

Università degli Studi di Padova
Dipartimento di Scienze Statistiche
Corso di Laurea Triennale in
Statistica e Gestione delle Imprese



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

**RELAZIONE FINALE
ANALISI DI DATI FACEBOOK:
STUDIO DELLE RELAZIONI TRA ATTIVITÀ IN FACEBOOK
E SCALE PSICOLOGICHE DI BENESSERE E PERSONALITÀ**

Relatore Prof. Livio Finos
Dipartimento di Scienze Statistiche

Correlatore Prof.ssa Claudia Furlan
Dipartimento di Scienze Statistiche
Correlatore Dott.ssa Claudia Marino
Dipartimento di Psicologia dello Sviluppo e della Socializzazione

Laureando: Anna Petrina
Matricola N 1077224

Anno Accademico 2015/2016

INDICE DEI CONTENUTI

CAPITOLO 1. Introduzione.....	3
CAPITOLO 2. Le scale psicologiche di personalità e benessere.....	5
2.1 Scala di Personalità-Stabilità Emotiva.....	5
2.1.1 Descrizione.....	5
2.1.2 Validazione.....	6
2.1.3 La scala nel questionario.....	6
2.2 Scala di Benessere Emotivo.....	7
2.2.1 Descrizione.....	7
2.2.2 Validazione.....	9
2.2.3 La scala nel questionario.....	9
2.3 Analisi descrittive: confronto tra scala di Personalità-Stabilità Emotiva e scala di Benessere Emotivo.....	12
CAPITOLO 3. Dati e libreria myFBr.....	17
3.1 Libreria myFBr.....	17
3.2 Presentazione dati.....	17
CAPITOLO 4. Analisi e modelli.....	23
4.1 Modelli scala di Personalità-Stabilità Emotiva.....	24
4.1.1 Modello con tutte le variabili.....	24
4.1.2 Modello con solo le variabili di analisi dei messaggi.....	28
4.2 Modelli scala di Benessere Emotivo.....	31

4.2.1	Modello con tutte le variabili.....	31
4.2.2	Modello con solo le variabili di analisi dei messaggi.....	34

CAPITOLO 5. Conclusioni.....37

Capitolo 1

Introduzione

Lo stage da cui nasce la presente relazione finale è stato realizzato grazie alla collaborazione con il Dipartimento di Psicologia dello Sviluppo e della Socializzazione dell'Università degli Studi di Padova, al fine di indagare come eventuali forme di dipendenze e di attività in Facebook possano essere legate a diversi stati emotivi.

Il dataset che viene utilizzato per le analisi statistiche ha origine dalla somministrazione di un questionario on-line agli studenti del corso di laurea in “Progettazione e Gestione del Turismo Culturale” e di “Psicologia dello Sviluppo e della Socializzazione” nel periodo ottobre-novembre 2015.

Il questionario ha lo scopo di indagare le abitudini e gli atteggiamenti di questi 245 studenti universitari circa Internet e Facebook.

La prima sezione del questionario è formata da quesiti riguardanti il possesso di apparecchi elettronici attraverso cui si può accedere a Internet e misura l'utilizzo di Internet e Facebook in termini di ore; successivamente vengono presentate una serie di domande specifiche sull'utilizzo di Facebook e si conclude con dei quesiti per indagare sulla personalità del rispondente tramite domande personali di atteggiamento, preferenze e interessi.

Ogni risposta del questionario prevede l'assegnazione di un punteggio secondo la scala Likert. Il questionario è stato suddiviso in gruppi di domande che corrispondono alla formazione di determinati test psicometrici.

Il seguente studio si interessa di due scale del questionario: la prima scala ha lo scopo di indagare sulla personalità degli studenti (che sarà successivamente denominata “scala di Personalità-Stabilità Emotiva”) e si trova nella sezione H del questionario; la seconda scala consiste nello studio del benessere della persona (denominata successivamente “scala di Benessere Emotivo”) e si trova nella sezione I del questionario.

Durante lo studio, sono state ricercate delle ulteriori variabili oltre a quelle derivanti dalle risposte al questionario. Queste variabili derivano dall'analisi dei profili Facebook delle persone a cui è stato sottoposto il questionario.

Grazie all'accesso alle informazioni e ai dati personali dei profili Facebook (ottenuti con il consenso degli utenti) sono state ricavate delle variabili tramite l'analisi del testo dei messaggi che gli utenti del profilo hanno inviato e ricevuto negli ultimi 6 mesi.

Prima di ricavare queste variabili, sono stati individuati 12 gruppi di emoticons e 19 parole inerenti alle descrizioni delle scale psicologiche di personalità e benessere che interessano lo studio.

Successivamente sono state ottenute delle variabili derivanti dal conteggio di queste emoticons e parole ricavate dai messaggi inviati e ricevuti dall'utente dei profili Facebook.

Inoltre, sono state individuate delle variabili che descrivono la frequenza di attività e inattività giornaliera e settimanale nella bacheca di Facebook e nello scambio di messaggi nelle chat.

Quindi, dopo aver ottenuto queste variabili, è stato costruito un unico dataset dal quale sono stati creati dei modelli di regressione lineare per studiare se le variabili Facebook possono spiegare i valori delle scale psicologiche di personalità e benessere.

Nel capitolo 2 verranno presentate le descrizioni delle scale psicologiche su cui si basa lo studio, con relative validazioni e analisi descrittive; il capitolo 3 sarà dedicato alla spiegazione del dataset su cui sono state svolte le analisi e della libreria di R da cui sono state recuperate le funzioni per queste analisi (libreria myFBr); nel capitolo 4 saranno presentate le analisi statistiche e i modelli e l'ultimo capitolo (capitolo 5) sarà dedicato alle conclusioni.

Capitolo 2

Le scale psicologiche di personalità e benessere

Con lo scopo di indagare sul benessere e sulla personalità dei rispondenti, nel questionario sono stati presentati 12 test psicometrici composti da batterie di domande.

Per lo studio presentato successivamente saranno utilizzate due batterie di domande che formano valori che descrivono due scale psicologiche: la scala di Personalità (che si trova nella sezione H del questionario), che si riferisce alla personalità dell'individuo, e la scala di Benessere (che si trova nella sezione I del questionario), che si riferisce invece al benessere dell'individuo.

2.1 Scala di Personalità-Stabilità Emotiva

2.1.1 Descrizione

La scala di Personalità-Stabilità Emotiva consiste in una serie di domande che permettono lo studio e l'identificazione della personalità di una persona, attraverso la teoria dei Big Five (BF). Questa teoria risulta essere, tra i modelli incentrati su un approccio nomotetico allo studio della personalità, uno dei più condivisi e testati sia a livello teorico che empirico. Identifica le dimensioni caratterizzanti e le differenze individuali attraverso analisi statistiche di tipo fattoriale. La personalità è infatti data dall'insieme dei tratti di un individuo che sarebbero in grado di spiegare il comportamento osservato. Mc Crae e Costa (1987) postulano la teoria denominata "Big Five" che presenta i tratti della personalità raggruppati in 5 parti:

1. ESTROVERSIONE-INTROVERSIONE: il polo positivo di questo fattore è rappresentato dall'emozionalità positiva e dalla socialità, mentre quello negativo è rappresentato dall'introversione, ossia dalla tendenza ad "esser presi" più dal proprio mondo interno che da quello esterno.
2. AMICALITÀ-DISTACCO: il polo positivo di questo fattore è rappresentato da cortesia, altruismo e cooperatività; il polo negativo da ostilità, insensibilità ed indifferenza.
3. COSCIENZIOSITÀ-NEGLIGENZA: questo fattore contiene nel suo polo positivo gli aggettivi che fanno riferimento alla scrupolosità, alla perseveranza, alla affidabilità ed alla autodisciplina e, nel suo polo negativo, gli aggettivi opposti.
4. STABILITÀ EMOTIVA-INSTABILITÀ EMOTIVA (NEVROTICISMO): il polo positivo di questo fattore è rappresentato da vulnerabilità, insicurezza ed instabilità emotiva. Il polo opposto è rappresentato dalla stabilità emotiva, dalla dominanza degli stati d'ansia e dalla sicurezza.
5. APERTURA MENTALE-CHIUSURA MENTALE: il polo positivo di questo fattore è rappresentato da creatività, anticonformismo ed originalità. Il polo opposto invece è identificato dalla chiusura all'esperienza, ossia dal conformismo e dalla mancanza di creatività ed originalità.

