplice presentazione delle imprese italiane nel mercato ICT abbiamo potuto vedere come, per una serie di ragioni, le dinamiche di investimento del governo e dei privati per le proprie realtà siano ancora in ampia crescita. Ciò che ne risulta oggi è una grande varietà di innovazione, sfruttata, in maniera direttamente proporzionale alle dimensioni, da aziende che hanno deciso di mettersi in gioco, sentendo l'esigenza di innovarsi, sollecitate dai competitor del proprio settore.

# Capitolo 2

## Steelco S.p.A.

In questo capitolo verrà illustrata la realtà nella quale ho intrapreso il mio stage, essa, è un esempio di come la costante innovazione e la voglia di progredire verso un miglioramento sistematico siano la chiave del progresso nel ventunesimo secolo.

Steelco S.p.A. è un'azienda italiana che figura tra i leader mondiali nella produzione di macchine e sistemi per lavaggio, disinfezione e sterilizzazione di apparecchiature mediche in ambito sanitario, di ricerca scientifica e nell'industria farmaceutica.

Il settore di competenza e le dimensioni aziendali raggiunte, hanno permesso all'impresa di raggiungere negli anni una maturità digitale tale da autodefinirsi "anticipatori del futuro", una corsa all'innovazione indotta dall'esigenza di adottare tecnologie in grado di fornire soluzioni specialistiche per ogni richiesta e farlo in maniera efficiente.

### 2.1 Storia

Nasce nel 1998 come realtà che lavora l'acciaio inossidabile, con sede produttiva a Riese Pio X (TV), in Veneto, per poi espandersi fino ad essere oggi l'azienda con il tasso di crescita più rapido sul suo mercato, con un CAGR – Compounded Average Growth Rate del 13,4% tra il 2019 e il 2023.

Nel 2001 l'azienda comincia la sua crescita, con l'introduzione della produzione di termo-disinfettori di strumenti ospedalieri, che dal 2005 diventa esponenziale. Dopo pochi anni riesce ad acquisire un'esperienza tale da permetterle di competere nei mercati farmaceutico e di ricerca, vantando un portafoglio diversificato di prodotti.

Nel 2013 avviene l'acquisizione di Icos Pharma S.p.A., azienda specializ-

zata nella produzione di autoclavi, un tassello essenziale per permettere a Steelco di offrire un servizio completo: lavaggio, disinfezione e sterilizzazione. Nel 2017 il gruppo Miele entra nel capitale sociale d'azienda, attraverso questa importante partnership con il colosso tedesco da più di 5 mld di fatturato e oltre 20000 dipendenti in tutto il mondo, avviene un'ulteriore spinta per l'espansione del marchio all'estero.

Oggi Steelco S.p.A. conta 10 stabilimenti diversificati per i settori di competenza, 14 filiali di vendita ed assistenza in tutto il mondo, oltre 175 brevetti registrati per un totale di quasi 800 dipendenti e più di 200mln di fatturato lo scorso anno.

### 2.1.1 Lean production

Dopo essere passata sotto l'ala del gruppo Miele, l'azienda ha tratto grossi benefici in termini produttivi ed economici, i quali però, necessitano di un sistema di gestione ben strutturato.

Per ottimizzare il carico di lavoro si è deciso di abbracciare la filosofia Kaizen del miglioramento continuo: dagli idiomi “Kai”, che significa cambiamento, e “Zen”, che significa bene, assumendo così il significato di “cambiamento verso il meglio”. Venne quindi instaurato nel dicembre del 2016 il “Kaizen Promotion Office” – KPO, un mezzo per una collaborazione trasversale tra i reparti con lo scopo di rilevare problemi e/o migliorie, attraverso il coinvolgimento di più funzioni, da analizzare e risolvere nel minor tempo possibile.

Grazie alla diffusione del “Lean Thinking” tra i dipendenti è possibile migliorare l'efficienza dei processi produttivi ed organizzativi, essi conoscono il proprio settore operativo nei minimi dettagli e svolgono, dunque, un ruolo cruciale nell'identificazione degli sprechi, in termini di tempo, materiali, risorse umane e finanziarie. Quest'attività, se fatta centralmente, perderebbe di potenza proprio poiché la supervisione, seppur minuziosa, non potrà mai portare all'accorgimento di ogni falla nel sistema quanto l'identificazione da parte degli attori stessi del teatro in questione. L'essenza del pensiero “Lean” – snello, nasce con la rivoluzione produttiva adottata da Toyota negli anni 50, dalla mente del membro del consiglio esecutivo Taiichi Ono e si basa su tre fondamentali principi:

1. Focus sul cliente e sul flusso di valore (Value Stream)
2. Eliminazione degli sprechi (Muda)
3. Miglioramento continuo (Kaizen)

In opposizione alla produzione di massa da catena di montaggio di Henry Ford, il Toyota Production System prevede di “fare di più con meno”. (Wikipedia, 2022)

Ci riesce grazie ad un approccio Pull anziché Push, dunque di produzione sulla base della domanda dei clienti anziché del mercato, e all’utilizzo di strumenti di comunicazione, come i cartellini Kanban, per regolare l’acquisto e la movimentazione di materiali.

L’implementazione della produzione Lean ha permesso a Steelco di incrementare la produzione e ridurre notevolmente il lead time, permettendo di fornire servizi personalizzati in tempi ridotti; ciò è risultato essenziale, ad esempio, per superare la grande richiesta d’aiuto da parte degli ospedali durante la pandemia Covid-19.

## 2.2 Mercati e prodotti

Questa grande azienda è presente in diverse aree di mercato, classificate nelle seguenti categorie sulla base dei prodotti offerti:

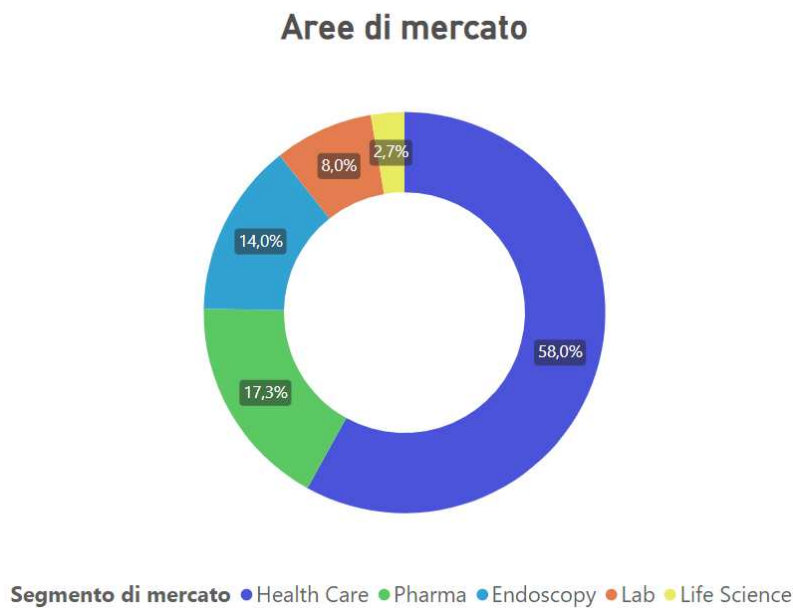


Figura 2.1: Donut chart delle aree di mercato suddivise per fatturato derivato.

Si può notare come nell’ambito di cura della persona – Healthcare l’azienda sia più presente, questo è suggerito, dalla presenza delle vendite più massive in termini sia di quantità che di prezzo della singola macchina, nonché dal fatto che rappresenta l’area di mercato nella quale



lavora da più tempo.

Steelco, ad oggi, offre una varietà di prodotti nel mercato che supera le 180 differenti unità di prodotti finali, una diversificazione possibile grazie alla versatilità nel riuscire ad offrire la soluzione giusta per ogni cliente. La produzione “Push” e la moltitudine di brevetti posseduti, sono caratteristiche fondamentali che le hanno permesso di essere efficace nei prodotti custom, garantendole un costante afflusso di nuovi clienti che a causa delle loro richieste stringenti non trovano soluzioni standard nel mercato. Un altro importante applicativo della produzione su misura è stato raggiunto in seguito alla partnership con Miele, ovvero la progettazione personalizzata di centrali di sterilizzazione – CSSD, dei progetti molto complessi ad elevato valore economico.

Sinteticamente, le soluzioni Steelco si possono suddividere secondo due fini funzionali: il processo di lavaggio-disinfezione e la sterilizzazione. Di seguito vengono presentati i principali prodotti commerciati, raggruppati per macro-famiglia di appartenenza, le stesse suddivisioni categoriche saranno poi implementate durante gli studi sviluppati.

Per lavaggio-disinfezione, solitamente, i prodotti vengono identificati dalla sigla finale WD – Washer-Disinfectors, le macro famiglie produttive più diffuse sono:

- Trolley WD o lava-carrelli, sono i sistemi più grandi, utilizzati per carrelli operatori o simili
- Large chamber WD, rappresentano la soluzione più spaziosa per il lavaggio di strumenti
- Small-medium WD, sono le soluzioni più ridotte in termini di capacità di contenimento di strumenti
- Bedpan WD o lava-padelle, sono termo-disinfettori di articoli quali padelle, pappagalli o recipienti
- AERs, sono le soluzioni per il lavaggio e disinfezione di endoscopi

Per la sterilizzazione:

- Steam sterilizers o sterilizzatori a vapore, racchiudono la maggioranza delle soluzioni di sterilizzazione di dispositivi medici
- Small sterilizers, sono le autoclavi da banco ad una unità di scarico, per piccole cliniche

- Low-temp sterilizers o sterilizzatori a bassa temperatura, sfruttano tecnologie al plasma

## 2.3 Il gestionale

In realtà come Steelco, o più in generale per le grandi imprese produttive, risulta vitale, per una gestione efficiente e competitiva, mantenere sotto controllo i dati ed i processi.

Per farlo è necessario usufruire di un gestionale, o, meglio ancora, di un software per la gestione ed automazione di processi aziendali e funzioni interne: l'ERP – Enterprise Resource Planning. Esso, non svolge solamente la funzione di raccoglitore centralizzato di informazioni dalle diverse aree funzionali, ma, controlla e gestisce i flussi aziendali grazie ad un'ampia flessibilità di configurazione per adattarsi ad essi e allo stesso tempo sfrutta modelli organizzativi ottimizzati che si rivelano utili soprattutto quando l'azienda deve innovare la propria organizzazione.

L'utilizzo di questo applicativo porta tre principali vantaggi che durante la mia esperienza ho potuto notare con trasparenza:

1. Accesso e condivisione totali in tempo reale: l'ERP è essenzialmente una piattaforma nella quale convergono informazioni, aggiornate, da tutte le funzioni aziendali: catena di approvvigionamento, produzione, commercio, finanza, risorse umane e report vari.
2. Automazione dei processi ed aumento dell'efficienza: svolge funzionalità di analisi ed elaborazione dei dati che gli permettono di capire quando e quali attività è più funzionale svolgere, per poi programmare un workflow da attuare capace di ridurre i tempi e gli errori legati alla gestione manuale. Potendo dialogare direttamente con la filiera produttiva e commerciale si occupa ad esempio di: inserimento delle fatture, controllo delle scorte e gestione degli ordini.
3. Controllo e visibilità: grazie a strumenti di data mining e business intelligence integrati, restituisce proiezioni utili al monitoraggio delle attività, considerando gli indicatori chiave di prestazione (KPI - Key Performance Indicators) permette al management di migliorare i processi decisionali strategici ed accorgersi tempestivamente di eventuali anomalie da risolvere prima che causino scompiglio.

In Steelco, al mio arrivo, si stava seguendo un processo di innovazione del sistema di gestione, poiché al momento non risultava ben centralizzato e ciò si rifletteva nel progetto di globalizzazione del business.

Attualmente l'azienda, come sopra citato, conta diversi stabilimenti, di questi: 2 sono sedi produttive, i restanti si occupano di sales & service. Per quanto riguarda le sedi produttive, l'Headquarter – HQ, è la sede di produzione per il lavaggio e la disinfezione, dove ho svolto lo stage, a Riese Pio X (TV), mentre a Cusano di Zoppola (PN) si trova la sede di produzione per la sterilizzazione: Icos Pharma S.p.A.

Nelle sedi produttive, il sistema di riferimento utilizzato è il gestionale IBM i – AS/400.

Nasce nel 1988 come computer di supporto per la struttura informatica aziendale; composto da un server (AS400 per l'appunto), un sistema operativo specifico (OS400), un database (DB2) e un linguaggio di programmazione nativo (RPG). Nei seguenti 30 anni IBM ha fornito aggiornamenti per modernizzare la piattaforma, rendendola in grado di offrire un'ampia gamma di funzionalità, quali: l'hosting di applicazioni aziendali (come software ERP), la gestione del database e l'elaborazione delle transazioni. Velocità di elaborazione, sicurezza e stabilità di questa piattaforma hanno spinto molte strutture sanitarie, banche e altre realtà che gestiscono una mole importante di dati sensibili, a dotarsi di questo strumento. (Borghetti, 2022)

Al contempo però, viene considerato un “sistema legacy”, in quanto alcuni aspetti risultano oggi obsoleti:

- L'accesso ai programmi con interfaccia a linea di comando Telnet (la vista “nero verde”).
- La mancanza sempre più elevata di personale che conosce il linguaggio RPG.
- Non garantisce tutte le funzionalità dei sistemi ERP.

Per questo motivo, spinte anche dall'influenza del gruppo Miele, le filiali di sales & service utilizzano sistemi SAP: una software house tedesca, leader di mercato nell'ambito degli ERP, che sfrutta funzionalità di AI e machine learning. Nello specifico, SAP-B1, un software ad interfaccia grafica “look and feel”, intuitiva, simile a quella dei programmi pc, suddivisa in due moduli:

- Sales, per visualizzare qualsiasi dato o report di tutto ciò che entra ed esce dall'azienda.

- Service, che sfrutta tool come “MyPlan” o “MySAPP” per i servizi post-vendita.

La direzione che sta prendendo Steelco assieme al gruppo Miele è quella di uniformare la gestione dei processi aziendali sotto le tecnologie SAP, tuttavia, l’inserimento di un nuovo ERP è una scelta molto delicata perché ha un’alta percentuale di fallimento ed in buona percentuale trascina al fallimento anche l’azienda stessa (nel 17% dei casi), dunque la roadmap prevede di effettuare il passaggio non prima dei prossimi tre anni.

Affinché questa transazione possa avvenire, devono essere condotti test per ogni processo critico e simulazioni di go-live per la sincronizzazione dei vecchi sistemi con il nuovo, inoltre, vanno risolti eventuali problemi di migrazione dei dati per far sì che il risultato non sia inquinato di errori. Nell’ufficio IT nel quale sono stato inserito durante il mio percorso, ho potuto vedere questo flusso di azioni svolte dai miei colleghi e ho preso coscienza dell’importanza della pulizia dei dati e della comunicazione tra le varie funzioni aziendali. Due lezioni mantra per la figura del business intelligence analyst, che ottiene dai dati l’unica fonte di conoscenza e fa della corretta comunicazione il fine del suo lavoro.

# Capitolo 3

## Business Intelligence

Per Business Intelligence si intende un insieme di processi e funzionalità per la raccolta e l'analisi di dati storici e correnti, utili a delineare un quadro della situazione aziendale e permettere al management di attuare soluzioni strategiche.

Per trasformare i dati grezzi in informazioni dettagliate, di facile comprensione e utilizzabili da tutti gli utenti dell'azienda si seguono quattro passaggi chiave: (PowerPlatform, 2023)

1. Raccolta e trasformazione dei dati da più origini.

In genere viene utilizzato il metodo Extract, Transform and Load – ETL: un processo di integrazione che combina i dati provenienti da più origini in un singolo archivio dati coerente. Esso prevede l'estrazione dei dati da varie origini (SQL server, sistemi ERP, CRM, ...) che vengono poi caricati in una staging area, dove avviene la trasformazione, una fase di vitale importanza. Qui avvengono i controlli per garantire la qualità dei dati e dell'analisi di conseguenza: filtraggio, pulizia, modifica e autenticazione dei dati. Una volta rimodellati verranno caricati presso il data warehouse di destinazione, un'ubicazione centrale dalla quale gli analisti potranno attingere per attuare le loro analisi.

2. Individuazione delle tendenze e delle incoerenze.

Dopo aver avuto la certezza di poter attingere ad una fonte certa, vengono utilizzati tutti gli strumenti di BI a disposizione per poter scavare nei dati: esplorativi, descrittivi, statistici e predittivi. L'EDA – Exploratory Data Analysis per apprendere dai dati, descriverli ed identificare ricorrenze; l'analisi statistica per poterli interrogare, definire modelli, testare ipotesi e prevedere accadimenti.

3. Utilizzo della data visualization per presentare i risultati.

La creazione di report di business intelligence utilizza la data visualization per semplificare la comprensione e la condivisione dei risultati, uno strumento che identifica i metodi ottimali per illustrarli, in modo da consentire ai lettori di avere chiara la panoramica d'azienda, tra questi: dashboard di dati interattivi, grafici, diagrammi e mappe.

4. Azione in base alle informazioni dettagliate in tempo reale.

I primi tre punti permettono di trarre conoscenza da ciò che l'azienda è stata in passato ed è oggi e gettano le basi per la fase decisionale finale, ovvero, su come muoversi per definire come sarà in futuro. Questo dipenderà dalle decisioni che i responsabili del business adotteranno, informati dagli strumenti di BI utilizzati ed aiutati da funzionalità e automatismi di supporto, come i DSS – Decision Support System, per integrare correzioni strategiche in tempo reale, migliorando inefficienze, problemi dell'offerta, del cliente e variazioni del mercato.

## 3.1 Data visualization

Come può un'analisi essere funzionale per l'azienda se viene capita solamente da chi l'ha svolta?

Attraverso la data science, gli analisti aziendali, o gli strumenti integrati di elaborazione delle informazioni, riescono a trarre conoscenza dai dati puri, che, così come si presentano all'interno del data warehouse, non hanno nessuna valenza benefica per il business prima di essere analizzati.

Una volta fatto però, si riscontra un secondo problema, di tipo comunicativo, tra l'esperto che ha svolto ricerche funzionali: sollevando ipotesi, evidenziando tendenze, riscontrando anomalie o effettuando previsioni; e il manager del settore interessato durante l'indagine. Quest'ultimo, ricoprirà una carica di spessore nell'ambito di competenza che sarà pane per i suoi denti: ne conosce il linguaggio, le dinamiche e probabilmente punti deboli e di forza.

Tuttavia, in quanto soggetto ai limiti umani, non può tenere a mente tutto ciò che succede e avere una chiara visione del presente e del passato, ma, soprattutto, per conoscere la vera entità di una minaccia o di un'opportunità dovrebbe essere in grado di quantificarla.

Il data analyst si occupa proprio di quantificare, attraverso i dati, le

informazioni. Lo fa sfruttando indici, modelli, forecast e altri strumenti statistici che il manager di settore non ha le competenze per decifrare. Per questo motivo, un business intelligence analyst, per svolgere bene il proprio compito, deve produrre dei risultati in grado di essere compresi da chiunque debba riceverli e per farlo si avvale del mezzo di comunicazione più antico e universale esistente: l'arte figurativa.

La Data Visualization, o rappresentazione grafica dei dati, è una forma scientifica di arte figurativa, utilizzata per tradurre le informazioni contenute nei dati e comunicarle attraverso strumenti visivi. La sua implementazione è in continuo sviluppo, di pari passo alla psicologia cognitiva sulle associazioni delle immagini mentali e alla scienza dei dati e delle rappresentazioni di statistiche descrittive.

Ogni strumento visivo si occupa di illustrare uno specifico argomento, nella maniera più opportuna, le tecniche più utilizzate sono: grafici, diagrammi, infografiche, matrici, mappe e schede KPI.

Essi possono essere raccolti in un unico cruscotto, per raccontare una storia o rispondere a una domanda rispetto ad un argomento comune, utilizzando le dashboard, delle interfacce grafiche contenenti più oggetti visivi in un solo pannello.

Dagli studi finora pubblicati possiamo definire delle combinazioni di strumenti che funzionano meglio, in base alle loro caratteristiche grafiche, per descrivere ogni specifico obiettivo visivo:

- Per visualizzare un ammontare, inteso come volume totale di una quantità rispetto ad una variabile categorica utilizziamo bar chart o dot plots; nel caso in cui vi siano più categorie: grouped bars, stacked bars oppure heatmaps.
- I grafici di proporzione sono utili per rappresentare, appunto, proporzioni del totale, in valori assoluti o percentuali, di una quantità rispetto ad una suddivisione in gruppi. Tra questi abbiamo pie, donut e bar chart quando ho una sola suddivisione in gruppi. Grouped bars, stacked bars e stacked densities per due variabili categoriali.
- Quando invece abbiamo suddivisioni categoriche che si sovrappongono (una unità fa parte di più gruppi), per le quali ci interessa anche la relazione che intercorre tra le variabili che formano i gruppi, allora parliamo di “nested proportion”, dove sono efficaci: mosaic Plots, treemaps e parallel sets.

- Se volessi graficare la distribuzione di una particolare variabile in un dataset ricorro a grafici quali: histogram, density plot, cumulative density, Q-Q plot; per la rappresentazione di più densità invece, è molto diffuso l'utilizzo di boxplot, ma anche violins e ridgeline plot.
- Solitamente è interessante capire se esistono (e in che misura) relazioni tra due o più variabili quantitative. Se si tratta di uno studio sulla relazione a due, di influenza una sull'altra, il grafico di riferimento è lo scatterplot, che può integrare delle modifiche per rappresentare dati appaiati o suddivisi rispetto ad una variabile categoriale (paired scatterplot e bubble chart).  
Qualora si voglia studiare contemporaneamente la relazione tra più di due variabili si può ricorrere a tecniche statistiche per la riduzione delle dimensionalità oppure sfruttare grafici più complessi come correlogrammi o matrici di scatterplot.
- I grafici di tendenza rappresentano i dati correlati alla dimensione temporale, qui la direzionale naturale dettata dal tempo impone un'evoluzione ai dati, i quali risultano facilmente leggibili poiché viene istintivamente evidenziato un successore ed un predecessore per ogni punto. Questi sono utilizzati per raffigurare serie storiche, anche multiple, ovvero per una variabile suddivisa in gruppi. Spesso sono anche accompagnate da funzioni che descrivono l'andamento dei dati nel tempo come medie mobili, trend e funzioni di regressione. I principali grafici di tendenza sono line graph e area graph.  
Se, però, la variabile tempo forma pochi raggruppamenti, come la visione di una quantità in un range di 5 anni, si può ricorrere a grouped bars, che ne evidenziano in maniera più netta le differenze.
- Infine, per comunicare le informazioni geo localizzate, infine, si ricorre alle mappe attraverso strumenti quali: choropleth e proportional symbol map.

La rappresentazione grafica dei dati è uno strumento essenziale per la comunicazione dei risultati di analisi, ricordando che il progetto di un BI analyst non verrà letto da un altro BI analyst, bensì da un manager, o da un qualsiasi addetto, inesperto, per cui vige la regola “less is more”. A significare che la presentazione non debba essere decifrata da chi la visualizza, inserendo il maggior numero di informazioni possibili, al contrario, dev'essere sintetica, chiara e memorabile. (Wilke, 2019)  
Perché questo avvenga è necessario comprendere chi sarà l'utente finale



e quali informazioni deve cogliere, per definire cosa graficare; sfruttando anche tutte le potenzialità “estetiche” possibili, come il corretto utilizzo di: titolo, descrizione, font, colori, unità di misura, sfondo, griglia, assi e altri.

## 3.2 Power BI

Nello sviluppo del mio progetto durante il percorso di stage ho potuto elaborare una serie di contenuti e rappresentazioni grafiche mediante l'utilizzo del software Power BI. Power BI è una raccolta di servizi di business intelligence creata da Microsoft nel 2010; è stata poi alimentata da costanti integrazioni ed aggiornamenti degli applicativi per poter connettere, elaborare e graficare con facilità i dati creando report interattivi. Il modo in cui questo applicativo viene utilizzato dipende anche dal ruolo che l'utente ha in un progetto, poichè il software è costituito da diversi elementi che interagiscono tra loro, consentendo agli utenti di sfruttare al meglio le informazioni aziendali.

Le offerte di utilizzo del software sono tre:

1. Power BI Desktop: un'applicazione desktop Windows;
2. Power BI Service: un servizio SaaS (Software as a Service) online;
3. Power BI App mobile: per dispositivi Windows, iOS e Android.

Il flusso di lavoro comune inizia connettendosi alle origini dati in Power BI desktop, il quale ha le potenzialità di fungere da dataware house ospitando informazioni da origini di dati disparate e di elaborare questi ultimi al fine di creare visualizzazioni grafiche personalizzate, presentate sottoforma di report. È quindi possibile pubblicare il report creato da Power BI desktop all'applicativo service in aree di lavoro condivise, in modo che gli utenti interessati possano contribuire alle modifiche, suggerire miglioramenti o visualizzarli per assimilare le evidenze che l'analista ha voluto comunicare. Attraverso l'applicazione mobile, inoltre, è possibile monitorare il business direttamente da uno smartphone, visualizzando le panoramiche aggiornate in tempo reale. (Learn, 2023)

Come abbiamo visto in precedenza, per le aziende moderne, adottare un approccio data-driven risulta essenziale per sfruttare lo “tsunami dei big data” per il miglioramento continuo dei processi e delle decisioni strategiche aziendali. Power BI nasce proprio per colmare il divario tra

dati e decisioni, permettendo di seguire il processo nella sua interezza, rendendolo più semplice e pratico.

### 3.2.1 Power BI desktop

L'applicativo desktop di Power BI è utilizzato dai progettisti che si occupano di cogliere dai dati le informazioni salienti e creare report interattivi per illustrare le dinamiche di interesse grazie alle varie funzionalità di business intelligence offerte. Questo servizio consente all'autore di essere accompagnato in ogni fase dell'implementazione del progetto, dal caricamento dei dati alla condivisione con gli utenti finali.

In particolare avvengono:

1. Connessione alle origini dei dati: in fase di importazione è possibile selezionare tra una vasta gamma di fonti di informazioni, tra cui database privati, servizi cloud, file locali e online.

2. Trasformazione dei dati con l'editor di Power Query: è fondamentale che i dati importati siano stati precedentemente processati e modificati ove necessario. Questo poiché lo strumento di ETL presente nel servizio permette trasformazioni a livello di colonna (per il totale dei valori della variabile trasformata) e non a livello di una singola cella.

L'editor, tuttavia, offre varie potenzialità di data cleaning sfruttando il linguaggio nativo "M" e spazi dedicati alla visione dello storico di modifica e all'organizzazione dei file.

Ogni origine di dati genera una tabella di cui si può modificare la forma: filtrando righe, rimuovendo colonne e modificando il tipo dei dati contenuti; oppure studiare la distribuzione, con processi automatici di verifica della qualità dei dati utili per assicurarsi che non ci siano anomalie prima del caricamento. Inoltre, si possono effettuare combinazioni di tabelle, da più origini di dati, attraverso comandi simili alle istruzioni JOIN in SQL. Le modifiche non influiscono sull'origine dati originale connessa, ma solo su questa vista specifica dei dati, in questo modo, è possibile aggiornare i file radice ed applicare automaticamente le modifiche salvate.

3. Data modelling: è la fase di definizione del modello relazionale presente nei dati, dove vengono definite gerarchie e relazioni tra le tabelle che presentano campi comuni, così da poterle elaborare congiuntamente attraverso grafici complessi.

4. Creazione di report interattivi: dalla finestra “Data view” è possibile utilizzare svariati tool di visualizzazioni grafiche, creazioni di colonne calcolate e misure necessarie all’analisi. Se il modello di organizzazione dei dati finora creato è la radice del progetto, il report costruito attraverso questa finestra sarà l’albero sul quale si ramificheranno le informazioni da cui cogliere i frutti della nostra analisi.

L’interfaccia è molto intuitiva poiché suggerisce una lista di oggetti visivi da utilizzare e misure statistiche rapide, inoltre, c’è la possibilità di interrogare i dati sfruttando l’AI, con l’opzione Q&A – question and answer, generando grafici in risposta alle esigenze. Allo stesso tempo, offre la possibilità di creare misure complesse mediante l’utilizzo del linguaggio DAX – Data Analysis Expression e importare grafici aggiuntivi per consentire di svolgere un’analisi completa, fedele ai principi precedentemente illustrati della data visualization. Gli utenti, dunque, possono aggiungere in ogni dashboard schede, tabelle, mappe e una varietà di grafici personalizzabili in termini di variabili coinvolte e formattazioni estetiche, per rappresentare i dati in modo efficace.

I grafici sono costituiti da punti dati, dunque si vi passa il puntatore, è possibile visualizzarne i dettagli o creare collegamenti ad altre pagine per vederne le informazioni correlate. La più grande potenzialità offerta dal software riguarda l’interattività tra gli oggetti visivi di una dashboard, ciò significa che se si seleziona uno o più valori in una visualizzazione, le altre visualizzazioni che usano lo stesso valore cambieranno in base alla selezione; come se ci fossero dei filtri applicati all’intera pagina, permettendo di focalizzarsi sul comportamento nei valori selezionati.

5. Condivisione e aggiornamento: Power BI Desktop e Service sono strettamente collegati, da qui è possibile pubblicare i contenuti nella piattaforma online, aggiornando i file di input per permettere agli utenti visualizzatori di vedere i KPI in tempo reale.

### **3.2.2 Power BI service**

Il servizio online di Power BI, nella visione interna d’impresa, è un luogo di scambio degli elaborati sviluppati per raccontare l’andamento di un determinato processo, con l’intento di comunicare alle varie funzioni aziendali le evidenze che sono state colte per poter agire in maniera costruttiva.

Qui, gli utenti, possono condividere report con altri all’interno o all’e-

sterno della propria organizzazione e collaborare con essi in tempo reale, mantenendo così, l'accuratezza e la rilevanza degli approfondimenti.

Per garantire che gli utenti vedano solo le informazioni a loro pertinenti, viene implementata la sicurezza a livello di riga, che limita l'accesso ai dati in base a ruoli e autorizzazioni utente.

Nel mondo dei big data, è vitale sfruttare tecnologie e tool che usano sistemi di data visualization per permettere all'impresa di risultare dinamica e riuscire cogliere opportunità di miglioramento.

Attraverso l'utilizzo integrato dei servizi di Power BI, i dati sono sempre a portata di mano, letteralmente, utilizzando gli applicativi mobile. Una comodità che nelle grandi imprese, come Steelco, nella quale mi sono trovato, rafforza la comunicazione permettendo di svolgere lavori in team di sviluppo di progetti altrimenti irrealizzabili. La comunicazione delle informazioni viaggia in tempo reale tra chiunque sia parte del progetto condiviso, come se l'utente artefice dell'elaborato stesse giocando una partita e i colleghi del team fossero in diretta streaming.

Queste sono alcune delle implicazioni per la quale l'utilizzo di questa piattaforma di business intelligence si è rilevata utile: analisi delle vendite (progetto dell'elaborato), gestione dell'inventario, monitoraggio delle prestazioni dei macchinari, analisi della clientela e delle promozioni.

# Capitolo 4

## Data Analysis

### 4.1 Conoscenza dei dati

Più di ieri e meno di domani le informazioni si moltiplicano, fornendo una sorgente di conoscenza inestimabile e allo stesso tempo molto complessa. Il termine “Big Data” identifica l’universo di dati generato quotidianamente dalle attività del panorama aziendale moderno, il quale, non si riferisce solo al volume in continua crescita, ma anche alla varietà e alla velocità con cui vengono generati.

Il modello “3V”, teorizzato in origine da Doug Laney e poi esteso in “5V”, è stato definito per comprendere tutti i fattori di crescita che influenzano la complessità dei dati, riassumibili secondo 5 variabili:

- Volume: riferito all’aumentare esponenziale della mole di dati generati.
- Varietà: in termini di sorgenti eterogenee di provenienza delle informazioni.
- Velocità: intesa come difficoltà nel cogliere tempestivamente le evidenze prima che perdano efficacia comunicativa.
- Veridicità: riferita alla qualità del dato da analizzare, che dev’essere affidabile.
- Valore: il fattore più importante, la caratteristica che quantifica la capacità di trasformare il dato in conoscenza utile per il business.

L’evoluzione del mondo dell’informazione, seppur ancora largamente inesplorato, cresce di pari passo alla curiosità dell’uomo di individuare tecniche capaci di sfruttarne il potenziale insito. (Wikipedia, 2024)

Una volta compresa la complessità del fenomeno e l'impossibilità di affrontarlo senza munirsi delle dovute tecnologie, l'integrazione di processi di data analysis e business intelligence rappresenta il faro guida per le organizzazioni che vogliono imbarcarsi nel mare burrascoso delle informazioni digitali.

Attraverso l'analisi dei dati è possibile individuare tendenze e relazioni, scavando nei dati grezzi per permettere alle organizzazioni di prendere "decisioni informate", un termine potente in quest'epoca, che illustra come l'applicazione di una strategia data-driven non sia più un'opzione, ma la bussola che le guida nel mare dei big data.

L'estrazione di conoscenza dai dati è articolata di vari sistemi di analisi, tradizionalmente, nella Data Science – scienza dei dati, il processo in questione è chiamato Data Mining, dall'inglese "to mine", che per l'appunto significa estrarre.

Esso è spesso parte integrante del processo più ampio di business intelligence. Mentre il data mining si concentra sull'analisi statistica per la scoperta di pattern nei dati, la business intelligence incorpora la raccolta, la presentazione e la distribuzione di informazioni aziendali mirata a fornire una visione più chiara possibile.