Tali dimensioni sono state individuate a partire da studi psicolessicali, secondo i quali l'uomo ha codificato in forma verbale tutte le esperienze significative per la comunità, comprese, in questo caso, parole che si riferiscono alle differenze individuali: le 5 dimensioni elencate, quindi, corrisponderebbero alle macro-categorie più usate, nel linguaggio, per descrivere le diversità tra individui.

2.1.2 Validazione

Seguendo le linee di Mc Crae e Costa (1987), Caprara et al. (1994) svilupparono un nuovo questionario per supportare quello che esisteva precedentemente, che era denominato "Five Factor Model of Personality" (FFM).

Successivamente è stato creato un nuovo questionario che seguiva la linea del FFM, denominato Big Five (BF). Lo scopo di questo nuovo questionario è di essere più parsimonioso nel numero di sfaccettature delle personalità che si riferivano ad ognuna delle 5 dimensioni e di essere il più possibile coerente con quanto espresso nella teoria del FFM.

Per fare la validazione, è stato selezionato un campione di circa 1000 italiani (Caprara et al., 1994). L'affidabilità delle sfaccettature delle 5 dimensioni è stata valutata tramite il coefficiente Alpha di Cronbach (Caprara et al., 1994), il cui range tra le scale trovato è risultato tra 0.60 e 0.90. [Il coefficiente Alpha di Cronbach descrive la coerenza interna dei raggruppamenti degli item in una scala Likert. Il valore soglia per accettare la scala in termini di coerenza interna e di adeguatezza di costrutto del test costruito è 0.7].

2.1.3 La scala nel questionario

Nel questionario analizzato in questo studio, per ognuna delle 5 dimensioni della personalità sono state individuate 4 domande, la cui media offre una valutazione riguardante quella dimensione. Le domande sono le seguenti:

- 1) **EXTRAVERSION** (estroversione)
 - a) H8: metto molta energia nelle cose che faccio
 - b) H10: generalmente riesco a ben argomentare le mie ragioni e a convincere gli altri della loro validità
 - c) H11: tendo a sostenere con energia i miei punti di vista
 - d) H13: sono una persona attiva e vigorosa.

- 2) **AGREEBLENESS** (gradevolezza, amicalità)
 - a) H6: non esito a dare un aiuto a chi ne ha bisogno, anche se si tratta di persone che conosco appena
 - b) H9: mi sforzo sempre di venire incontro alle esigenze altrui

- c) H12: tendo a comportarmi in maniera cordiale con gli altri
 - d) H16: sono molto sensibile alle difficoltà altrui.
- 3) CONTIOUS (coscienziosità)
- a) H14: sono molto scrupoloso/a
 - b) H15: gli ambienti disordinati mi infastidiscono
 - c) H18: cerco di organizzare la mia vita in modo sistematico
 - d) H19: sono una persona metodica e ordinata.
- 4) EMOTIONAL STABILITY (stabilità emotiva)
- a) H3: riesco a dominare gli stati d'ansia
 - b) H5: il mio umore è piuttosto stabile
 - c) H7: mi capita raramente di entrare in agitazione per qualcosa
 - d) H17: sono una persona rilassata.
- 5) OPENESS (apertura mentale)
- a) H1: mi attraggono i programmi di informazione culturale e/o scientifica
 - b) H2: mi interessa conoscere i modi di vita e le usanze di paesi e popoli diversi dal nostro
 - c) H4: dedico molto tempo alla lettura
 - d) H20: mi piace acquisire conoscenze anche in ambiti che non sono strettamente attinenti ai miei interessi di studio e di lavoro.

A ognuna di queste domande il ragazzo/a risponde con una valutazione da 1 a 5, in cui: 1=assolutamente vero per me, 2=abbastanza vero per me, 3=né vero né falso per me, 4=piuttosto falso per me, 5=assolutamente falso per me. Dunque punteggi bassi dati alle domande del questionario vanno a supportare una personalità che tende al polo positivo di ognuna delle 5 dimensioni (estroversione, gradevolezza, coscienziosità, stabilità emotiva, apertura mentale), e punteggi alti una tendenza verso il polo negativo (introversione, sgradevolezza, negligenza, nevroticismo, chiusura mentale).

Per quanto riguarda la scala di Personalità-Stabilità Emotiva, quindi, valori bassi indicano: capacità di dominanza degli stati d'ansia, sicurezza, stabilità emotiva, rilassatezza; valori alti invece indicano: vulnerabilità, insicurezza, instabilità emotiva, agitazione.

2.2 Scala di Benessere Emotivo

2.2.1 Descrizione

La scala di Benessere Emotivo estratta dal questionario, ha lo scopo di fornire una misurazione relativa al benessere della persona. Si basa su degli indici descritti da Furlong et al. (2014). Questi indici nascono dal concetto di "Covitality", cioè un indicatore capace di misurare il benessere di un individuo. Inoltre si parla del SEHS ("Social and Emotional Health Survey" = "Questionario sul benessere sociale ed emotivo"), usato tradizionalmente

per la misurazione della salute mentale dei giovani e adolescenti. È un modello unidimensionale, cioè studia solamente l'indicatore positivo della salute mentale, in quanto trova una corrispondenza tra la diminuzione dei sintomi emotivi e comportamentali dei giovani e l'aumento del benessere e della qualità della loro vita. Tra il 2008 e il 2011, grazie a degli studi di Shaffer et al. (2010), si è passati a un modello bidimensionale, cioè un modello che considera congiuntamente la valutazione del malessere e del benessere per dare un significato più preciso del benessere della persona. È quindi un modello che presenta un indicatore sia positivo che negativo della salute mentale. L'applicazione di questo secondo modello al contesto scolastico risulta più efficace, poiché sia la presenza di malessere psicologico che l'assenza di benessere psicologico sono associate all'abbassamento della performance scolastica. Infatti, a supporto di questo, Furlong et al. (2014) affermano che il benessere di un individuo è determinato da altri fattori esterni (oltre che alle caratteristiche personali): famiglia, scuola, rapporto con i coetanei.

Sono presenti due modelli principali di questionario che seguono la linea dello studio bidimensionale della salute mentale della persona per la previsione dell'andamento scolastico. Il primo si chiama "40 developmental assets model". La sua struttura è formata dalla valutazione di un insieme di attività sia interne (famiglia, relazioni e scuola) che esterne (motivazione per raggiungere i risultati, impegno a scuola) che, studiate insieme, descrivono il benessere mentale del ragazzo/a. Il secondo modello, che sarà quello sfruttato nello studio descritto successivamente, è chiamato "Covitality" ed è formato dall'effetto sinergico di 4 concetti di primo ordine, derivanti a loro volta da 12 indicatori (o concetti di secondo ordine). I 12 concetti di secondo ordine sono: Self-efficacy (consapevolezza di essere efficaci e adeguati), Self-awareness (consapevolezza di sé), Persistence (persistenza, tenacia), School support (appoggio trovato nel contest scolastico), Family coherence (coerenza, compattezza familiare), Peer support (supporto tra coetanei), Emotional regulation (regolazione emotiva), Empathy (empatia, comprensione dell'altro), Behavioral self-control (autocontrollo), Optimism (ottimismo), Gratitude (gratitudine), Zest (entusiasmo).