Il data mining si rivela prezioso soprattutto quando si parla di grandi quantità di dati, che nascondono, di conseguenza, più intuizioni sepolte al loro interno e difficilmente estraibili. Lo fa sfruttando tecniche avanzate di analisi statistica, machine learning e intelligenza artificiale, per identificare pattern nascosti, trend, correlazioni o, in generale, regole nei dati. (SAP, 2023)

Questa sfida diventa tanto ardua poiché le realtà moderne si trovano a fronteggiare informazioni provenienti da fonti tra le più disparate, risultando difficile cogliere le relazioni tra variabili appartenenti ad un insieme che è indefinito.

Pensiamo ad esempio al processo di sviluppo e migioria di un prodotto. Clienti e dispositivi con cui interagiscono le aziende ogni giorno offrono enormi quantità di informazioni utili come:

- Transazioni commerciali, anche geo-localizzate, che aiutano a definire le preferenze di acquisto dei clienti e il loro grado di fedeltà.
- Feedback sui prodotti che indirizzano i progettisti nell'ideazione di migiorie, recepiti attraverso recensioni, grado di coinvolgimen-

to delle campagne di marketing e pubblicità, comunicazioni mail e non.

- Qualità del prodotto ed efficienza della sua produzione, monitorati analizzando dati di riparazione, log generati dai server e dalle applicazioni aziendali, dati da sensori e dispositivi connessi alla catena di montaggio.

Allo stesso tempo è possibile optare per modelli di acquisto basati su variabili esterne come possono essere dati economici, dati demografici o tendenze rilevate attraverso l'utilizzo di social media e altri comportamenti online.

Da questo esempio è intuibile come ogni processo sia influenzato da una moltitudine di fattori e quanto sia importante spingersi oltre la semplice analisi descrittiva, per rispondere non solamente alla domanda “cosa è successo”, ma indagare sul “perché è successo”.

L'obiettivo del data mining è proprio quello di “pensare fuori dagli schemi” per colmare questa lacuna, sfruttando appieno i big data per consentire alle aziende di avere una visione completa dei loro clienti, dei loro processi produttivi e dell'ambiente circostante, permettendogli di intraprendere decisioni informate.

## 4.2 CRISP-DM

Esistono vari tipi di approccio all'estrazione di conoscenza dai dati, in base alle esigenze, che possono includere attività di classificazione, regressione, clustering, associazione o analisi delle anomalie.

Ciò che comunemente è utilizzato, per guidare i data miner nelle fasi di estrazione della conoscenza, è un modello standard che scandisce gli step necessari per una corretta analisi: il CRISP-DM – Cross Industry Standard Process for Data Mining. (Wirth e Hipp, 2000)

La rappresentazione grafica del ciclo è la seguente e sintetizza il processo in 6 fasi principali.

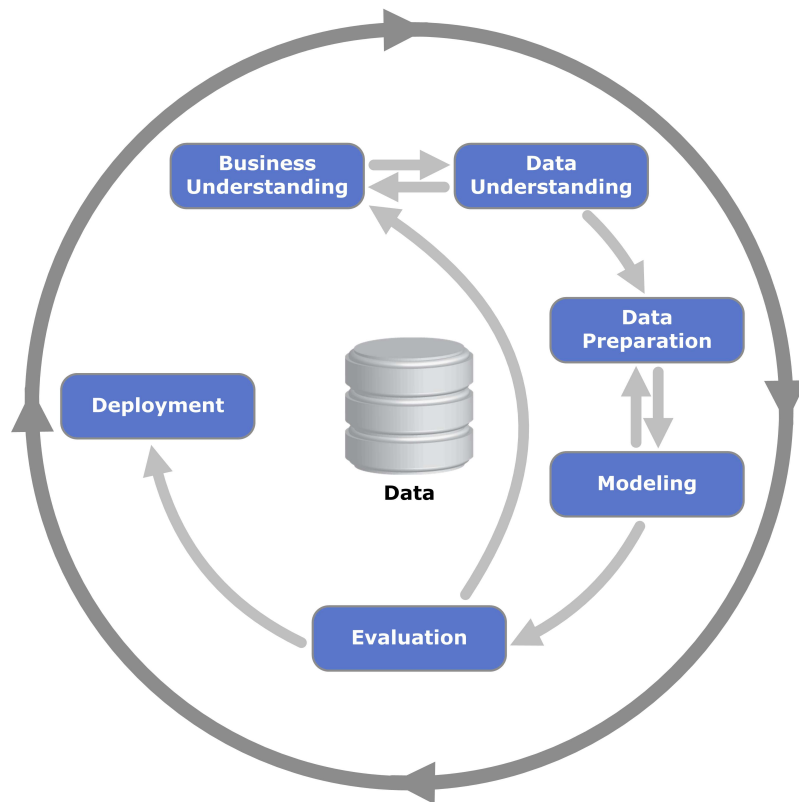


Figura 4.1: Diagramma del ciclo CRISP-DM.

1. Business Understanding: È il progetto del progetto. Il punto di partenza necessario per identificare gli obiettivi di ogni analisi, basati sul problema che l'azienda vuole mitigare. Nello schema del progetto di data mining è fondamentale capire quali stakeholder sono coinvolti e qual è la situazione aziendale in termini di risorse, in modo da definire costi e benefici potenziali.

2. Data Understanding: Questa è la fase più importante per far sì che il progetto di analisi sia condotto in maniera efficiente. Senza comprendere l'origine dei dati, il processo di creazione e le misurazioni, è impossibile ottenere una chiara inquadratura del problema da risolvere e di conseguenza i mezzi per venirne a capo. Durante questa fase, gli analisti esplorano le sorgenti dei dati raccolti e svolgono diverse attività per familiarizzare con questi, al fine di valutarne la qualità e di selezionare le caratteristiche chiave che saranno oggetto di studio. Inizialmente i dati vengono raccolti e descritti, in modo da sapere il significato che si



trova dietro ad ogni "numero" registrato, in seguito vengono esplorati per identificarne eventuali sfide o limiti. Vengono analizzate la presenza di dati mancanti, possibili valori anomali o errati, la distribuzione delle variabili e le relazioni tra loro esistenti.

3. **Data Preparation:** In questo step vengono create le fondamenta del progetto, coinvolge tutte le attività che portano alla costruzione del dataset che sarà oggetto di modellazione, partendo dai dati grezzi. Il risultato di questa fase è determinante per la qualità dell'analisi, per questo vengono ripetute più e più volte tecniche di manipolazione e trasformazione dei dati, al fine di ottenere una base solida e affidabile per lo sviluppo del progetto. Qui avviene la selezione delle variabili da considerare, sfruttando raggruppamenti e suddivisioni per ridurre la complessità del modello e variabili di sostegno, come le dummy, create per rappresentare categorie specifiche. Il tutto ricorrendo ricorsivamente alle attività di data cleaning – pulizia dei dati, che includono la gestione dei dati mancanti, errati o anomali, la standardizzazione delle variabili e altre trasformazioni capaci di rendere il set di dati finale robusto e coerente.

4. **Modeling:** La fase di modeling è il succo del progetto, ciò da cui effettivamente verrà estratta la conoscenza. Inizialmente, avviene la selezione delle tecniche di modellazione più adatte alla natura del problema e agli obiettivi prestabiliti, ricordando però, che la soluzione non è unica. È necessario valutare a posteriori l'output generato dai vari modelli implementati, così da verificare attraverso test di prestazione, se risponde correttamente alle domande poste e se lo fa in maniera efficiente, ricorrendo ad algoritmi di ottimizzazione, come la regolazione degli iperparametri e la validazione incrociata. È compito del data miner capire quale modello si adatta meglio ai dati e agli obiettivi, a maggior ragione poiché egli ha seguito il processo di understanding e preparation, che si mostrano essere i requisiti primari. Infatti, questa fase è strettamente collegata alle precedenti e solitamente spinge l'analista a rivedere le scelte prese in fase di preparazione del set di dati e ad apportare modifiche. In generale, è fondamentale rivedere i vari step durante tutto il ciclo esecutivo del processo, che è un iterative flow, non lineare, ovvero assomiglia più alla pianificazione di una vacanza che all'assemblaggio di un mobile.

5. **Evaluation:** Dopo questa fase di valutazione, saremo giunti alla fine

dell'analisi, il modello di data mining sarà definito e le domande avranno una risposta specifica. È importante, allora, capire se il risultato che il modello genera è utile o meno, approfondendo ogni passaggio attuato nelle fasi precedenti e valutando l'output mediante l'utilizzo di test set di dati per confrontare i risultati attesi con i reali.

6. Deployment: Una volta svolta l'analisi completa, attraverso un modello ottimizzato che risponde agli obiettivi aziendali, è il momento di chiedersi: “come può essere integrato?”. L'ambiente operativo nel quale si troverà è popolato di persone che devono essere in grado di cogliere gli effetti benefici del modello, perché ciò accada è necessario pianificare accuratamente la sua distribuzione, documentare minuziosamente e formare gli addetti. Il modello verrà quindi presentato, implementato secondo il deployment plan, inserito all'interno dei processi aziendali coinvolti consentendo l'estrazione di conoscenza, monitorato il suo funzionamento e valutati i risultati ottenuti per garantire il valore aggiunto.

### 4.3 Analisi esplorativa

Come abbiamo visto, il data mining è un processo articolato che utilizza tecniche statistiche per estrarre informazioni complesse, nascoste nei dati. Una volta costruito il progetto di analisi, dalle conclusioni tratte mediante l'utilizzo dei modelli costruiti risulteranno piani di efficientamento cuciti ad hoc sullo scheletro dell'impresa; ma non è così semplice, per un addetto all'estrazione di conoscenza, acquisire una familiarità con le dinamiche d'impresa tale da saper creare modelli analitici su misura. Dobbiamo fare un passo indietro.

Quando si vuole avviare un processo di data mining, soprattutto in una grande società, è necessario avere un'immagine limpida del proprio funzionamento: quali sono le funzioni aziendali coinvolte nei processi? quali informazioni sono disponibili? quali problemi si stanno verificando e quali opportunità sono incolte?

Per rispondere a queste domande sono essenziali la comunicazione e l'esplorazione dei dati.

L'esplorazione fornisce una panoramica dettagliata sullo stato attuale dei processi e dei dati derivati per le diverse aree di business, ognuna di esse però, gestisce dati provenienti da diverse fonti e conosce legami che senza la comunicazione resterebbero nell'ombra.

L'analisi esplorativa diventa quindi uno strumento cruciale per incoraggiare le varie unità aziendali a condividere i loro problemi, il loro stato attuale e le prospettive future attraverso report dettagliati. Da questo scambio di informazioni vengono comprese le esigenze specifiche di ciascuna funzione, consentendo una progettazione mirata dello schema di lavoro del data analyst e garantendo che gli obiettivi siano allineati con le reali esigenze dell'azienda.

L'EDA si differenzia dalla statistica inferenziale in quanto l'obiettivo non è quello di generalizzare delle conclusioni a popolazioni più ampie, bensì di riepilogare le informazioni nascoste nel set di dati oggetto di studio per descriverlo. Lo fa utilizzando misurazioni di sintesi dei fenomeni come medie o varianze e vuole evidenziare le caratteristiche salienti attraverso l'uso di rappresentazioni grafiche conformi alle regole suggerite dalla data visualization.

I dati fanno parte di un set osservato, reale e completo, non vengono dunque utilizzati modelli stocastici che potrebbero averli generati o strumenti come test di ipotesi e intervalli basati su un grado di incertezza. Inoltre, la modalità e lo schema di rilevazione dei dati suggerisce che non siano "perfetti", al contrario potrebbero essere inaccurati, osservati in modo incompleto o con errori, pertanto, verrà svolta una fase consistente di pulizia dei dati preliminare all'analisi esplorativa. (Cicchitelli *et al.*, 2017)

Dopo la fase di data cleaning, si può supporre che le informazioni colte siano attendibili, dunque, vengono generate visualizzazioni in grado di riassumere l'andamento dei fenomeni misurati seguendo i KPI guida sulla quale l'azienda fa affidamento in base ai benchmark scelti.

Nella fase di esplorazione, l'obiettivo generale è cogliere le seguenti caratteristiche dai dati, in modo da poterle comunicare con consapevolezza:

- Identificazione di outliers: dati anomali che si discostano significativamente dalla norma e possono influenzare negativamente la precisione di modelli statistici futuri.
- Pattern e tendenze: Attraverso l'analisi di dati correlati temporaneamente è possibile definire andamenti ricorrenti, trend, stagionalità e ciclicità.
- Comportamento delle variabili: Analizzate singolarmente per capirne la distribuzione e misurarla, se qualitativa, oppure osservandone

la frequenza di ogni categoria, se qualitativa. Inoltre, vengono condotti studi sulle relazioni tra coppie di variabili per evidenziarne correlazioni o dipendenze.

- Piste per analisi future: avvengono le verifiche di ipotesi iniziali di interesse e sono individuati suggerimenti sui quali basare analisi più approfondite o modelli statistici che potrebbero essere appropriati.

### 4.3.1 Time series

Attraverso le analisi temporali dei dati si scorgono informazioni utili sull'evoluzione del comportamento di una variabile, come può essere la vendita di un determinato prodotto, nel tempo. Le quali suggeriscono spunti interessanti riguardo analisi mirate, da condurre sfruttando tutte le potenzialità di data analysis disponibili, per capire, ad esempio, il motivo della crescita delle vendite relative ad un prodotto piuttosto di un altro nello storico aziendale.

Poiché nulla è un punto di riferimento maggiore del tempo, l'analisi delle serie storiche è il metodo più efficace per determinare quanto un processo aziendale sia stabile ed è essenziale per sollevare questioni riguardanti problemi e opportunità identificate.

Ogni processo generatore dei dati, misurato secondo osservazioni sequenziali nel tempo di una variabile numerica, è composto di due fattori che ne determinano il risultato: una componente deterministica e una componente stocastica.

Per componente deterministica intendiamo tutto ciò che segue uno schema prevedibile, dunque che possa essere modellata; mentre la componente stocastica viene anche definita "rumore" in quanto rappresenta la parte imprevedibile del fenomeno, dovuto a fluttuamenti casuali o errori di misurazione. Non essendo questi fattori osservabili, il compito dell'analisi time series è riuscire a misurarli, in modo da poter descrivere il fenomeno attraverso un modello ed attuare previsioni basate sui valori osservati.

Un evento temporale  $Y$  viene definito nel seguente modo:

$$Y_t = f(t) + U(t) \quad (4.1)$$

Dove:

$f(t)$  è la componente sistematica o deterministica

$U(t)$  è la componente stocastica o casuale

In particolare, la parte sistematica può essere decomposta in 3 fattori che concettualmente modificano l'evoluzione temporale della serie, essi sono:

1. Trend: la tendenza di fondo nel lungo periodo.
2. Ciclo: la componente legata all'alternarsi di fasi di espansione e recessione del sistema in un periodo di durata media.
3. Stagionalità: la componente periodica che si esaurisce nell'arco di un anno, dovuta ad oscillazioni con frequenza periodica ( $s$ ) e ampiezza costante.

Le medie mobili sono uno strumento statistico utile per evidenziare la componente di trend ed attenuare le componenti periodiche di periodo medio-breve e le componenti di rumore presenti nei dati.

Si tratta di un approccio non parametrico per il "lisciamento delle serie storiche": avviene una trasformazione della serie di dati tramite l'utilizzo di una funzione di calcolo specifica (media mobile) che filtrerà la componente stagionale al fine di stimare il ciclo-trend.

Vengono generalmente utilizzati due tipi di medie mobili: la media mobile semplice o SMA calcola la media aritmetica dei valori in un determinato periodo di tempo, la media mobile esponenziale o EMA assegna pesi crescenti ai dati in base alla loro età, i dati più recenti hanno un peso maggiore rispetto ai più vecchi.

L'ordine di una media mobile, ovvero il numero di punti di dati storici coinvolti nel calcolo, è responsabile della capacità di questa di evidenziare variazioni nei dati, più sarà elevato, più sarà preferibile per catturare tendenze di lungo termine, al contrario coglierà cambiamenti immediati. Efficace è l'utilizzo di medie mobili aritmetiche semplici, come vedremo poi nel progetto, per eliminare la stagionalità, quando stiamo trattando periodicità dovute a dati trimestrali o mensili.

Questo poiché una SMA è capace di eliminare componenti stagionali di periodo pari al proprio ordine e conservare trend lineari, inoltre attenua la componente erratica secondo un rapporto di riduzione della varianza pari all'inverso dell'ordine.

La trasformazione di un dato temporale ottenuta mediante SMA, centrata al tempo  $t$ , di periodo  $s$ , è calcolata come media dei valori nei tempi:

$$\left[ t - \frac{s-1}{2}; t + \frac{s-1}{2} \right] \quad (4.2)$$

Ad esempio, considerando dati quadrimestrali, si ottiene dal calcolo:

$$Y_t^* = \frac{1}{3} \times (Y_{t-1} + Y_t + Y_{t+1}) \quad (4.3)$$

Che in generale diventa:

$$Y_t^* = \frac{1}{s} \times \sum_{i=t-\frac{s-1}{2}}^{i=t+\frac{s-1}{2}} Y_i \quad (4.4)$$

Dove:

$Y_t^*$  è la media mobile centrata al tempo  $t$

$S$  è il periodo di riferimento

Quando il periodo di riferimento è di ordine pari (ad esempio con dati mensili,  $s = 12$ ), per ottenere gli stessi risultati sopra analizzati, è necessario effettuare un'operazione di centratura attraverso il calcolo della media di due medie mobili. (DiFonzo e Lisi, 2005)

# Capitolo 5

## Il Progetto

### 5.1 Problema e obbiettivi

All'interno di una grande azienda sono presenti grandi dinamiche funzionali, diverse, che necessitano di una formazione a 360 gradi per poter essere comprese fino in fondo. Per facilitare la possibilità di prendere decisioni informate da parte di chi si occupa di attuare le strategie di business, abbiamo visto come sia indispensabile, al giorno d'oggi, l'integrazione di tecnologie ICT, in particolar modo, di sistemi di analisi di dati e funzionalità di business intelligence.

Ad oggi, Steelco, sta vivendo un processo evolutivo completo, dopo l'acquisizione di Icos e l'avvento di Miele, ha integrato metriche di business che le hanno permesso di diventare grande, tuttavia, una casa non può diventare un grattacielo poggiando sulle stesse fondamenta.

All'interno dell'ufficio IT – Information Technology, nel quale sono stato inserito, mi è stata illustrata la roadmap di evoluzione della società, dei processi coinvolti e della complessità nel gestirli. La velocità di mutamento degli input spesso produce degli “angoli ciechi” anche per chi, come gli specialisti IT, si nutre di dati. Perciò, in questo particolare momento di transazione evolutiva, si è deciso di impiegare delle risorse nell'analisi esplorativa delle dinamiche d'azienda; per avere una base solida di supervisione degli eventi su cui fondare potenziali ipotesi da testare, poi, attraverso l'utilizzo di tecniche avanzate per la creazione di modelli di efficientamento.

È bene ricordare come, all'interno di una grande azienda, sia naturale che gli addetti all'analisi siano più creativi ed efficaci con i loro studi man mano che acquisiscono maggiore familiarità con gli strumenti, i dati e le dinamiche aziendali. Allo stesso modo, senza conoscere ogni lato della personalità di una persona, è facile definirla secondo un modello

comportamentale sbagliato, tuttavia, esplorando gli atteggiamenti che adotta nei confronti di persone diverse o situazioni particolari, è possibile identificarne i tratti distintivi, rilevare “trend comportamentali” e relazioni specifiche che li uniscono o li diversificano. Per questo motivo è importante conoscere i propri limiti e sfruttare ogni possibilità per ottenere delle analisi, magari meno complesse, ma precise.

Sulla base di questi principi, sono stato avvicinato alle funzioni di BI Analysis, nell’ottica di poter alimentare la corretta comunicazione per favorire il processo strategico decisionale. In particolare, mediante l’utilizzo di Power BI, il sistema comunicativo scelto all’interno della rete aziendale, per la condivisione di report, in modo che ogni settore possa condividere KPI relativi al proprio ambito con chi di dovere attraverso la piattaforma service di Power BI.

Una volta deciso di integrarmi come risorsa nel ruolo di Junior IT specialist, mi è stata fatta presente la problematica di perfezionamento della gestione delle scorte di magazzino, un meccanismo molto complesso, che va sezionato e studiato per parti. Avendo abbracciato le ideologie lean e, in particolare, un approccio push alla produzione, il risultato AS IS di ottimizzazione del magazzino è buono, ma può ancora essere migliorato. In un momento storico come questo, dove i prezzi dei beni nel mercato sono alle stelle e gli investimenti in innovazione e tecnologie ICT sono essenziali, è fondamentale riuscire a sfruttare al meglio le disponibilità economiche aziendali. Trattandosi di una società altamente coinvolta nella produzione specializzata, lavora a contatto con istituti sanitari, universitari e di ricerca, per i quali offre soluzioni elaborate ad una gamma di prezzi elevati. È indispensabile, perciò, mantenere un livello di scorte ottimizzato, in modo da poter disporre di una maggiore disponibilità finanziaria per gestire l’insieme dei processi.

Due dei sistemi che si è deciso di adottare per affrontare questa difficoltà sono il MES – Manufacturing Execution System e il WMS – Warehouse Management System, dei software di gestione del processo di produzione e del magazzino. Il reparto IT, tra i vari compiti, sta svolgendo attività di integrazione delle funzionalità verso questi nuovi sistemi, occupandosi della migrazione di dati e funzioni, mantenendo allineate le divisioni aziendali e garantendo lo svolgimento delle attività presenti. Marginalmente, ho contribuito all’implementazione di test e simulazioni di go-live per la sincronizzazione dei vecchi sistemi con i nuovi; ciò mi ha permesso di familiarizzare con il sistema gestionale AS400, dal quale ho recuperato tutte le informazioni necessarie per le mie analisi.

Vivendo un periodo di passaggio delle strutture di gestione dei meccani-



smi coinvolti nell'ottimizzazione del magazzino, si è deciso di incentrare le mie ricerche sui dati di vendita, una dinamica altamente correlata, tenendo a mente che: “per sapere quanto produrre, devo sapere quanto vendo”.

L'obbiettivo prefissato è dunque quello di esplorare il mondo delle vendite, riassumendo gli accorgimenti di maggior interesse secondo report volti a descrivere i concetti chiave, per evidenziare ricorrenze, anomalie o relazioni. Il tutto, sfruttando gli applicativi di Power BI, per generare visualizzazioni personalizzabili e interattive, che consentono di scavare più a fondo nei dati in modo da offrire spunti per future analisi più complesse; ed alimentare la condivisione dei risultati nelle piattaforme di comunicazione che prenderanno presto piede in azienda.

## 5.2 Dal gestionale al software

Una volta definito il problema presente e il contributo possibile da poter dare per risolverlo, ho intrapreso un periodo di “familiarizzazione”, per capire come portare a termine gli obbiettivi prefissati.

All'arrivo in azienda, mi è stato dedicato del tempo per poter familiarizzare con il funzionamento dell'impresa, consapevole del fatto che per poter sviluppare un lavoro di analisi coerente e pensato, è doveroso andare in profondità nelle sorgenti dei dati in modo da riuscire ad interpretarli.

Per alimentare la mia cultura sui meccanismi d'azienda, ho avuto dei meeting con i colleghi dei reparti Business Process e KPO, che mi hanno aiutato a comprendere la value chain di Steelco, essenziale per definire le opportunità di miglioramento, in modo da massimizzare il valore generato. Il reparto IT, invece, è il covo delle informazioni, con cui ho potuto prendere confidenza mediante il sistema gestionale AS400. Questo applicativo viene utilizzato nel 90% dei processi d'impresa: dalla produzione alla vendita, dal marketing alla gestione del magazzino.

All'interno dell'ufficio, sono stato reso partecipe dei vari meccanismi di manutenzione del suo funzionamento. Alla fine dei conti, concettualmente, è paragonabile alla troupe che permette lo svolgimento di uno show, facendo sì che tutto proceda senza intoppi, cercando di prevenirli ed eventualmente costruendo soluzioni veloci per aggiustare errori. Qui, ho appreso il significato dei dati: ogni procedimento svolto nelle diverse aeree viene tracciato e converge nel sistema, attraverso interfacce

specialistiche è possibile accedere alle categorie di interesse e tutte le variabili in esse coinvolte, le cui informazioni vengono codificate attraverso tabelle di rappresentazione dei campi.

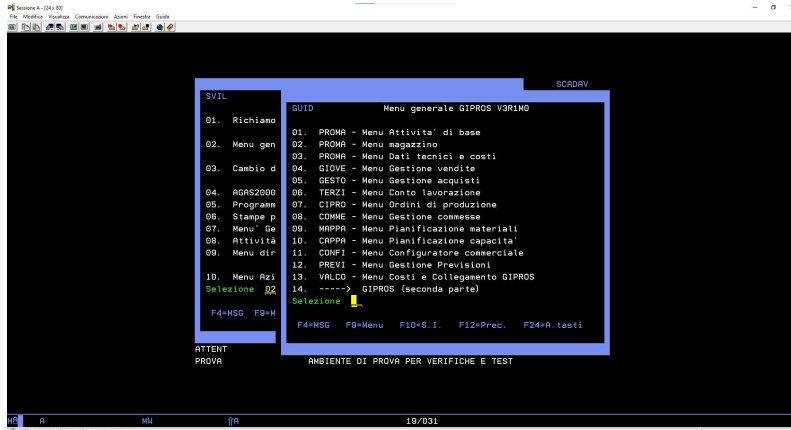


Figura 5.1: Gestionale AS400, interfaccia menù

Da qui, ho ereditato tutte le informazioni coinvolte nel progetto, estraendole mediante query di interrogazione nel linguaggio SQL. All'interno di una cartella dedicata, interna al datacenter aziendale, sono stati estratti ed inseriti i dati grezzi, sotto forma di file tabellati Excel suddivisi per gruppi di informazioni, e alcuni file di codifica delle variabili, per comprendere il significato dei valori interni ai campi. Di seguito un esempio per le righe di fatturazione a sistema, tabellate e spiegate.

J	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	?
1	RSNRAZ	RSANNU	RSRIAL	RSRIFE	RSRFNR	RSRIGA	RSORDI	RSORDO	RSORRI	RSCAFI	RSMAGC	RSOAR	RSUMIS	RSVA	RSQZC	
2	5645100	21	017479		490		21002695	910 02	VA	9991979	NR	52				RSNRAZ RIGBFOOF Numero azione
3	5645100	21	017479		500		21002695	920 02	VA	9991549	NR	52				RSANNU RIGBFOOF Annullamento record
4	5645100	21	017479		510		21002695	925 02	VA	9991817	NR	52				RSRIAL RIGBFOOF Rif. Utente 2
5	5645100	21	017479		520		21002695	930 02	VA	9991887	NR	52				RSRIFE RIGBFOOF Nr.Progr.Rifer.2
6	5645100	21	017479		530		21002695	940 02	VA	9991972	NR	52				RSRFNR RIGBFOOF Nr. Riferim.2
7	5645100	21	017479		540		21002695	950 02	VA	9991891	NR	52				RSRIGA RIGBFOOF N.rifer.ordine
8	5645100	21	017479		550		21002695	960 02	VA	9991902	NR	52				RSORDI RIGBFOOF Riferim. Ordine
9	5645100	21	017479		560		21002695	965 02	VA	9991949	NR	52				RSRORD RIGBFOOF Rifer. Ordine 2
10	5645100	21	017479		570		21002695	971 02	VA	9991252	NR	52				RSORRI RIGBFOOF Riga Rifer.ordine
11	5645100	21	017479		580		21002695	970 02	VA	9991178	NR	52				RSCAFI RIGBFOOF Cani Standard Contrab.
12	5645100	21	017478		20		99999999	0 TR		04050000	52					RSMAGC RIGBFOOF Codice magazzino
13	5611987	21	017480		10		21017507	10 R4	01	012796	NR	52				RSOAR RIGBFOOF Codice articolo
14	5645100	21	017481		10		21016223	10 R2	01	5992087	NR	52				RSUMIS RIGBFOOF Unità di misura
15	5645100	21	017481		20		21016223	20 TR		04050000	52					RSVA RIGBFOOF Abiquota I.V.A.
16	5611987	21	017482		10		21017319	10 R4	01	509053	NR	52				RSQZC RIGBFOOF Q.ta 2 ordinata
17	5611987	21	017482		20		21017319	20 R4	01	010109	NR	52				RSQZTS RIGBFOOF Q.ta 2 spedia
18	5611987	21	017482		30		21017319	30 R4	01	020007	ML	52				RSFCON RIGBFOOF Fattore convers.
19	5701737	21	017483		10		21017043	10 R1	01	505083	NR	52				RSQZMA RIGBFOOF Un.mis.acquisto
20	5645100	21	017484		10		21017325	10 R2	01	010979	NR	52				RSQTOR RIGBFOOF Q.ta 1 ordinata
21	5645100	21	017484		20		21017325	20 R2	01	010980	NR	52				RSQTSR RIGBFOOF Q.ta 1 spedia
22	5645100	21	017484		30		21017325	30 R2	01	010209	NR	52				RSQTSB RIGBFOOF Q.ta calc. saldi
23	5645100	21	017484		40		99999999	0 TR		04050000	52					RSQTSV RIGBFOOF Sconto articolo
24	5611987	21	017485		10		21017270	10 R4	01	661911	NR	52				RSSCAR RIGBFOOF Sconto cliente
25	5611990	21	017486		10		21017343	10 C6	01	054903	NR	52				RSSCMO RIGBFOOF Sconto Mod.Pagament
26	9346649 A	21	017487		10		99999999	0 S9	40	013728	NR	52				RSSCTB RIGBFOOF Sconto tabellare
27	5645100	21	017488		10		21017354	10 R2	01	010941	NR	52				RSPEPR RIGBFOOF Perc. Provvigioni
28	5645100	21	017488		20		99999999	0 TR		04050000	52					RSRPER RIGBFOOF Provv. campagna
29	5751671	21	017489		10		21017295	10 R1	01	PL001032	NR	52				RSDEAR RIGBFOOF Descrizione articolo
30	5751671	21	017489		20		21017295	20 R1	01	PL001032	NR	52				RSDEAS RIGBFOOF Descr. supplm.
31	5751671	21	017489		30		21017295	30 R1	01	PL001032	NR	52				RSRDCR RIGBFOOF Data cons.conf
32	5751671	21	017489		40		21017295	40 R1	01	PL001032	NR	52				RSDDCO RIGBFOOF -Confemata
33	5751671	21	017489		50		21017295	50 R1	01	PL001032	NR	52				RSNUBO RIGBFOOF Numero bolla
34	5751671	21	017489		60		21017295	60 R1	01	PL001032	NR	52				RSDABO RIGBFOOF Data bolla amg
35	5751671	21	017489		70		21017295	70 R1	01	PL001032	NR	52				RSNLFA RIGBFOOF N.fattura
																RSNFA RIGBFOOF Data fattura

Figura 5.2: A sinistra file Excel tabellato con le informazioni di riga, a destra file note con le codifiche delle variabili

Al fine di produrre delle comunicazioni efficaci, inoltre, sono stato inserito nel progetto condiviso sull'applicativo service di Power BI del dipartimento commerciale di Steelco. Potendo entrare solamente in visualizzazione, grazie alla sicurezza di riga, ho captato le informazioni che si vuole comunicare più che il “come farlo”, una risposta a cui sono giunto solamente attraverso una fase consistente di Data understanding direttamente dalla fonte delle informazioni. Mettendomi in contatto con loro, ho potuto usufruire di tabelle Excel esterne al gestionale, “scritte a mano”, per la classificazione dei prodotti secondo area di mercato, macro famiglia di prodotto e modello di macchina, costruite sfruttando i codici di codifica associati ad ogni singolo prodotto venduto. Nel momento in cui ho avuto l'accesso alle informazioni a sistema, capito le dinamiche di vendita e produzione presenti internamente e contattato i commerciali, ovvero i “clienti finali” del mio progetto, sono entrato a capofitto sui dati, per riuscire a tradurre ogni numero in un evento. Dalla miniera di informazioni accessibili, inizialmente, ho deciso di trarne alcune di fondamentali per esplorare le dinamiche di vendita, riassunte secondo le seguenti tabelle:

- Righe di fattura
- Testate di fattura
- Anagrafiche prodotti
- Anagrafiche clienti
- Classificazioni commerciali

Il processo di analisi segue un flow iterativo, il che prevede che spesso si ritornerà sui propri passi per valutare strade alternative, solitamente, variabili escluse che possono rivelarsi utili o informazioni aggiuntive che rendono il programma di analisi meno rapido e verranno accantonate.

Nella fase di familiarizzazione con i dati, uno step utile per valutarne la qualità è analizzare i valori assunti dai campi, poiché se compresi, faranno sorgere dei dubbi sulla veridicità delle informazioni e scateneranno attività di pulizia per renderli coerenti. In primis, sono sorti degli accorgimenti su tre importanti casi che generano la mancanza delle informazioni di vendita in riga di fattura, la tabella riferimento principale per l'analisi, alla quale si collegheranno tutte le altre. Nello specifico quando:

- La variabile “Data Fattura” è valorizzata a Zero.

- La variabile “Riga da non fatturare” è valorizzata a 1.
- La variabile “Stato” è valorizzata “A” (Annullato).

Perciò, si è deciso di filtrare i dati escludendo le righe con le sopra citate caratteristiche, per avere un dataset pieno; successivamente, sono stati modificati alcuni errori di allineamento dei sistemi che hanno generato mancanza di informazioni o anomalie nei dati. Tra questi: tassi di cambio e classi di sconto errate, clienti senza codice associato, ecc.