I 4 concetti di primo ordine che si ricavano da questi indicatori sono:

- 1) Belief-in-self, formato da "Self-efficacy", "Self-awareness" e "Persistence";
- 2) Belief-in-others, formato da "School support", "Family coherence" e "Peer support";
- 3) Emotional competence, formato da "Emotional regulation", "Empathy" e "Behavioral self-control";
- 4) Engaged living, formato da "Optimism", "Gratitude" e "Zest".

Il "Covitality" può essere considerato quindi come predittore del benessere, come dimostrato dagli studi di Furlong et al. (2014). Nel caso preso in esame nello studio oggetto di queste analisi, si tratta del benessere degli studenti a cui viene sottoposto il questionario.

Per quanto riguarda la scala di Benessere, in questa tesi ci si concentrerà sul concetto di primo ordine "Emotional competence", che corrisponde alla scala di Benessere Emotivo.

2.2.2 Validazione

La validazione è stata fatta nell'articolo di Furlong et al. (2014).

Il campione utilizzato consiste in studenti adolescenti (età media: 15,1 anni) di 12 scuole della California Centrale. Il numero di questionari ritenuti validi alla fine della somministrazione, avvenuta nelle classi delle scuole con la sorveglianza di un insegnante o un coordinatore, sono stati 4109 (che corrisponde quindi alla numerosità del campione).

È stato misurato per ogni studente il valore di "Covitality", cioè come costrutto formato dalle 12 sottoscale tramite l'operatore somma. Successivamente è stata studiata la sua distribuzione nell'intero campione attraverso una standardizzazione. Usando i dati standardizzati sono stati ricavati 4 livelli di "Covitality": molto bassa (per $z < -1$, $n=623$), bassa (per z tra -1 e 0 , $n=1310$), alta (per z tra 0 e 1 , $n=1523$), molto alta (per $z > 1$, $n=653$). Il "Covitality" è stato misurato distintamente all'interno di queste classi (considerate come variabili indipendenti). Inoltre è stata effettuata un'analisi inferenziale tra questi 4 gruppi e 4 variabili dipendenti che formano il "Covitality": 1) Risultati scolastici; 2) Percezione della sicurezza a scuola; 3) Uso di sostanze; 4) Sintomi depressivi.

Le analisi confermano la presenza di differenze significative tra i vari gruppi, e quindi una relazione lineare tra i risultati ottenuti nelle 4 variabili che formano il "Covitality" e il livello di quest'ultimo.

Si conferma infatti la presenza di relazione lineare tra livello di "Covitality" e le variabili risultati scolastici, percezione della sicurezza a scuola, uso di sostanze e sintomi depressivi.

In conclusione, grazie a questo studio possiamo affermare che:

- ognuna delle 12 sottoscale può essere considerata come un costrutto distinto che descrive il benessere della persona;
- il concetto di "Covitality" è in relazione con alcune variabili esterne (che poi interesseranno la ricerca svolta successivamente), tra cui il comportamento/rendimento a scuola e i sintomi depressivi.

Dunque il "Covitality" è un buon indicatore del benessere.

Queste conclusioni andranno a supportare e giustificare la formazione della scala sul benessere utilizzata nel questionario considerato per l'analisi descritta in seguito.

2.2.3 La scala nel questionario

Nel questionario sottoposto agli studenti per la dipendenza da Internet e Facebook sono presenti delle domande che rimandano ai concetti espressi dai 4 costrutti primari di questa scala (Belief in self, Belief in others, Emotional Competence ed Engaged Living). Le domande sono le seguenti (analizzate distintamente per ognuno dei 4 costrutti primari e articolato nei 12 indicatori secondari):

- 1) BELIEF-IN-SELF, come media di:
 - a) SELF-EFFICACY, formato dalle seguenti domande del questionario:

- a) IA1: sono in grado di risolvere i miei problemi
 - b) IA2: riesco a fare la maggior parte delle cose se ci provo
 - c) IA3: ci sono molte cose che faccio bene;
 - b) SELF-AWARENESS, formato dalle seguenti domande del questionario:
 - a) IB1: la mia vita ha uno scopo
 - b) IB2: riconosco i miei sentimenti e il mio umore
 - c) IB3: capisco perché faccio quello che faccio;
 - c) PERSISTENCE, formato dalle seguenti domande del questionario:
 - a) IC1: quando non capisco qualcosa, chiedo chiarimenti all'insegnante finché non capisco
 - b) IC2: provo a rispondere a tutte le domande fatte in classe
 - c) IC3: quando cerco di risolvere un problema, non mi fermo finché non trovo una soluzione definitiva.
- 2) BELIEF-IN-OTHERS, come media di:
- a) SCHOOL SUPPORT, formato dalle seguenti domande del questionario:
 - a) ID1: nella mia scuola c'è un insegnante o un altro adulto ...che vuole sempre che io faccia del mio meglio
 - b) ID2: ...che mi ascolta quando ho qualcosa da dire
 - c) ID3: ...che crede che avrò successo;
 - b) FAMILY COHERENCE, formato dalle seguenti domande del questionario:
 - a) IE1: i membri della mia famiglia si aiutano e sostengono davvero a vicenda
 - b) IE2: c'è sentimento di unione nella mia famiglia
 - c) IE3: la mia famiglia va molto d'accordo;
 - c) PEER SUPPORT, formato dalle seguenti domande del questionario:
 - a) IF1: ho un amico/a della mia età a cui importa davvero di me
 - b) IF2: ho un amico/a della mia età con cui parlare dei miei problemi
 - c) IF3: ho un amico/a della mia età che mi aiuta quando ho un periodo difficile.
- 3) EMOTIONAL COMPETENCE, come media di:
- a) EMOTIONAL REGULATION, formato dalle seguenti domande del questionario:
 - a) IG1: mi assumo la responsabilità per le mie azioni
 - b) IG2: ammetto quando faccio un errore
 - c) IG3: sopporto quando mi dicono di no;
 - b) EMPATHY, formato dalle seguenti domande del questionario:
 - a) IH1: mi dispiace quando qualcuno si sente ferito
 - b) IH2: provo a comprendere cosa stanno passando le altre persone
 - c) IH3: cerco di capire cosa sentono e pensano le altre persone;
 - c) BEHAVIORAL SELF-CONTROL, formato dalle seguenti domande del questionario:
 - a) II1: posso aspettare per quello che voglio
 - b) II2: non disturbo gli altri quando sono impegnati

c) II3: penso prima di agire.

4) ENGAGED LIVING, come media di:

a) OPTIMISM, formato dalle seguenti domande del questionario:

a) IJ1: ogni giorno non vedo l'ora di divertirmi un sacco

b) IJ2: di solito mi aspetto di aver una buona giornata

c) IJ3: in generale, mi aspetto che succedano più cose positive che negative;

b) GRATITUDE, formato dalle seguenti domande del questionario:

a) IK1: da ieri ho provato ...gratitudine

b) IK2: ...riconoscenza

c) IK3: ...apprezzamento;

c) ZEST, formato dalle seguenti domande del questionario:

a) IL1: in questo momento provo la sensazione: ...energico/a

b) IL2: ...attivo/a

c) IL3: ...vivace.