Una volta compreso il significato dei dati e selezionate le fonti di maggior interesse per raggiungere gli obiettivi, vengono costruite le fondamenta per l’analisi: il dataset di riferimento. Nei campi ICT, è gergo comune utilizzare il termine “GIGO – gargbage in, garbage out” quando ci si riferisce ai problemi legati a dati input, il termine deriva proprio dai tempi dei primi programmi computer, dove, se viene fornito un insieme di dati in entrata palesemente erronei o insensati (garbage in) verrà prodotto un risultato erroneo o insensato (garbage out). (Wikipedia, 2023b)

Per questo motivo, la fase di preparazione dei dati, spesso seguita con superficialità, è stata condotta nel dettaglio, cercando di produrre delle analisi coerenti, che facciano riferimento su una base solida.

Lo strumento di ETL utilizzato dal software, Power Query, ha contribuito al corretto svolgimento di questa fase grazie alle sue funzionalità. Inizialmente, è stata creata una connessione alle sorgenti di dati sottoforma di file Excel, questi sono stati poi importati all’interno di una staging area dove è possibile attuare le operazioni di modifica, strutturata nel seguente modo:

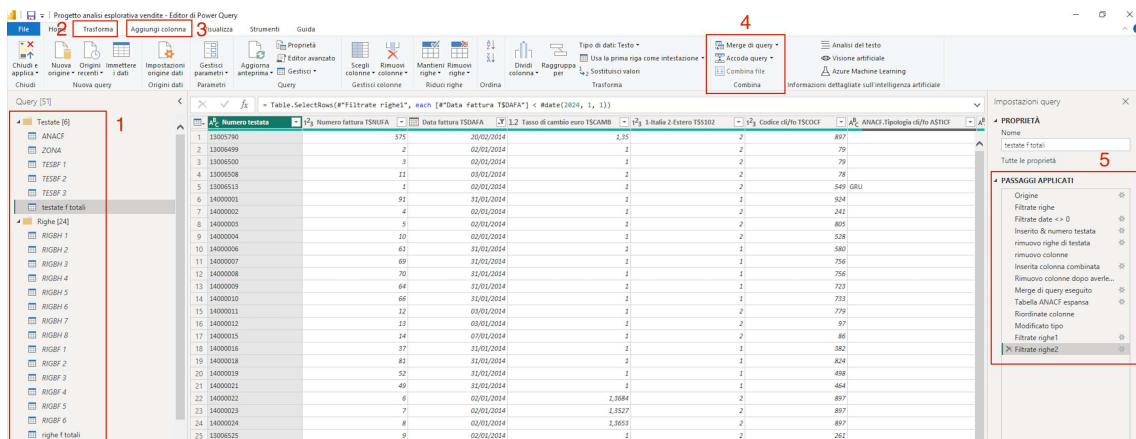


Figura 5.3: Interfaccia dell’editor Power Query

Nella Sezione 1 vediamo le tabelle importate, nel nostro caso quelle illustrate precedentemente. Ogni foglio di un file Excel contiene un numero limitato di righe, per questo motivo verranno importate in più tabelle, che saranno unite per ottimizzare gli spazi.

Dalla Sezione 2 è stato possibile attuare le trasformazioni, tra cui: verifica del tipo di dato, rinomina delle variabili, rimozione di variabili superflue, verifica del completamento dei campi, applicazione di filtri sistematici di riga.

Mediante Sezione 3 sono state create colonne di supporto, come l'unione delle informazioni di spedizione sotto un'unica variabile indirizzo, la distinzione per continenti o la variabile numero testata, prima codificata da un anno e un progressivo, poi combinati.

Nella Sezione 4 sono stati sviluppati comandi di join, in linguaggio M, per due tipi di operazioni: L'accodamento, quando i record di due tabelle vengono appunto accodati sotto la stessa famiglia di variabili in una unica, svolto ad esempio con le varie tabelle di riga. Il merge, prevede che due o più tabelle composte di variabili differenti (ma una chiave) vengano unite in una unica. Ad esempio, le righe di fattura vengono combinate con le anagrafiche clienti e prodotti.

Infine, la sezione 5 permette di ripercorrere le elaborazioni svolte.

Una volta completata la fase di pulizia, i dati vengono caricati all'interno del software, nel quale è possibile definire e modificare il modello relazionale creato, visualizzando i legami tra le tabelle. Nel nostro caso:

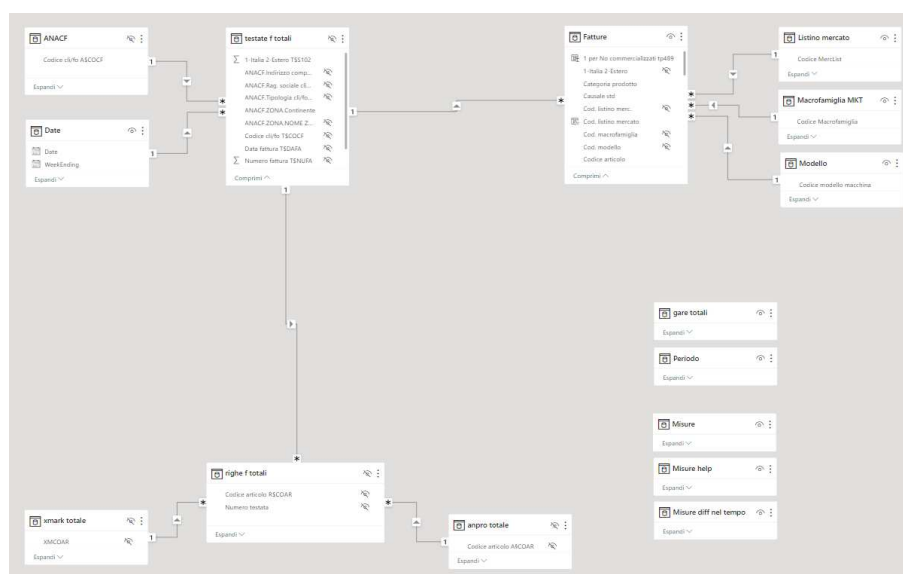


Figura 5.4: Modello relazionale delle tabelle di dati caricate.

## 5.3 Steelco nel mondo

Il dataset impostato pone le sue basi su un indicatore di vendita estremamente importante e di interesse; chiunque voglia conoscere l'andamento dell'azienda in termini di vendite, vuole sentire parlare di fatturato. È proprio dalle fatture di vendita che nasceranno le analisi del progetto, da ognuna di esse ho potuto trarre informazioni, secondo il modello relazionale creato, sul cliente, sul prodotto venduto, sulla quantità, sul prezzo, sugli sconti, ecc.

La prima panoramica che ho deciso di creare vuole rappresentare l'espansione globale che Steelco ha avuto negli ultimi 10 anni, dopo l'acquisizione di Icos e la partnership con Miele. Il dataset considerato, riferisce proprio alle fatture emesse nell'ultima decade, tuttavia, per considerare finestre di tempo più ridotte per analisi mirate, è sufficiente utilizzare un semplice filtro temporale. Ne viene illustrato un esempio dove, in particolare, vengono illustrati gli indicatori di fatturato e quantità dell'anno in corso e il loro incremento percentuale rispetto all'anno precedente.



Figura 5.5: Filtro per modificare la finestra temporale e indicatori di fatturato e quantità, espressi in valore assoluto e come incremento percentuale rispetto all'anno precedente.

Lo strumento più efficace a livello visivo, per farsi un'idea iniziale sulle aree maggiormente influenti per il business, è una mappa, più densamente colorata per le aree a maggior valore commerciale.

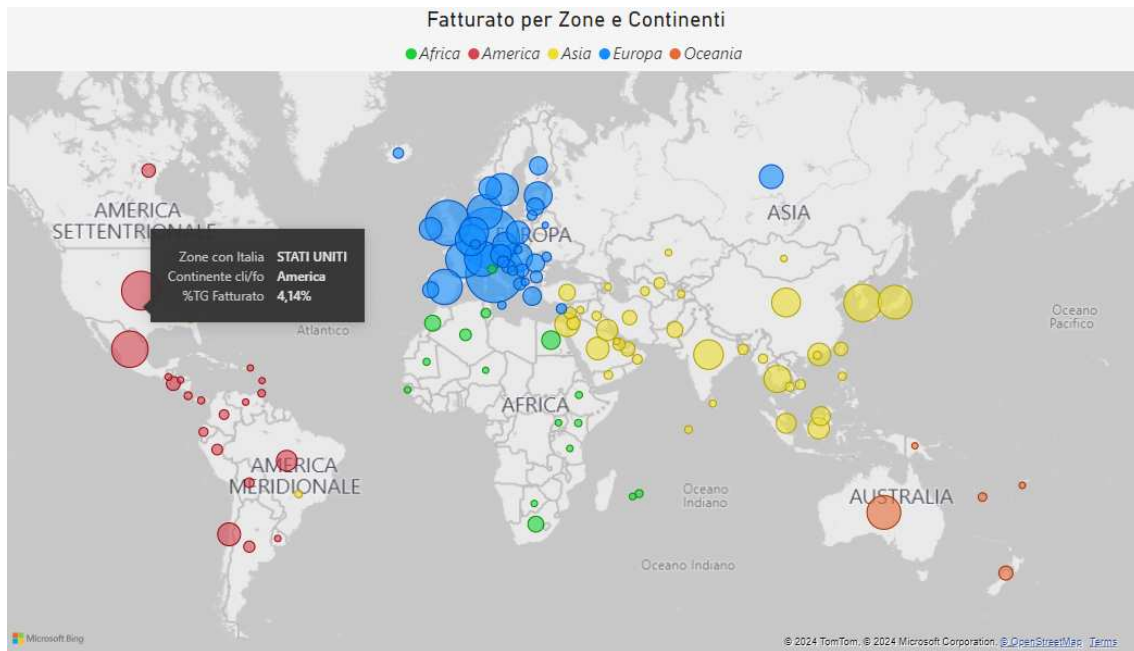


Figura 5.6: Proportional symbol map del fatturato nel mondo espresso in valori percentuali, rappresentato da cerchi di grandezza proporzionale al contributo dello stato sul fatturato totale, colorati diversamente per continente.

Salta subito all'occhio la preponderante attività commerciale presente in Europa, responsabile del 63% del fatturato totale, seguono Asia (22%) e America (10%), Africa e Oceania insieme conferiscono solo il 5% degli introiti. Uno scenario prevedibile, vista la presenza di 10 sedi commerciali su 14 in Europa e le restanti 4 distribuite tra Cina, Emirati, USA e Messico. Per entrare nel dettaglio, riguardo il fatturato derivante da ogni paese, si è deciso di generare una tabella di profondità, sotto forma di matrice, nella quale è possibile indagare prima a livello di continente, poi di stato, fino ad arrivare al cliente finale. In aggiunta, è stato indotto un approfondimento sulla fedeltà dei clienti coinvolti nei rapporti commerciali, attraverso la creazione di una tabella di supporto e di misure calcolate attraverso il linguaggio DAX. Questa variabile, chiamata *loyalty – fedeltà*, prende in pasto un periodo di riferimento selezionato dall'utente (nell'esempio 24 mesi) e viene definito il comportamento dei clienti secondo il loro grado di fedeltà, nel modo seguente:

- Se un cliente ha acquistato nel periodo di riferimento e non aveva mai acquistato prima, è un “cliente nuovo”.

- Se un cliente ha acquistato nel periodo di riferimento, però, aveva già acquistato precedentemente, è un “cliente fedele”.
- Invece, se un cliente non ha acquistato nel periodo di riferimento ma prima lo faceva non aveva, esso è un “cliente perso”.



Figura 5.7: Matrice di profondità a tre livelli: continente, stato, ragione sociale del cliente/fornitore. Per ogni record viene visualizzato il fatturato in valore assoluto e il grado di fedeltà del cliente. Sulla sinistra è presente il filtro di selezione del periodo su cui basare la variabile loyalty.

Da questa matrice è stato possibile approfondire il comportamento di alcuni clienti ed è utile ai ricercatori di opportunità di vendita per inoltrarsi nel dettaglio sui clienti di interesse. Inoltre, è risultato come una enorme fetta del fatturato americano sia derivato da vendite di prodotti verso le filiali (71%) piuttosto che al cliente finale direttamente, questo poiché il mercato americano è soggetto a vincoli di certificazioni e procedure più stringenti.

Da queste prime visualizzazioni abbiamo potuto vedere come la realtà di Steelco sia attiva in tutto il mondo, però ci siamo chiesti quanto effettivamente questo progresso sia stato indotto dalle partnership avvenute nell’ultima decina d’anni. Per rispondere, si è deciso di rappresentare la sua crescita delle vendite, all’estero, nel tempo.





Figura 5.8: Line graph del fatturato medio annuale in Italia e all'estero.

Si nota come la crescita del fatturato mondiale negli anni sia a tutti gli effetti una scalata, della quale si evidenziano i gradini percorsi ogni anno. Allo stesso tempo, vediamo come le vendite in Italia, pur mantenendo un trend di leggera crescita, non stiano facendo un salto di quantità, conseguenza del progetto di globalizzazione messo in atto dall'impresa.

### 5.3.1 L'Italia

Come nel mondo vediamo aree più densamente popolate di vendite, è d'interesse capire come cambia il fatturato in Italia, nelle diverse regioni. È stata svolta una analisi simile alla precedente, filtrando i dati per gli ultimi 5 anni in Italia; si ottiene:



Figura 5.9: Dashboard delle visualizzazioni geo-localizzate per l'Italia.

La heatmap sulla sinistra svolge un ruolo importante nella comunicazione diretta dell'influenza di ogni regione nel fatturato totale complessivo, che viene poi approfondita navigando nella matrice gemella alla precedente per la visione globale.

Si può affermare che al nord Italia sono convogliate la quasi totalità delle opere commerciali, con un fatturato percentuale sul totale italiano di oltre il 68%, il 23% solo in Veneto.

Un dato illustrato maggiormente dal grouped bars plot, suddiviso per macro regione italiana, nel quale vediamo un picco di vendite per il centro Italia nel 2020, anno del covid, dovuta ad importanti investimenti da parte della regione Sardegna.

Inoltre, sulla destra viene illustrata la proporzione tra acquisti da parte di privati e di pubbliche amministrazioni, pari a tre quarti circa per i privati e un quarto per le PA, una tematica interessante per la quale si entrerà poi nel dettaglio.

## 5.4 Le vendite nel tempo

Abbiamo visto come il trend di espansione del business globale sia positivo, in Italia più al nord che nel centro-sud, ma vogliamo trarre qualche informazione in più sulla distribuzione delle vendite nel tempo.

Sono presenti ricorrenze o stagionalità?

L'obiettivo è quello di esplorare i dati su fatturato e quantità nel tempo, cercando di cogliere, attraverso indicatori più specifici, una rappresentazione dell'andamento delle vendite, scarna di influenze che generano rumore nei dati.

Per poter approfondire questa analisi, si è deciso di suddividere i dati di vendita in totali mensili, così da valutarne i cambiamenti stagionali e calcolare misure specifiche attraverso le tecniche di analisi delle serie storiche viste in precedenza.

Prima di tutto, si è deciso di visualizzare la situazione temporale, degli ultimi 5 anni, attraverso una forma tabellare di heatmap, dove i dati del fatturato vengono suddivisi nel tempo e colorati più intensamente in base alla loro dimensione.

Mese	FATTURATO					Totale
	2019	2020	2021	2022	2023	
gennaio	8.771.247,84	8.478.787,38	8.486.240,88	8.240.074,99	8.884.840,84	53.794.773,83
febbraio	7.867.840,88	8.088.794,27	8.867.800,07	7.830.398,00	8.798.893,88	51.834.771,87
marzo	8.113.738,83	8.148.871,38	8.807.142,87	7.827.888,88	8.383.778,88	51.160.088,81
aprile	8.849.888,88	8.888.888,88	8.871.188,81	7.888.888,88	8.178.888,88	48.198.888,88
maggio	8.188.888,88	7.888.888,88	7.888.888,88	7.888.888,88	8.188.888,88	47.888.888,88
giugno	7.888.888,88	7.888.888,88	7.888.888,88	7.888.888,88	7.888.888,88	47.888.888,88
luglio	8.788.888,88	7.888.888,88	7.888.888,88	7.888.888,88	8.788.888,88	48.288.888,88
agosto	7.888.888,88	8.488.888,88	8.488.888,88	7.888.888,88	8.888.888,88	48.188.888,88
settembre	7.188.888,88	8.888.888,88	8.188.888,88	7.888.888,88	7.188.888,88	48.188.888,88
ottobre	8.888.888,88	7.888.888,88	8.888.888,88	7.888.888,88	8.888.888,88	48.888.888,88
novembre	8.888.888,88	7.888.888,88	8.888.888,88	7.888.888,88	8.888.888,88	48.888.888,88
dicembre	7.188.888,88	7.888.888,88	7.888.888,88	7.188.888,88	7.188.888,88	47.188.888,88
<b>Totale</b>	<b>100.000.000,00</b>	<b>100.000.000,00</b>	<b>100.000.000,00</b>	<b>100.000.000,00</b>	<b>100.000.000,00</b>	<b>600.000.000,00</b>

Figura 5.10: Heatmap del fatturato mensile negli anni 2019-2023.

L'evidenza principale, qui, è la crescita lineare del fatturato e della densità di fatturazione negli ultimi mesi dell'anno; per verificare la veridicità di questa ipotesi è stato svolto uno studio temporale approfondito. Per evidenziare il trend di crescita, è stata calcolata una centratura di due medie mobili aritmetiche semplici, di ordine 12. Sottraendo il trend calcolato alla distribuzione dei dati veri, si sono ottenuti i coefficienti specifici di stagionalità, che rappresentano il discostamento dei valori osservati rispetto a quelli di tendenza; poi, vengono calcolati i coefficienti prima grezzi poi ideali. I coefficienti ideali – CI, descrivono la stagionalità del processo, ad esempio, per una gelateria, saranno molto positivi in estate e negativi in inverno. Nel nostro caso:



Figura 5.11: Bar chart dei coefficienti ideali di stagionalità del fatturato di vendita, espressi come fatturato mensile in valore assoluto.

Notiamo che le evidenze riguardo la vendita maggiore negli ultimi mesi, più che nei primi, dell'anno, che avevamo colto in tabella vengono certificate.

Inoltre, attraverso gli indicatori calcolati, è stato possibile ottenere una rappresentazione approfondita dell'andamento delle vendite nel tempo, graficato secondo il seguente line graph composto.

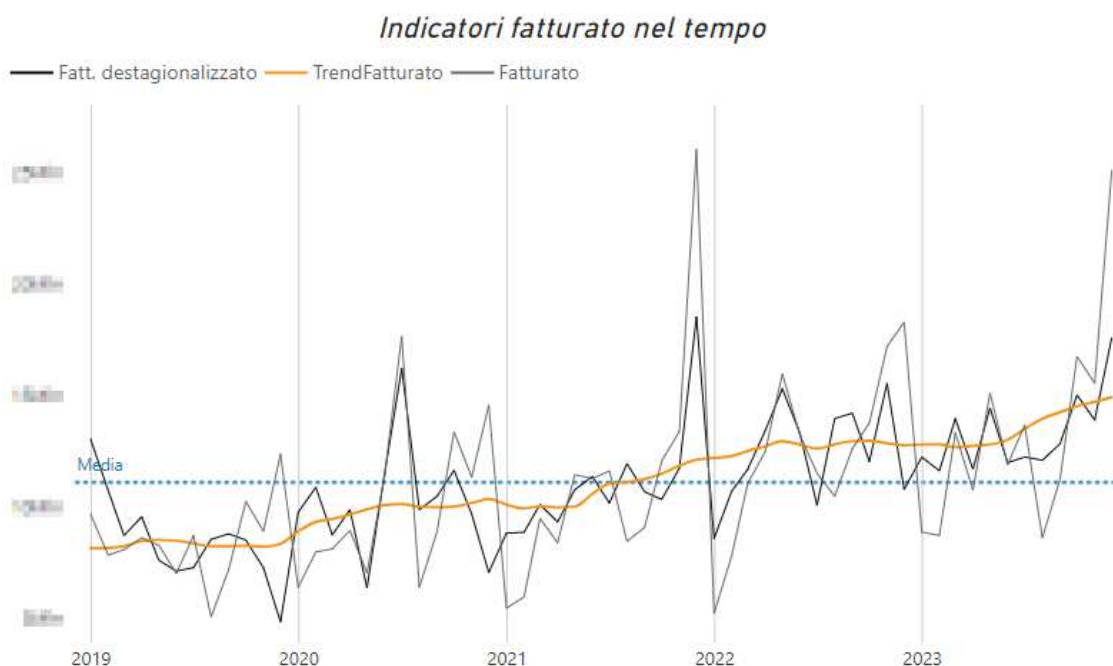


Figura 5.12: Line graph composto dei principali indicatori del fatturato negli anni 2019-2023, fatturato espresso come valore mensile.

Da questa visualizzazione vediamo le osservazioni rappresentate con intenzioni differenti:

- Le linee nera e grigia, da un lato, figurano il fatturato e il fatturato spogliato della componente stagionale, entrambe rappresentazioni “pure” del dato, del quale si possono visualizzare in maniera nitida i picchi di vendita.
- Le linee azzurra e arancione, invece, rappresentano la media e il trend, ottenuto come centratura di due medie mobili, due indicatori chiave per cogliere l'andamento generale delle vendite.

Da esse vediamo, rispettivamente, la “soglia”, intesa dal mondo vendita come l’obiettivo minimo da raggiungere e il modello attuale, che sintetizza l’andamento del progresso, idealmente da mantenere linearmente positivo.

In aggiunta, è stata costruita una pagina collegata di approfondimento, che appare come un pop-up non appena si passa il cursore in una finestra temporale di interesse della heatmap tabellata precedentemente descritta. Così, è possibile evidenziare le differenze di fatturato generato, rispetto al mese e all’anno precedenti, due indicatori spesso utilizzati internamente come riferimento per i commerciali.

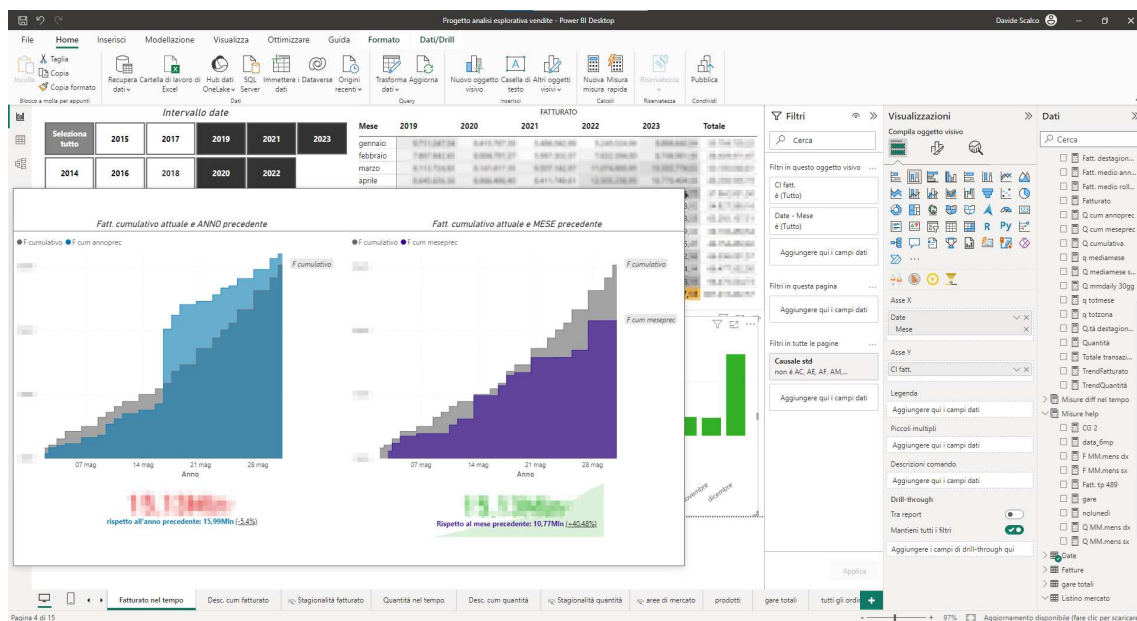


Figura 5.13: Area graph doppio, sulla sinistra fatturato cumulativo in relazione all’anno precedente, sulla destra fatturato cumulativo in relazione al mese precedente. Ai piedi del grafico, schede KPI con l’incremento percentuale rispetto al periodo di riferimento.

### 5.4.1 AERs

La stessa tipologia di analisi è stata condotta anche per lo studio delle quantità vendute nel tempo, dove, approfondendo l’analisi per tipologia di prodotto venduto, si sono cercate anomalie e/o particolarità nella distribuzione delle vendite.

Un interessante spunto di riflessione è nato dalle ricorrenze anomale di vendita presenti per la categoria AERs – Automated Endoscopy Re-processor, ovvero la fascia di prodotti specifici per l’endoscopia, che si presenta nel seguente modo:

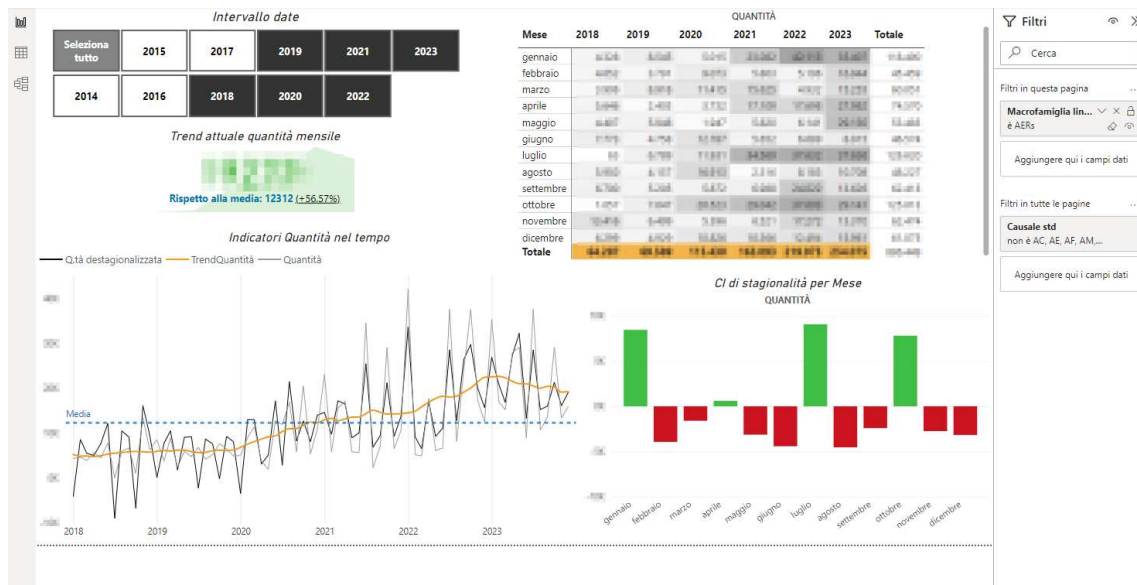


Figura 5.14: Dashboard sulle quantità vendute, filtrato per macro famiglia prodotto = AERs.

Vediamo una tendenza crescente, molto positiva, attualmente addirittura pari ad un incremento del 57% sulla media, con un totale di macchine vendute che dal 2018 al 2023 è quadruplicato.

Allo stesso tempo scorgiamo una forte variabilità delle vendite nell’arco dell’anno, contraddistinte da una stagionalità forte nei mesi di gennaio, luglio e ottobre, negativa negli altri casi; una situazione particolare, per la quale sono state approfondite le dinamiche di vendita della macro famiglia AERs, visualizzandone lo storico di vendita e i dettagli.

È saltato subito all’occhio un importante contributo al quantitativo di vendite dovuto alle fatture di noleggio, una tipologia di contratto in crescita negli ultimi anni e particolarmente diffusa per questo segmento di prodotti.

Nel dettaglio, costruendo il grafico a cascata seguente, si evidenzia la preponderanza della macro famiglia produttiva AERs nel totale dei noleggi, del quale firmano il 95% dei contratti.

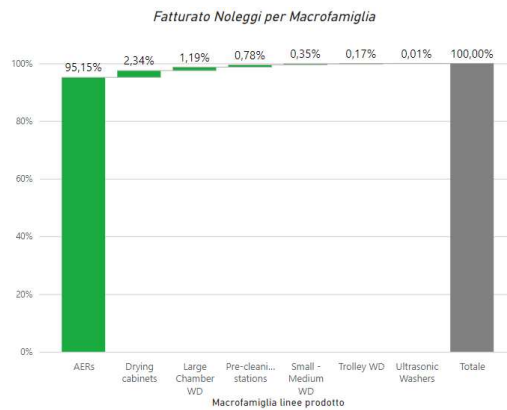


Figura 5.15: Waterfall chart del fatturato percentuale sul totale dei noleggi, suddiviso per macro famiglia di linee prodotto.

Essendo dei contratti meno dispendiosi e più frequenti rispetto ad una vendita ordinaria, è naturale che vi siano dei picchi in corrispondenza delle scadenze di contratto, che nella maggioranza dei casi sono per l'appunto: gennaio, luglio e ottobre. Vediamo come cambierebbe il quantitativo di vendite per il gruppo AERs, senza considerare i noleggi:

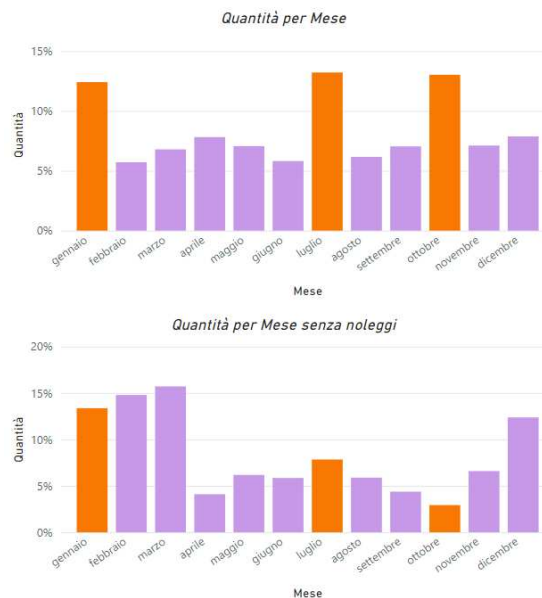


Figura 5.16: Bar chart doppio della quantità percentuale di vendite mensili, della macro famiglia produttiva AERs.

Vediamo una sostanziale differenza, nello specifico, ottobre passa ad essere il mese con meno macchine vendute, a sottolineare il contribu-

to che i noleggi danno alla vendita di questa categoria di macchine. I noleggi prevedono contratti longevi, con introiti sicuri nel medio-lungo periodo, il che è un ottimo segnale per la stabilità dell'azienda e, soprattutto, per l'ottimizzazione del magazzino. Alla firma del contratto, solitamente, vengono definite tutte le specifiche: dal numero di macchine coinvolte, alla possibile integrazione successiva di ulteriori, fino ai programmi di manutenzione, ricambi e chimici necessari per garantirne il funzionamento. Calendarizzare la produzione, secondo gli accordi stipulati, è essenziale per poter definire quantitativo di scorta minimo, punti di riordino e altre dinamiche di ottimizzazione del magazzino.

## 5.5 Panoramica di mercati e prodotti

Così come i prodotti per l'endoscopia, navigando tra le informazioni disponibili sarà possibile scorgere altre anomalie nella distribuzione delle vendite. Si è deciso, dunque, di riassumere le vendite secondo le aree di mercato nel quale Steelco è attiva e i principali prodotti in esse venduti, tenendo a mente che storicamente il business si divide in due tipologie:

1. Il processo di lavaggio e disinfezione, con il quale nasce l'attività imprenditoriale
2. La sterilizzazione, integrata in seguito all'acquisizione di Icos

Le aree di mercato vengono segmentate in modo da vedere il contributo di ogni divisione all'area di riferimento come segue:

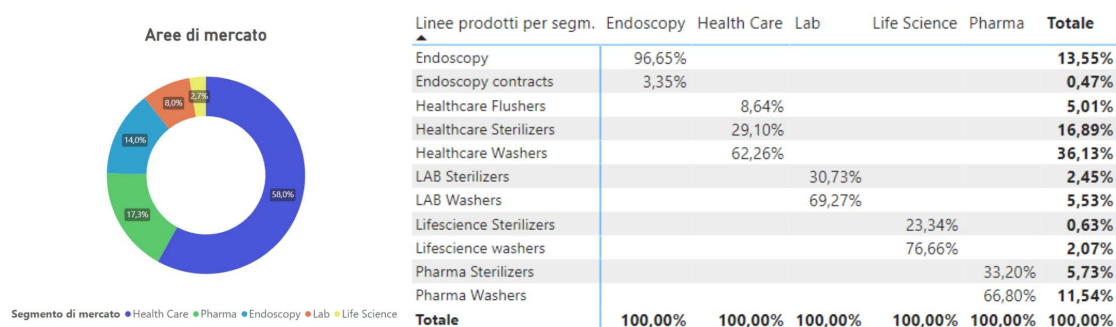


Figura 5.17: A sinistra donut chart del fatturato suddiviso in percentuale tra le aree di mercato. A destra tabella di approfondimento delle aree di mercato suddivise ulteriormente per linea prodotto, in cella valore percentuale del fatturato sul totale di colonna.



Per washers, si intendono i macchinari “WD – washer-disinfectors”, dunque responsabili dei processi di lavaggio e disinfezione, i quali, occupano la porzione maggiore di ogni area di mercato. Per l’endoscopia, vista l’importanza, è stata creata una classificazione apposita, composta solamente però da strumenti di lavaggio e disinfezione. Nello specifico, ad oggi, i prodotti WD generano il 74% del fatturato, il restante 26% riguarda gli sterilizzatori, un risultato comunque importante, poiché raggiunto nel corso degli ultimi 10 anni.