La parte della scala che viene considerata per questa tesi è il costrutto "Emotional competence" (che corrisponde al nome assegnato alla scala: Benessere Emotivo).

Per quanto riguarda l'indicatore di Regolazione Emotiva ("Emtional Regulation") si valuta: la responsabilità delle proprie azioni, l'ammettere quando si fa un errore, la sopportazione dei "no"; quindi la capacità di avere un comportamento responsabile e maturità emotiva.

Per quanto riguarda l'indicatore Empatia ("Empathy") si valuta: il dispiacersi del dolore degli altri, il comprendere gli altri, il comprendere i sentimenti degli altri; quindi la capacità di entrare in empatia con gli altri, comprendendoli e mettendosi nei loro panni.

Per quanto riguarda l'indicatore Autocontrollo ("Behavioral self-control") si valuta: l'aspettare quello che si vuole, il non disturbare gli altri quando sono impegnati, il cercare di comprendere cosa vogliono le altre persone; quindi la capacità di sapersi moderare e saper pazientare.

Anche queste domande sono valutate attraverso una scala Likert, con i valori che vanno da 1 a 4, in cui 1=per niente, 2=un po', 3=abbastanza, 4=molto. Quindi, punteggi bassi corrispondono al polo negativo della personalità, e punteggi alti al polo positivo (esattamente il contrario di come si valuta la scala di Personalità-Stabilità Emotiva).

2.3 Analisi descrittive: confronto tra scala di Personalità-Stabilità Emotiva e scala di Benessere Emotivo

Sono state eseguite alcune analisi descrittive per valutare media, varianza, range di valori e distribuzioni delle scale di Personalità-Stabilità Emotiva e di Benessere Emotivo.

	Scala Personalità- Stabilità Emotiva	Scala Benessere Emotivo
Media	3.048	3.164
Mediana	3.000	3.222
Varianza	0.684	0.177
Deviazione standard	0.827	0.421
Coefficiente di variazione	0.271	0.133

Tabella 1: media, mediana, varianza, deviazione standard e coefficiente di variazione delle scale di Personalità-Stabilità Emotiva e di Benessere Emotivo

	Regolazione Emotiva	Empatia	Autocontrollo
Media	3.083	3.444	2.966
Mediana	3.000	3.667	3.000
Varianza	0.290	0.352	0.352
Deviazione standard	0.539	0.593	0.593
Coefficiente di Variazione	0.175	0.172	0.200

Tabella 2: media, mediana, varianza, deviazione standard e coefficiente di variazione delle sottoscale della scala di Benessere Emotivo

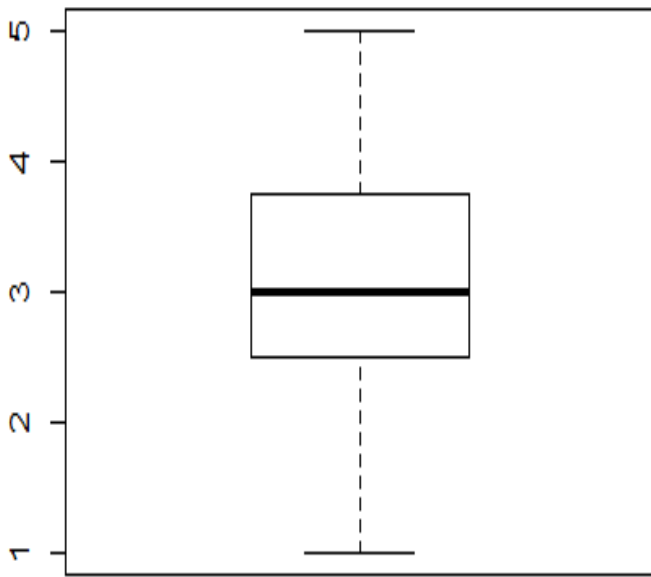


Figura 1: Boxplot scala di Personalità-Stabilità Emotiva

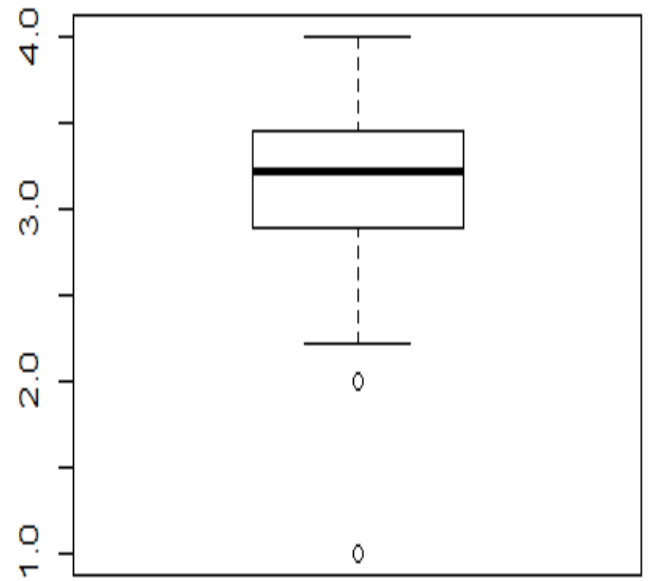


Figura 2: Boxplot scala di Benessere Emotivo

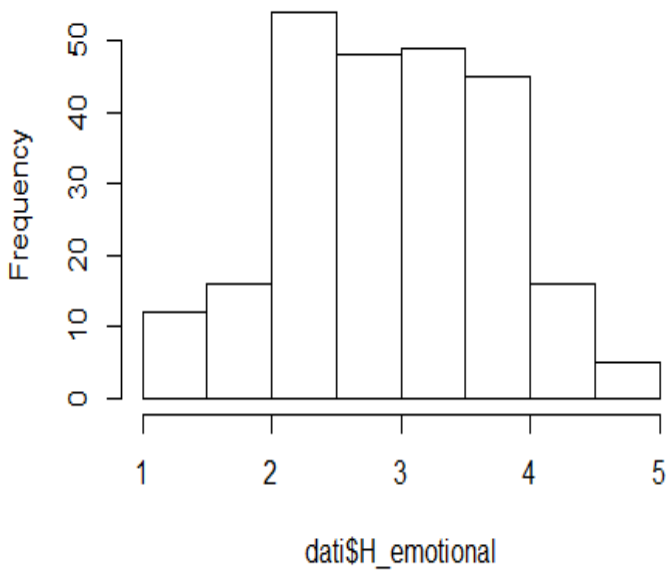


Figura 3: Istogramma scala di Personalità-Stabilità Emotiva

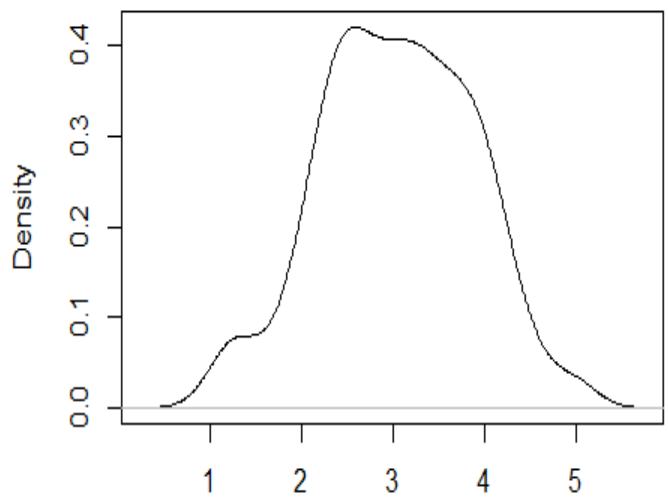


Figura 4: Densità distribuzione scala di Personalità-Stabilità Emotiva

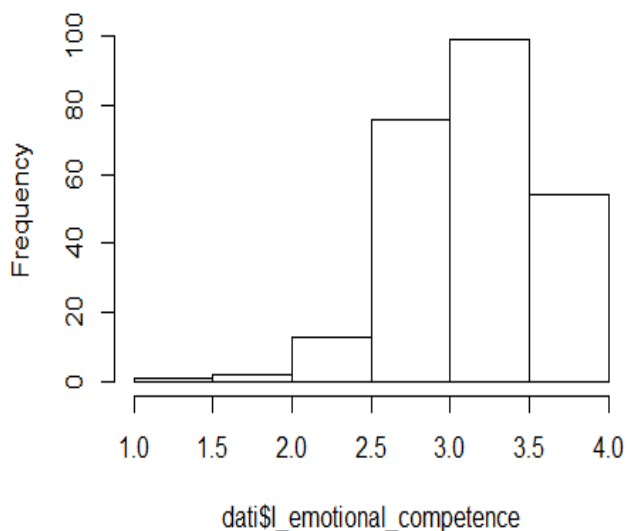


Figura 5: Istogramma scala di Benessere Emotivo

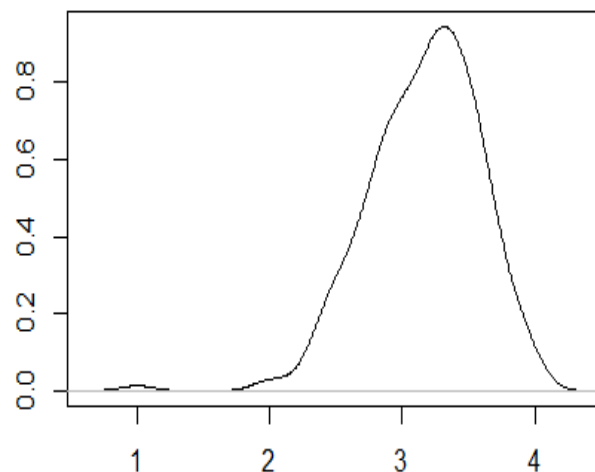


Figura 6: Densità distribuzione scala di Benessere Emotivo

Dall'analisi di media, mediana e varianza delle scale (tabella 1 e 2), possiamo notare che la scala di Personalità-Stabilità Emotiva presenta più variabilità della scala di Benessere Emotivo. Questo si valuta tramite i valori del coefficiente di variazione: infatti il coefficiente di variazione per la scala di Personalità-Stabilità Emotiva vale 0.271 ed è maggiore del coefficiente di variazione della scala di Benessere Emotivo, che vale 0.133. Quindi i valori osservati per la scala di Personalità-Stabilità Emotiva sono più dispersi attorno alla media rispetto ai valori osservati per la scala di Benessere Emotivo.

Questo può essere spiegato dal fatto che i valori della scala di Personalità-Stabilità Emotiva possono coprire tutto il range di possibili valori (quindi da 1 a 5), mentre la scala di Benessere Emotivo è la media di tre valori compresi tra 1 e 4, quindi il range finale di possibili risposte è più ristretto.

Possiamo trovare la conferma di questo anche dall'analisi degli istogrammi delle due scale: l'istogramma della scala di Benessere Emotivo (figura 5), infatti, presenta le maggiori frequenze intorno alla classe 3-3.5.

Media e mediana della scala di Benessere Emotivo sono superiori rispetto alla scala di Personalità-Stabilità Emotiva.

Dato che per la scala di Benessere Emotivo valori alti corrispondono a punteggi positivi, notiamo che la media dei punteggi dati a questa scala si concentrano in valori positivi. Infatti il picco della curva di densità della distribuzione delle medie dei punteggi dati alla scala di Benessere Emotivo si posiziona nella parte destra del grafico, e la distribuzione copre essenzialmente i valori da 2 a 4, con una leggera coda sinistra allungata (troviamo conferma nella presenza di valori anomali, cioè che stanno al di fuori dei baffi del boxplot).

Invece il picco della curva di densità di distribuzione della scala di Personalità-Stabilità Emotiva è centrale nel grafico e la distribuzione copre essenzialmente i valori tra 1 e 5. Analizzando le sottoscale della scala di Benessere Emotivo (quindi Regolazione Emotiva, Empatia e Autocontrollo), possiamo notare che i valori medi sono un po' eterogenei tra loro.

Capitolo 3

Libreria myFBr e costruzione del dataset

3.1 Libreria myFBr

MyFBr è una libreria creata da Davide Meneghetti, Simone Righetto e dal Prof. Livio Finos nel 2015 (Meneghetti et al., 2015): consiste in un pacchetto che consente di leggere e visualizzare le informazioni dei profili Facebook.

MyFBr contiene infatti una serie di funzioni che permettono di estrarre le diverse informazioni presenti sui profili Facebook di ciascun utente.

Queste funzioni sono:

getAnagrafica, getDataDownload, getDataReg, getDati, getEvents, getFriends, getInfo, getMessages, getNAccessi, getNPhotos, getNPost, getPlaces, getSex, getStati, getWall, getWallPosts.

Per l'analisi descritta successivamente, è stata presa in considerazione la funzione getMessages per estrarre tutti i messaggi ricevuti e inviati da ogni profilo. In particolare, la funzione getMessages estrae le seguenti variabili:

\$user: il nome degli utenti che hanno scritto il messaggio;

\$time: il giorno e l'ora in cui è stato inviato il messaggio;

\$text: il testo del messaggio;

\$thread: il numero di user coinvolti nella chat in cui è stato inviato quel messaggio.

3.2 Presentazione dati

All'inizio del progetto si aveva a disposizione un dataset che comprendeva le informazioni base di ciascuno dei profili Facebook degli studenti che, dopo aver compilato il questionario, hanno dato la disponibilità a consultare i loro dati di questo social network.

Si aveva quindi a disposizione un dataset formato da 245 osservazioni, corrispondenti a 245 diversi profili Facebook, e 32 variabili.

È stato effettuato un lavoro di pulizia dati che prevedeva la codifica di alcune variabili quali la situazione sentimentale in un numero di classi ristretto (per esempio nel caso della situazione sentimentale si avevano 55 diverse modalità, e con la pulizia dei dati queste modalità sono state riassunte in 3 classi: impegnato, single e non dichiarato).

Successivamente, al dataset iniziale è stata unita una seconda parte di variabili, ricavate attraverso myFBr, dai messaggi ricevuti e inviati dall'utente del profilo Facebook:

- 1) il numero di emoticons e di alcune parole inerenti agli atteggiamenti spiegati dalle scale psicologiche di Personalità-Stabilità Emotiva e di Benessere Emotivo;
- 2) alcuni indicatori di posizione e di scala per monitorare l'attività/inattività (ad esempio la frequenza di utilizzo di Facebook in termini di ore e di giorni).

Il dataset finale è dunque composto da 236 unità statistiche e 260 variabili.

Alcune unità (precisamente 9) sono state perse durante il processo di elaborazione dei profili Facebook per ricercare, tramite funzioni create appositamente, i conteggi delle parole selezionate per lo studio. Oltre alla perdita di queste unità nel dataset finale sono presenti alcuni valori mancanti (precisamente 3105 NA nell'intero dataset, che corrisponde circa al 5% del totale dei dati). Questi NA sono concentrati nelle variabili che riguardano la frequenza di attività e inattività in Facebook.