Vediamo com’è cresciuta la forza di questo business negli anni:

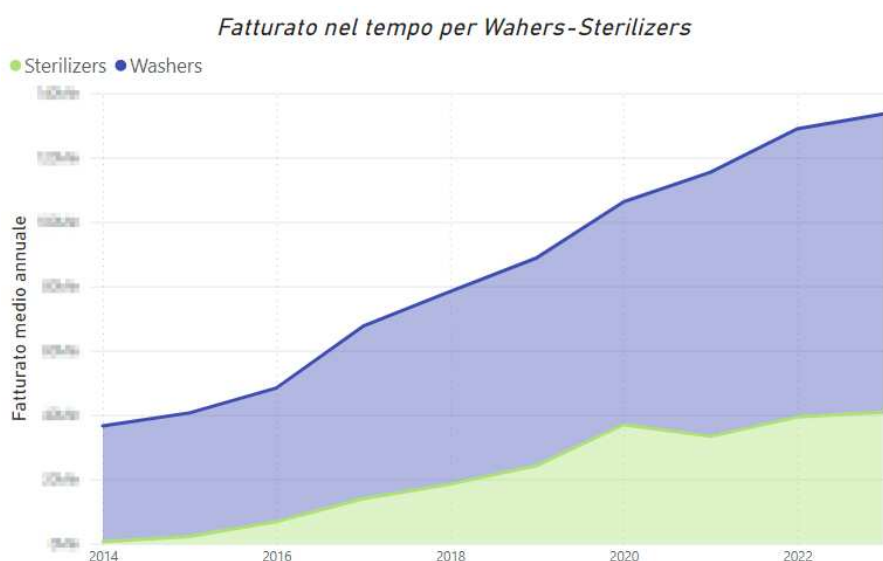


Figura 5.18: Area graph del fatturato medio annuale, negli anni 2014-2023, suddiviso per linee prodotto washers-sterilizers.

Dal grafico ad aree viene sottolineata la crescita della vendita di sterilizzatori dopo l’acquisizione di Icos nel 2013, che, infatti, parte da un livello approssimabile a zero, per poi raggiungere la quota di un quarto del fatturato.

In seguito, si è voluto vedere nel dettaglio quali categorie di prodotti alimentano maggiormente il business, ad esempio, se avvengono vendite massive di macchine di piccole dimensioni, o se l’impresa lavora più con i grandi sistemi.

Per farlo rappresentiamo le macro famiglie dei prodotti in un grafico a barre, dove le colorazioni differenti delle colonne riescono a comunicarci più rapidamente se si tratta di dispositivi di lavaggio-disinfezione o sterilizzazione.

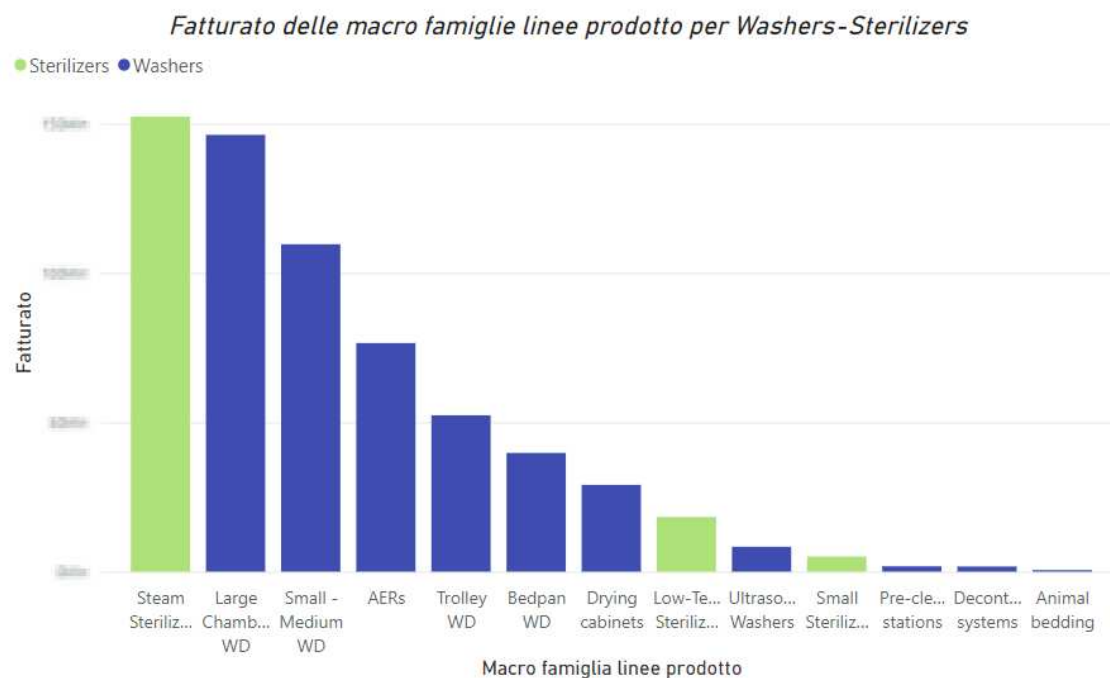


Figura 5.19: Bar chart del fatturato suddiviso per macro famiglia produttiva e linee prodotto washers-sterilizers.

Ciò che risulta dal grafico, è l'evidenza che la categoria degli sterilizzatori a vapore da un contributo molto importante alla fatturazione totale: degli sterilizzatori in primis, poiché gli introiti generati dalle altre categorie di sterilizzatori sono trascurabili; ma anche delle entrate generali, classificandosi al primo posto tra le macro famiglie di prodotti venduti.

Le vendite per le macchine WD invece, sono più equamente distribuite tra prodotti di grandi e piccole dimensioni.

In ogni categoria, sono presenti svariati modelli di macchina e un aspetto di fondamentale interesse per gli operatori del mondo vendite è vedere quali prodotti sono più venduti e quali generano le entrate maggiori, in modo da definire le strategie di marketing più efficaci.

Perciò, si è deciso di rappresentare i migliori 10 modelli per fatturato generato e per quantità di macchine vendute, questi ultimi sono stati filtrati dei noleggi poiché, come analizzato precedentemente, le stime per

la categoria AERs sarebbero state errate, influenzando sulla comunicazione visiva generata.

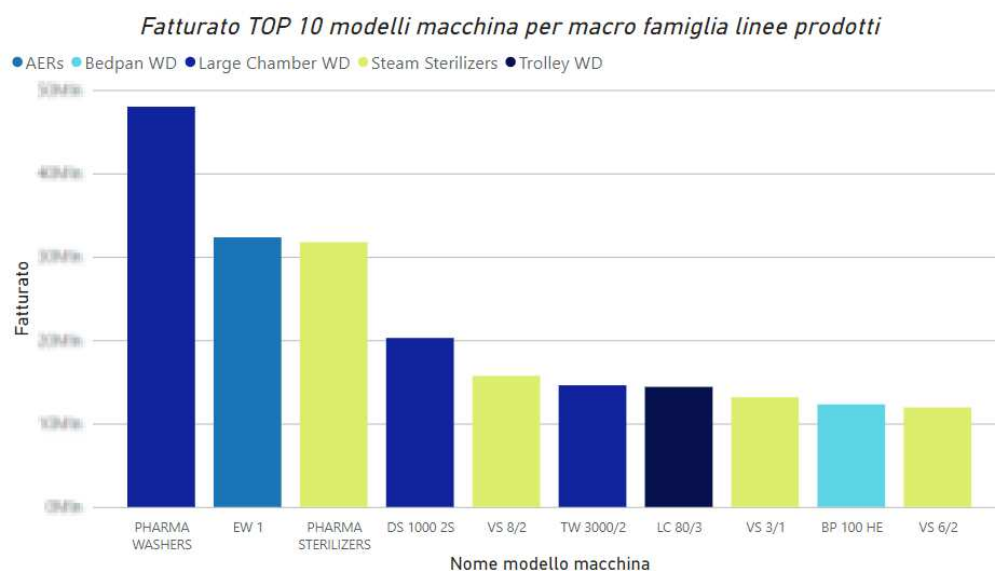


Figura 5.20: Bar chart del fatturato, suddiviso per modello macchina e macro famiglia produttiva.

Da questa rappresentazione notiamo la prevalenza dei modelli Pharma tra i più venduti, che occupano le posizioni prima e terza. Essi sono fautori, di conseguenza, di un grosso contributo al fatturato derivante dalle loro macro famiglie di appartenenza, ovvero, Steamsterilizers e Large chamber WD, che avevamo visto occupare la prima e seconda posizione delle macro famiglie produttive più vendute.

Questa classificazione modello pharma – macro famiglia, non sempre è esatta; per convenzione vengono associati tutti i prodotti washers alla linea “large chamber WD” e gli sterilizers a “steam sterilizers”, però, non è corretto nel 100% dei casi, generando così, dei valori sovrastimati.

Tuttavia, si può dire con certezza che dai prodotti di maggiori dimensioni derivano la maggioranza delle entrate; nei primi 10 prodotti a maggior fatturato generato troviamo:

- Per i prodotti WD, 4 su 6 fanno parte delle linee di dimensioni più grandi, mentre i prodotti EW1 e BP100HE sono, rispettivamente, lava-endoscopi e lava-padelle. Vengono rappresentati con sfumature di blu più scure i prodotti più grandi e costosi, sfumando man mano verso l’azzurro.

- Per gli sterilizzatori invece, vediamo solamente prodotti appartenenti alla famiglia degli strumenti di maggiori dimensioni, le altre, come visto precedentemente, danno un contributo irrisorio.

Se da questa rappresentazione vediamo come i prodotti di dimensioni più grandi generino la fetta maggiore del fatturato, attraverso il focus per quantità di macchine vendute, si nota la prevalenza dei modelli di dimensioni, dunque fascia di prezzo, più ridotte.

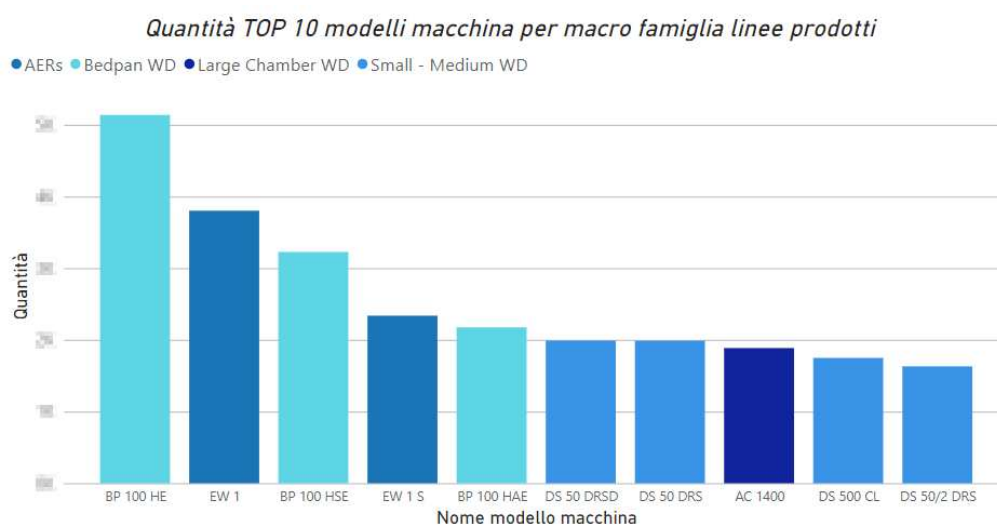


Figura 5.21: Bar chart della quantità venduta, suddiviso per modello macchina e macro famiglia produttiva.

Qui si evidenzia come gli sterilizzatori siano sistemi costosi, venduti in quantità più ridotte, infatti, nessun modello di sterilizzatore è presente tra i migliori 10 per quantità. Lo stesso vale per il segmento washer, ad eccezione delle AC1400, per il quale vediamo una TOP5 dominata dalle famiglie produttive di lava-endoscopi (AERs) e lava-padelle (Bedpan WD), a seguire sistemi di lavaggio di medio-piccole dimensioni.

Gli unici modelli presenti in entrambe le rappresentazioni sono EW1 e BP100HE, modelli di spinta per il business aziendale, per il quale è bene tenere a mente che la richiesta sarà sempre alta in termini di quantità, ed è bene che sia soddisfatta visti i risultati in termini di fatturato.

## 5.6 La pubblica amministrazione in Italia

Vista la presenza di un reparto interno dedicato agli appalti pubblici, si è deciso di approfondire le vendite generate dalle pubbliche amministrazioni italiane. Attraverso un incontro con gli addetti a queste pratiche di vendita, riassumendo, sono state tratte le seguenti informazioni:

Tutti gli ordini di acquisto (e le fatture derivate) conseguiti da PA italiane, derivano dallo svolgimento di una gara d'appalto. Spesso si ricevono delle comunicazioni, precedentemente alle gare, finalizzate a consultazioni di mercato per vedere le offerte disponibili, o manifestazioni di interesse per informarsi sui fornitori disposti a competere. In seguito, vengono generate le gare pubbliche per le quali un solo vincitore si aggiudicherà l'appalto delle forniture richieste. Può esserci una base d'asta, come no, e la competizione non è solamente sul prezzo, ma anche sulla qualità. In seguito, sono stati recuperati dei file Excel parzialmente compilati dai colleghi dell'ufficio gare, integrati con le informazioni registrate a sistema, in modo da avere un'altra base su cui poter indagare. Dopo averli puliti di errori e caricati nel software, ho filtrato i protocolli di gara archiviati, quelli riferiti a consultazioni di mercato e a manifestazioni di interesse.

Ne è risultato un dataset di quasi 1000 gare a cui Steelco ha partecipato negli ultimi 8 anni, distribuite nel tempo nel seguente modo:



Figura 5.22: Area graph del quantitativo di gare negli anni 2016-2023, asse y scalato a quantità 95 di partenza anziché 0.

La griglia di rappresentazione dell'area graph, riscalata nell'asse delle ordinate, evidenzia il maggior numero di richieste da parte delle PA italiane durante gli anni 2020-2021, colpiti dall'avvento del Covid-19. Ogni anno, comunque, si parla di almeno un centinaio di gare, mediamente 122, che conferiscono una base solida di fatturazione per l'azienda, considerando che il 40% delle gare a cui partecipa vengono vinte.

Le gare sono state categorizzate nel seguente modo:

- Washers, comprendono tutte le tipologie lavaggio e disinfezione ad eccezione di lava-endoscopi e lava-padelle, per i quali, vista la richiesta, sono state fatte delle classificazioni apposite.
- Sterilizers, comprendo tutte le tipologie di sterilizzatori
- AERs per lava-endoscopi e forniture di supporto
- Bedpand per lava-padelle e forniture di supporto
- Accessories & Furnitures, per quanto riguarda forniture di accessori, strumenti, sistemi di tracciabilità, servizi di manutenzione e chimici
- Misto, quì convergono le richieste di prodotti provenienti da diverse di queste categorie, di particolare interesse sono le centrali di sterilizzazione – CSSD, che prevedono progetti estremamente complessi, possibili negli ultimi anni grazie all'avvento di Miele.

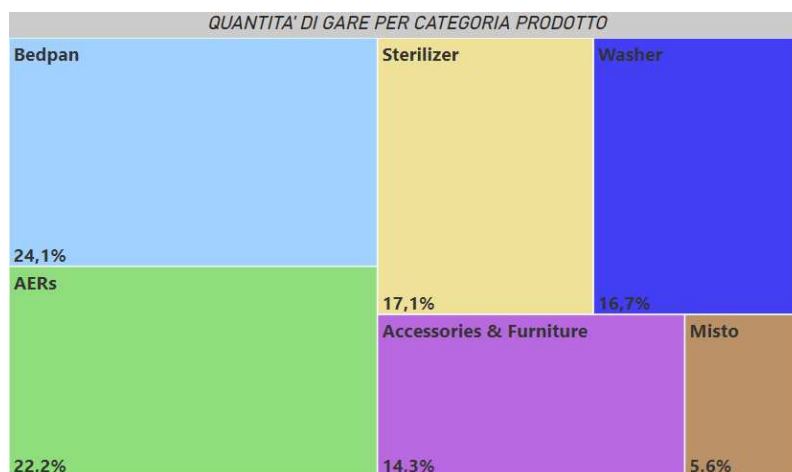


Figura 5.23: Treemap del quantitativo di gare percentuali, suddivise per categoria prodotto.

Vediamo come le richieste dei prodotti AERs e Bedpand, siano le più diffuse, così da dare un contributo importante alla vendita totale di questi modelli di prodotto, che precedentemente abbiamo visto essere i più venduti in termini di quantità.