Le 260 variabili si possono suddividere in 3 categorie: variabili derivanti dall'estrazione delle informazioni dai profili Facebook di ciascun utente, le variabili provenienti dalle risposte date alle domande al questionario (riguardanti esclusivamente le scale di Personalità-Stabilità Emotiva e di Benessere Emotivo) e le variabili derivanti dall'analisi dei testi dei messaggi inviati e ricevuti su Facebook da ciascun utente.

Di seguito sono analizzati più nello specifico questi gruppi di variabili.

- Informazioni dai profili Facebook:
 - sesso
 - data di nascita
 - situazione sentimentale
 - richieste di amicizia accettate
 - richieste di amicizia effettuate
 - richieste di amicizia rifiutate
 - amici rimossi
 - numero totale di amici
 - numero di stati pubblicati in bacheca
 - numero di "mi piace" inseriti
 - numero di condivisioni in bacheca
 - accesso a link di altri utenti
 - accesso a link propri
 - numero di eventi a cui l'utente parteciperà
 - numero di post pubblicati in bacheca
 - numero di volte in cui è stata aggiornata l'immagine del profilo
 - numero di foto pubblicate
 - numero di post totali
 - numero di eventi confermati
 - numero di eventi in "forse"
 - numero di eventi rifiutati
 - numero di eventi a cui non si è data risposta
 - numero di foto totali
 - numero di album totali
 - numero di posizioni segnalate
 - numero di messaggi totali.
- Scale psicologiche: formato dalle medie dei punteggi dati al questionario:
 - scala di Personalità-Stabilità Emotiva ("H_emotional")
 - scala di Benessere Emotivo ("I_emotional_competence")
 - sottoscala della scala di Benessere Emotivo: Regolazione Emotiva ("I_emotional_regulation")

- sottoscala della scala di Benessere Emotivo: Empatia (“I_empathy”)
- sottoscala della scala di Benessere Emotivo: Autocontrollo (“I_behavioral_self_control”).
- Analisi del testo: insieme delle variabili recuperate tramite il software R e la libreria myFBr:
 - conteggio delle emoticons scritte e ricevute nelle chat sia personali che di gruppo di Facebook: emoticons che esprimono felicità (“EMOTEGOOD” e “EMOTEWINK”), amore (“EMOTELOVE”), tristezza (“EMOTEBAD”), stupore (“EMOTESHOCK”), divertimento (“EMOTEAMAZE”).
 - conteggio delle parole inerenti alle scale psicologiche di Personalità-Stabilità Emotiva e Benessere Emotivo scritte e ricevute nelle chat sia personali che di gruppo di Facebook: “bello”, “sicuro”, “ragione”, “giusto”, “dispiace”, “ascolto”, “capisco”, “disturbo”, “dimmi”, “tu”, “bene”, “ok”, “non lo so”, “boh”, “bho”, “aiuto”, “agitato”, “ansia”, “scusa”, “male”;
 - variabili descrittive riguardanti la frequenza di attività in bacheca e la frequenza di ricezione e invio di messaggi sia giornaliera che settimanale.

Le variabili conteggio delle emoticons e delle parole sono suddivise in 4 gruppi:

- 1) conteggio delle parole scritte dall’utente del profilo Facebook in chat personali (la parte iniziale del nome di queste variabili è “user_personal.”);
- 2) conteggio delle parole scritte dall’utente del profilo Facebook in chat di gruppo (la parte iniziale del nome di queste variabili è “user_grp.”);
- 3) conteggio delle parole scritte dalle persone che inviano messaggi all’utente del profilo Facebook in chat personali (la parte iniziale del nome di queste variabili è “others_personal.”);
- 4) conteggio delle parole scritte dalle persone che inviano messaggi in chat di gruppo in cui è presente anche l’utente del profilo Facebook (la parte iniziale del nome di queste variabili è others_grp.”).

Con queste variabili sono stati creati dei modelli di regressione lineare in cui le variabili dipendenti sono rappresentate dalle scale psicologiche.

Questi modelli sono stati creati per valutare se insieme di variabili Facebook, quindi l’attività di un utente in Facebook per quanto riguarda la frequenza e la scrittura di determinate parole nelle chat, possono spiegare la loro personalità dal punto di vista emotivo e il loro benessere emotivo.

In seguito sono presentate alcune analisi descrittive di alcune variabili Facebook.

Le variabili selezionate sono le seguenti:

- numero di messaggi totali inviati e ricevuti (“mess”);
- conteggio delle volte in cui l’utente del profilo Facebook ha inserito in chat personali emoticons che esprimono felicità (“user_personal.EMOTEGOOD”);
- proporzione di inattività settimanale dell’utente nelle chat di Facebook (“Inattività messaggi settimanale”);

- conteggio delle volte in cui l'utente del profilo Facebook ha scritto la parola "bello" nelle chat personali ("user_personal.bello");
- conteggio delle volte in cui l'utente del profilo Facebook ha scritto la parola "scusa" nelle chat personali ("user_personal.scusa");
- conteggio delle volte in cui l'utente del profilo Facebook ha scritto la parola "ansia" nelle chat personali ("user_personal.ansia").

	mess	user_personal. EMOTEGOOD	Inattività messaggi settimanale
Media	5.182013e+03	115.953	0.212
Mediana	2.089500e+03	40.500	0.118
Varianza	2.286700e+08	35872.198	0.052
Deviazione standard	1.512184e+04	189.400	0.228
Coefficiente di variazione	2.918	1.633	1.075

Tabella 3: media, mediana, varianza, deviazione standard e coefficiente di variazione delle variabili Facebook:

numero di messaggi totali inviati e ricevuti ("mess"), conteggio delle volte in cui l'utente del profilo Facebook ha inserito in chat personali emoticons che esprimono felicità ("user_personal.EMOTEGOOD"), proporzione di inattività settimanale dell'utente nelle chat di Facebook ("inattività messaggi settimanale")

	user_personal. bello	user_personal. scusa	user_personal. ansia
Media	22.309	13.691	1.661
Mediana	8.000	6.000	0.000
Varianza	2353.576	732.036	12.948
Deviazione standard	48.514	27.056	3.598
Coefficiente di variazione	2.175	1.976	2.166

Tabella 4: media, mediana, varianza, deviazione standard e coefficiente di variazione delle variabili Facebook:

conteggio delle volte in cui l'utente del profilo Facebook ha scritto la parola "bello" nelle chat personali ("user_personal.bello"), conteggio delle volte in cui l'utente del profilo Facebook ha scritto la parola "scusa" nelle chat personali ("user_personal.scusa"), conteggio delle volte in cui l'utente del profilo Facebook ha scritto la parola "ansia" nelle chat personali ("user_personal.ansia").

Dalla tabella 3 possiamo notare che per la variabile numero totale di messaggi inviati e ricevuti (“mess”) si ha una variabilità molto elevata (il valore della varianza è $2.286700e+08$). Quindi ci sono dei profili Facebook in cui l’utente non scrive messaggi nelle chat ed altri che ne inviano un numero elevatissimo. Questo è spiegato dal valore del range: (0 : 205546), cioè esistono profili in cui sono stati scambiati più di 200000 messaggi nei 6 mesi di attività di cui si sono studiati i profili Facebook.

Anche per quanto riguarda le variabili che contano il numero di volte in cui viene scritta nelle chat una determinata parola (per esempio, dalla tabella 4, si possono vedere alcune statistiche descrittive del conteggio delle volte in cui l’utente del profilo scrive le parole “bello”, il cui nome della variabile è “user_personal.bello”, “scusa”, il cui nome della variabile è “user_personal.scusa” e “ansia”, il cui nome della variabile è “user_persona.ansia”) la variabilità è elevata. Infatti dalla tabella 4 possiamo notare che per esempio la varianza della variabile che conta il numero di volte in cui l’utente scrive in chat personali la parola “bello” vale 2353.576.

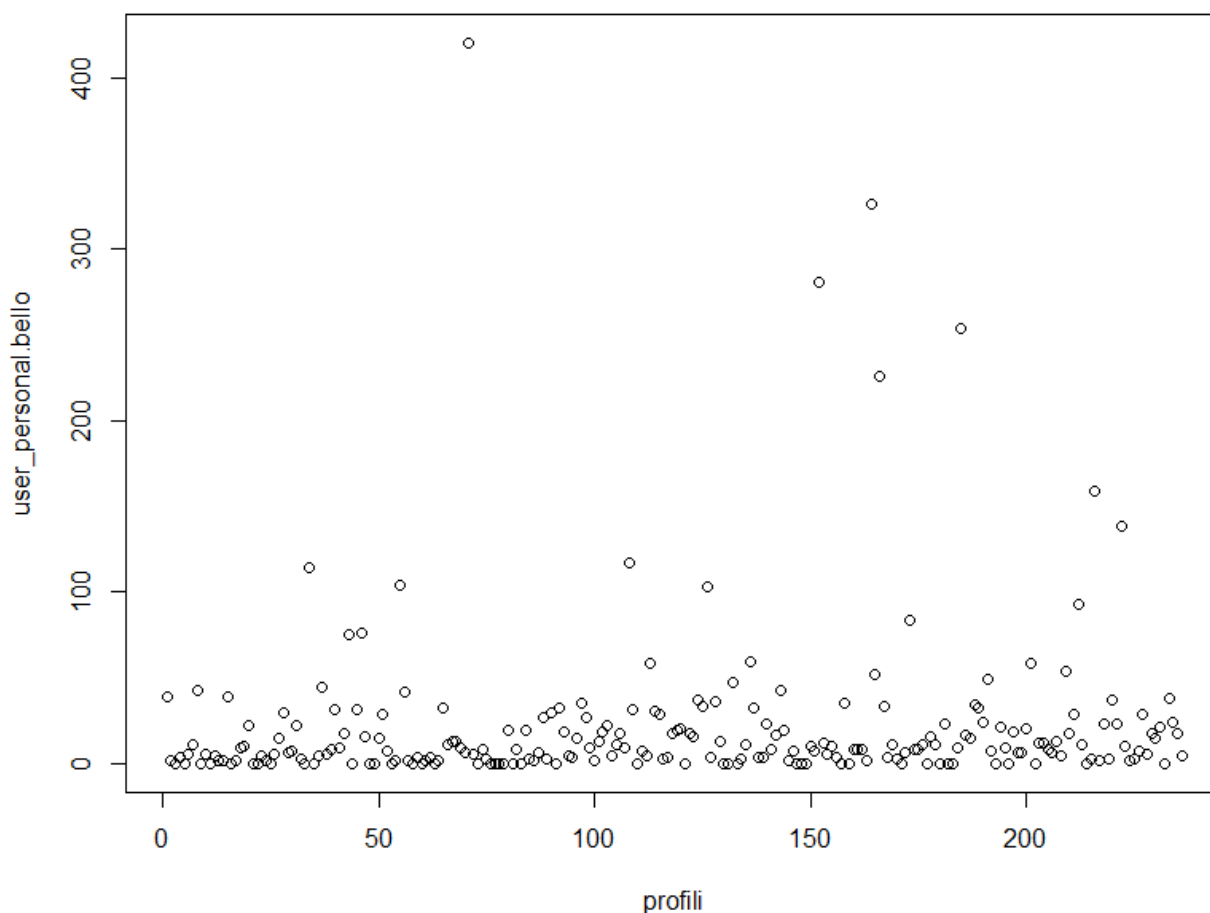


Figura 7: conteggio della parola "bello" scritta nei 236 profili Facebook in messaggi personali

La figura 7 mostra come sono distribuiti i valori della variabile conteggio della parola “bello” scritta nei 236 profili Facebook in chat personali. Si nota quindi un’alta variabilità. Alcuni utenti non hanno mai scritto questa parola, altri più di 300 volte.

Risultati simili valgono anche per le altre variabili che contano parole ed emoticons scritte nelle chat di Facebook.

Capitolo 4

Analisi e modelli

Per quanto riguarda le analisi, come variabili dipendenti sono state considerate la scala psicologica di Personalità nella parte riguardante la sfera emotiva (Personalità-Stabilità Emotiva) e la scala di Benessere Emotivo.

Per quanto riguarda la scala di Personalità-Stabilità Emotiva si ha a disposizione un valore corrispondente alla media delle seguenti 4 domande:

- a) H3: riesco a dominare gli stati d'ansia
- b) H5: il mio umore è piuttosto stabile
- c) H7: mi capita raramente di entrare in agitazione per qualcosa
- d) H17: sono una persona rilassata.

Per quanto riguarda la scala di Benessere Emotivo, si ha il valore della scala completa, cioè quella formata dalla media dei valori delle 3 sottoscale, e anche i valori delle 3 sottoscale considerate singolarmente.

Quindi, la scala complessiva Benessere Emotivo (“I_emotional_competence”) è formata dalla media delle seguenti sottoscale, formate a loro volta dalla media di 3 domande:

- a) Regolazione emotiva (“Emotional Regulation”), media delle seguenti domande del questionario:
 - a) IG1: mi assumo la responsabilità per le mie azioni
 - b) IG2: ammetto quando faccio un errore
 - c) IG3: sopporto quando mi dicono di no;
- b) Empatia (“Empathy”), media delle seguenti domande del questionario:
 - a) IH1: mi dispiace quando qualcuno si sente ferito
 - b) IH2: provo a comprendere cosa stanno passando le altre persone
 - c) IH3: cerco di capire cosa sentono e pensano le altre persone;
- c) Autocontrollo (“Behavioral self-control”), media delle seguenti domande del questionario:
 - a) II1: posso aspettare per quello che voglio
 - b) II2: non disturbo gli altri quando sono impegnati
 - c) II3: penso prima di agire.

Con questo dataset a disposizione sono stati eseguiti dei modelli lineari, con il software R, per vedere quanto le variabili Facebook potessero spiegare le scale psicologiche di Personalità-Stabilità Emotiva e di Benessere Emotivo, considerate quindi variabili dipendenti.

È stato individuato il modello migliore tramite il metodo della stepwise regression, cioè un metodo che, a partire dalle variabili che si inseriscono in un modello, sceglie la combinazione migliore di queste tramite il valore dell’AIC, cioè il criterio di informazione di Akaike. Questo metodo fornisce una misura della qualità della stima di un modello statistico tenendo conto

sia della bontà di adattamento che della complessità del modello. È inoltre un metodo che permette di confrontare modelli anche non annidati.

La regola è quella di preferire i modelli con il valore dell'AIC più basso.

Successivamente è stata valutata la bontà del modello tramite il valore del coefficiente di determinazione corretto (R^2 corretto).

Si tratta di una variante del coefficiente di determinazione (R^2), usata per l'analisi di regressione lineare multipla. Questo indice misura la frazione di devianza spiegata, cioè la proporzione di variabilità della variabile dipendente (in questo caso le scale psicologiche) spiegata dalle variabili esplicative del modello (in questo caso le variabili Facebook), corretto da un termine di penalizzazione che tiene conto della numerosità campionaria e del numero di esplicative utilizzate.