Successivamente, sempre in merito alle suddivisioni categoriali, si è deciso di chiarire in quali Steelco riesce ad essere più competitiva. Per farlo, si sono sfruttati i dati riguardo l'importo presunto a base d'asta e il prezzo effettivamente proposto dall'azienda, creando una misura di differenza domanda-offerta per ognuna delle categorie discusse. Il risultato di quest'analisi è raffigurato secondo un boxplot multiplo, nel quale viene rappresentata la distribuzione di queste differenze nelle varie categorie.

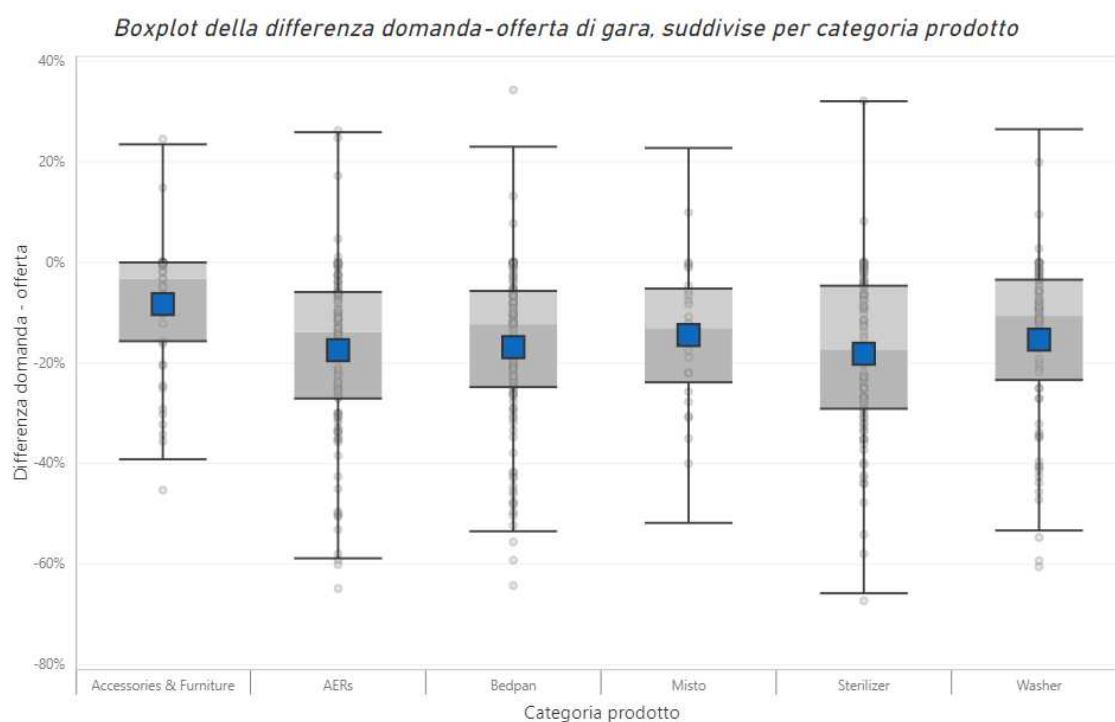


Figura 5.24: Boxplot della differenza domanda-offerta di ogni gara, suddivise per categoria.

Da questa visualizzazione, non emerge una categoria prodotto per la quale Steelco riesce ad offrire soluzioni particolarmente economiche. Mediamente, il ribasso sulla base d'asta aggira intorno al 16%, leggermente più economico per quanto riguarda lava-padelle, lava-endoscopi e sterilizzatori, le cui gare, effettivamente, vengono vinte in percentuale più spesso.

Un altro fattore particolarmente importante, è la distribuzione delle strutture appaltatrici all'interno del territorio italiano, nonché, della rete di distributori con cui Steelco lavora nelle varie regioni italiane.

La presenza di maggiori progetti di sviluppo e strutture nell'economia del nord Italia, suggerisce che qui vengano indotte un numero maggiore di gare, decrescenti man mano dal centro verso il sud.

La situazione è la seguente:

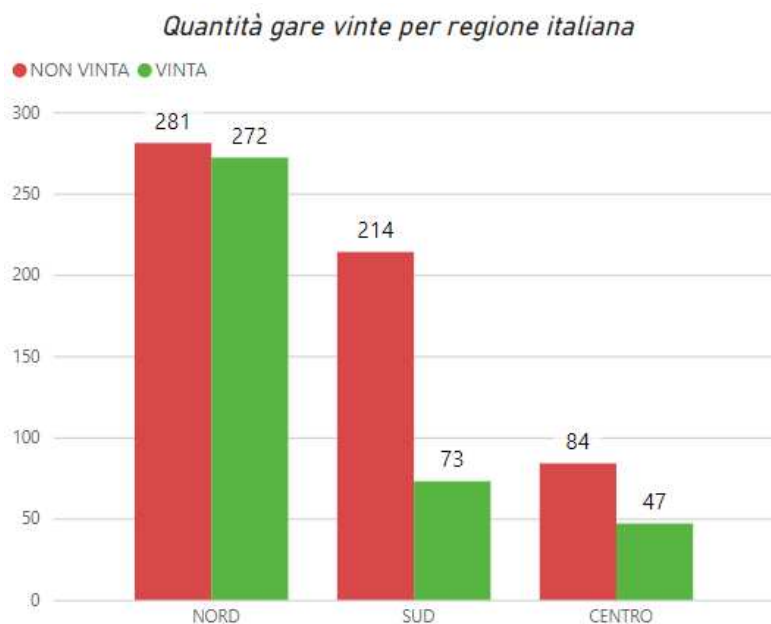


Figura 5.25: Grouped bar plot del quantitativo di gare per macroregione italiana, colonne affiancate per la suddivisione vinta - non vinta .

Come presunto, vediamo una netta distinzione tra le macro regioni italiane, con un fatturato generato dalle gare nel nord Italia pari all'85% del totale. Ciò è dovuto all'alta percentuale di vincita delle gare d'appalto, quasi una su due, conseguenza di una storica presenza nel territorio, nonché di una più strutturata rete di distributori.

I distributori, offrono all'azienda maggiori opportunità di riuscire a partecipare alle gare su invito e una disponibilità di risorse più immediata, utile per soddisfare bisogni urgenti come nel periodo da poco trascorso di pandemia. Steelco, oggi, dispone di una rete di distributori in tutta Italia, più ramificata al nord ma comunque presente in ogni regione, ad eccezione della Toscana.



Con il seguente grafico a barre orizzontali, si è voluta approfondire la situazione per regione, osservando la distribuzione percentuale del fatturato generato dalle gare suddiviso per regione.

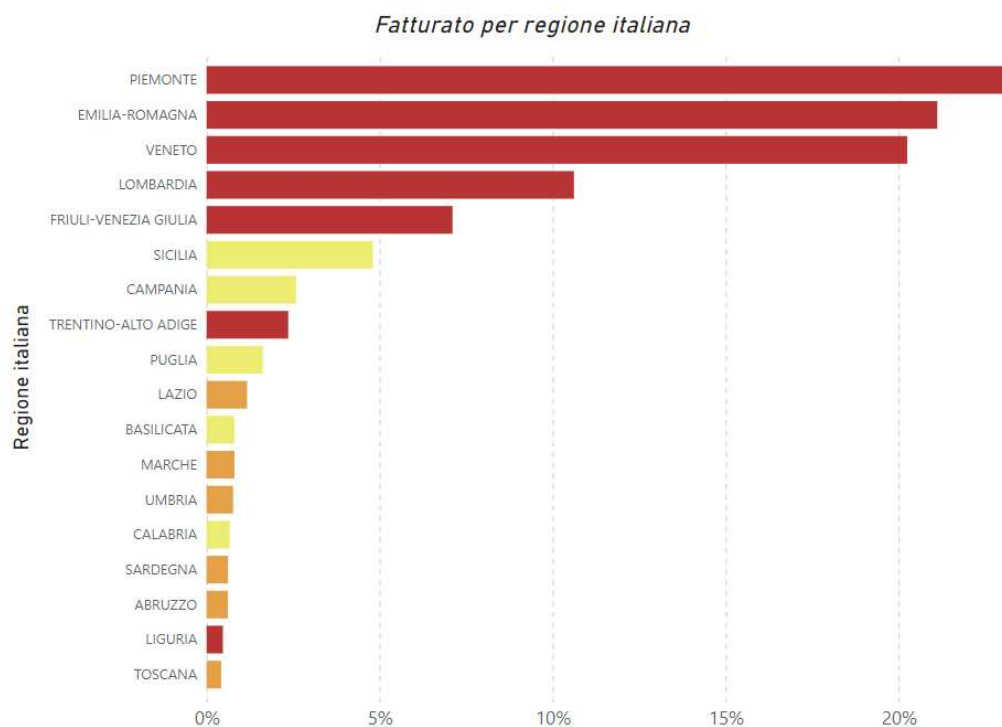


Figura 5.26: Bar chart del fatturato percentuale sul totale generato da gare, suddiviso per regione italiana, barre colorate in base alla macro regione geografica.

In questo grafico, viene evidenziata la preponderanza delle regioni del nord che occupano le prime 5 posizioni, per un totale dell'82% del fatturato totale generato da gare.

Inoltre, viene confermata l'ipotesi precedentemente descritta.

La regione Toscana, vista la mancanza di distributori attivi nel territorio, occupa l'ultima posizione tra le 20 regioni nel contributo percentuale al fatturato generato da gare.

Un dato particolare, essendo comunque responsabile del 2% delle vendite totali in Italia (non generate solamente da gare), che è un valore nella media per le regioni del centro-sud.

Dalla dashboard riferita alle vendite in Italia, cliccando sulla regione Toscana è possibile sfruttare le potenzialità di Power BI per generare la seguente modifica interattive dei grafici.

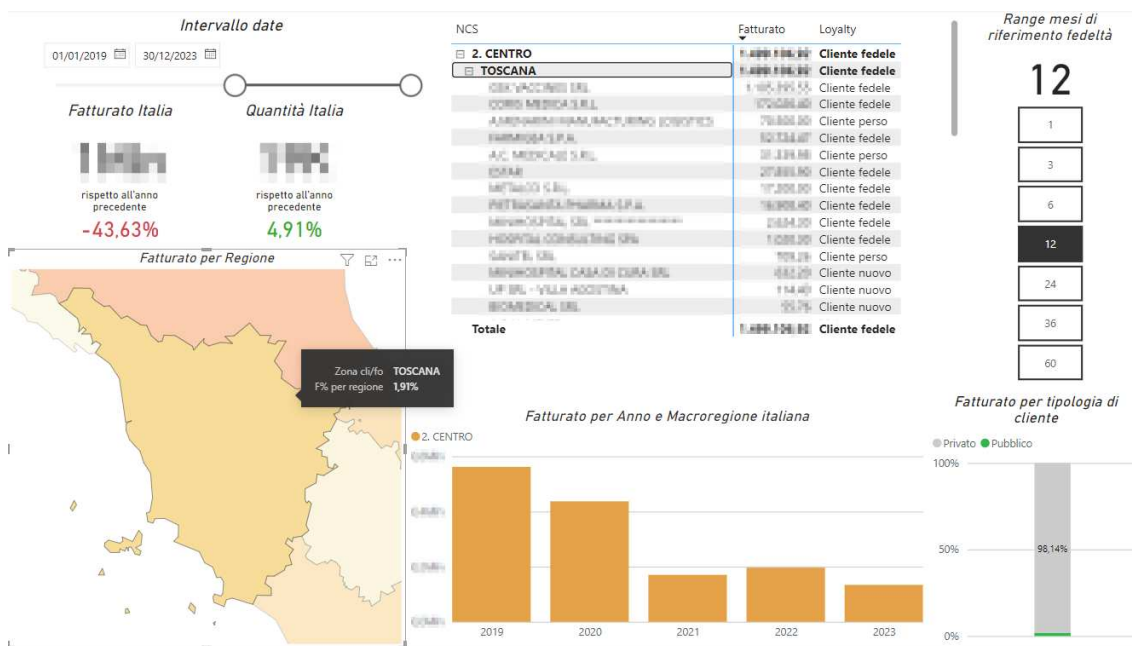


Figura 5.27: Dashboard delle visualizzazioni geo-localizzate in Italia, filtrate per la regione Toscana.

Da questa panoramica, vediamo in rilievo la Toscana, che è responsabile, più precisamente, dell'1,91% del fatturato totale italiano negli ultimi 5 anni. Invece, dal grafico sulla tipologia cliente pubblico/privato, si evidenzia la presenza quasi nulla (2%) delle vendite a istituzioni pubbliche che, per esempio, raggiunge un valore circa la metà (48%) per l'Emilia-Romagna.

## 5.7 Conclusioni

I fattori da analizzare per riuscire ad avere chiara la situazione che il mondo delle vendite Steelco sta vivendo sono molteplici, l'approccio esplorativo - descrittivo utilizzato per cercare di coglierli è stato molto funzionale ed esplicativo. Tuttavia, è naturale che alcune dinamiche siano riuscite a sfuggire e debbano essere approfondite per poter sciogliere il problema alla fonte.

Il mio ruolo come risorsa per l'analisi esplorativa delle vendite si è focalizzato nel sintetizzare gli indicatori riguardanti la sfera commerciale, così come altre risorse hanno potuto approfondire aspetti del business differenti attraverso ulteriori analisi.

Attraverso gli studi svolti, abbiamo potuto riscontrare delle dinamiche interessanti, da utilizzare come spunto di analisi più complesse, come, per esempio, la gestione dei noleggi e delle vendite di gara per l'ottimizzazione del magazzino. Grazie agli introiti "sicuri" generati dai modelli analizzati (lava-endoscopi EW1 e lava-padelle BP100HE) attraverso contratti di noleggio o richieste ricorrenti da parte delle PA, sarà possibile migliorare la gestione di scorte dei macchinari in causa, delle forniture di manutenzione necessarie e calcolare i cicli di chimici previsti per il loro funzionamento.

Allo stesso tempo, è stato svolto un importante compito di comunicazione delle informazioni, grazie alle quali è possibile che intuizioni diverse siano colte dagli utenti visualizzatori delle ricerche, che probabilmente, avendo maggiore esperienza, riusciranno ad interrogare i grafici più nel dettaglio. L'importanza, dunque, non ricade solamente sulle evidenze ottenute nel corso dell'analisi, ma sulla capacità descrittiva delle variabili coinvolte nei processi di vendita. Esse sono state sintetizzate in dashboard intuitive, create seguendo le leggi della corretta rappresentazione visiva e allineando misure e indicatori utilizzati alle comunicazioni attualmente attive tra i commerciali. La continua misurazione delle performance, attraverso gli strumenti di business intelligence, fornisce una panoramica costantemente aggiornata dello stato dell'organizzazione, che è la chiave per ottimizzare i processi e risparmiare risorse.

Il progetto di comunicazione attraverso le piattaforme online di Power BI, mi ha permesso di condividere i report creati con i diretti portatori di interesse, i quali potranno avere un quadro più nitido della situazione attuale nel mondo vendite. Le panoramiche che erano caricate precedentemente nell'ambiente server, si limitavano a rappresentazioni tabellari e pochi semplici grafici, creati e aggiornati periodicamente grazie alle funzionalità del software. Quindi, la condivisione di queste nuove visualizzazioni vuole essere uno spunto di riflessione sulle analisi condotte, nonché, una soluzione per alimentare la corretta comunicazione dell'andamento dei processi, in modo da fornire supporto nelle decisioni strategiche.

Nell'economia moderna, in breve, ciò che è necessario che Steelco faccia per progredire, è sfruttare le innovazioni tecnologiche integrate per riuscire ad organizzare le informazioni, così da avere chiari i punti di forza e di debolezza del proprio business.

Poiché, senza conoscere il proprio funzionamento, non può pensare di migliorarlo.

# Bibliografia

- Borghetti P. (2022). La storia di as400 e la figura del capocentro.
- CE E. C. (2022). Indice di digitalizzazione dell'economia e della società (desi) 2022.
- Cicchitelli G.; d'Urso P.; Minozzo M. (2017). *Statistica: principi e metodi*. Pearson.
- DiFonzo T.; Lisi F. (2005). *Serie storiche economiche, Analisi statistiche e applicazioni*. Carocci editore.
- Istat (2019). Digitalizzazione e tecnologia delle imprese italiane.
- Istat (2023a). Censimento permanente delle imprese 2023: primi risultati.
- Istat (2023b). Nel 2023 piccole e medie imprese ancora penalizzate in attività specialistiche di digitalizzazione.
- Learn M. (2023). Che cos'è power bi?
- PowerPlatform M. (2023). Funzionamento della business intelligence.
- SAP (2023). What is data mining?
- Wikipedia (2022). Toyota production system.
- Wikipedia (2023a). Era dell'informazione.
- Wikipedia (2023b). Garbage in, garbage out.
- Wikipedia (2024). Big data.
- Wilke C. O. (2019). *Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures*. O'Reilly Media, Inc.
- Wirth R.; Hipp J. (2000). Crisp-dm: Towards a standard process model for data mining.