Si calcola nel seguente modo:

$$R^2_{\text{corr}} = 1 - [(1 - R^2) * (n - 1) / (n - k - 1)]$$

dove k è il numero di variabili indipendenti presenti nel modello di regressione e n è la numerosità campionaria.

È quindi utile per effettuare confronti tra modelli con diverso numero di variabili esplicative.

4.1 Modelli scala di Personalità-Stabilità Emotiva

4.1.1 Modello con tutte le variabili

Per la scala della Personalità-Stabilità Emotiva è stata effettuata la stepwise regression a partire da un modello che comprendeva tutte le variabili descrittive dei profili Facebook, tutta la parte relativa al conteggio di parole e di emoticons e le variabili di frequenza di attività in Facebook. Sono state escluse le variabili del conteggio di parole/emoticons nei messaggi scritti da persone diverse dall'utente del profilo Facebook nelle chat di gruppo, poiché in questi casi non possiamo sapere se le parole scritte dall'altra persona siano rivolte all'utente del profilo Facebook, quindi non si ha la certezza che si riferiscano a conversazioni in cui è coinvolto l'utente del profilo.

In totale, sono state incluse 106 variabili.

Il modello ridotto a seguito della stepwise regression è il seguente:

```
modello 1 = lm(scala di Personalità-Stabilità Emotiva ~
  datiFine$ sesso +
  datiFine$ accettati +
  datiFine$ richEff +
  datiFine$ richRic +
  datiFine$ rimossi +
  datiFine$ piace +
  datiFine$ linkAltri +
```

```

datiFine$nuovaFoto +
datiFine$postTotali +
datiFine$eventiConf +
datiFine$eventiForse +
datiFine$nPhoto +
datiFine$nAlbum +
datiFine$nPosti +
datiFine$user_personal.EMOTEWINK +
datiFine$user_personal.EMOTEAMAZE.1 +
datiFine$user_personal.bello +
datiFine$user_personal.sicuro +
datiFine$user_personal.ragione +
datiFine$user_personal.dimmi +
datiFine$user_personal.tu +
datiFine$user_personal.bene +
datiFine$user_personal.NonLoSo +
datiFine$user_personal.bho +
datiFine$user_personal.agitat +
datiFine$user_grp.EMOTEGOOD +
datiFine$user_grp.EMOTESHOCK.1 +
datiFine$user_grp.EMOTEAMAZE.1 +
datiFine$user_grp.giusto +
datiFine$user_grp.capisco +
datiFine$user_grp.tu +
datiFine$user_grp.boh +
datiFine$user_grp.bho +
datiFine$user_grp.scusa +
datiFine$others_personal.EMOTEGOOD +
datiFine$others_personal.EMOTELOVE +
datiFine$others_personal.EMOTEWINK +
datiFine$others_personal.EMOTESHOCK.1 +
datiFine$others_personal.ragione +
datiFine$others_personal.giusto +
datiFine$others_personal.dispiace +
datiFine$others_personal.capisco +
datiFine$others_personal.disturbo +
datiFine$others_personal.NonLoSo +
datiFine$others_personal.boh +
datiFine$others_personal.ansia +
datiFine$mess_activity_time_summary.week_off.prop_inactive)

```

Sono state selezionate quindi 47 variabili tra le 106 variabili iniziali, che formavano il modello prima della stepwise regression.

Il valore stimato dei coefficienti, il valore del test t e il valore del p-value dei coefficienti di ciascuna variabile sono descritti nella seguente tabella:

	Nome delle variabili	Valore stimato	Test t	p-value
1	(Intercept)	3.143375e+00	17.845874	5.019599e-39
2	datiFine\$ sessoF	1.902409e-01	1.324301	1.874041e-01

3	datiFine\$accettati	-2.192883e-04	-1.475733	1.420978e-01
4	datiFine\$richEff	1.169750e-02	2.734827	6.988532e-03
5	datiFine\$richRic	9.765226e-04	1.573663	1.176583e-01
6	datiFine\$rimossi	7.521051e-04	1.754918	8.130150e-02
7	datiFine\$piace	1.189545e-03	2.320331	2.166192e-02
8	datiFine\$linkAltri	-4.556085e-02	-2.537630	1.217458e-02
9	datiFine\$nuovaFoto	-4.000536e-03	-1.747204	8.263503e-02
10	datiFine\$postTotali	-1.976919e-04	-2.042524	4.283911e-02
11	datiFine\$eventiConf	4.089557e-03	1.453310	1.482128e-01
12	datiFine\$eventiForse	-3.341871e-02	-3.003108	3.128826e-03
13	datiFine\$nPhoto	5.083797e-03	1.978398	4.970314e-02
14	datiFine\$nAlbum	-1.130417e-01	-1.503890	1.346983e-01
15	datiFine\$nPosti	6.973635e-01	1.573977	1.175856e-01
16	datiFine\$user_personal.EMOTEWINK	1.471651e-03	1.262444	2.087353e-01
17	datiFine\$user_personal.EMOTEAMAZE.1	-2.695558e-03	-2.097877	3.758151e-02
18	datiFine\$user_personal.bello	9.546240e-03	2.314441	2.199141e-02
19	datiFine\$user_personal.sicuro	-3.448084e-02	-3.248807	1.428475e-03
20	datiFine\$user_personal.ragione	-9.444223e-02	-2.275385	2.428989e-02
21	datiFine\$user_personal.dimmi	-2.917613e-02	-1.719403	8.759017e-02
22	datiFine\$user_personal.tu	1.420173e-03	1.977490	4.980667e-02
23	datiFine\$user_personal.bene	-6.207500e-03	-2.208126	2.874315e-02
24	datiFine\$user_personal.NonLoSo	1.111806e+00	1.877819	6.233429e-02
25	datiFine\$user_personal.bho	-5.611092e-02	-4.654171	7.064466e-06
26	datiFine\$user_personal.agitat	2.387770e-01	2.838023	5.164034e-03
27	datiFine\$user_grp.EMOTEGOOD	-2.651963e-02	-2.294752	2.312508e-02
28	datiFine\$user_grp.EMOTESHOCK.1	-4.786737e-01	-2.723180	7.227544e-03
29	datiFine\$user_grp.EMOTEAMAZE.1	1.488631e-01	3.214376	1.598730e-03
30	datiFine\$user_grp.giusto	4.605787e-01	3.334692	1.074546e-03
31	datiFine\$user_grp.capisco	-3.022780e+01	-3.786603	2.199002e-04
32	datiFine\$user_grp.tu	-1.540608e-02	-1.740305	8.384274e-02
33	datiFine\$user_grp.boh	-6.763660e-01	-3.709078	2.916936e-04
34	datiFine\$user_grp.bho	8.452407e-01	3.156333	1.929054e-03
35	datiFine\$user_grp.scusa	1.941427e-01	2.368538	1.912424e-02
36	datiFine\$others_personal.EMOTEGOOD	5.389025e-03	4.419061	1.882554e-05
37	datiFine\$others_personal.EMOTELOVE	-2.454022e-03	-3.881393	1.547828e-04
38	datiFine\$others_personal.EMOTEWINK	-7.065038e-03	-2.250504	2.586147e-02
39	datiFine\$others_personal.EMOTESHOCK.1	9.764623e-03	2.904927	4.225637e-03
40	datiFine\$others_personal.ragione	-9.399977e-02	-2.222360	2.774551e-02

41	datiFine\$others_personal.giusto	-5.232884e-02	-2.632136	9.366402e-03
42	datiFine\$others_personal.dispiace	6.827961e-02	4.990076	1.642483e-06
43	datiFine\$others_personal.capisco	1.459971e-01	2.070971	4.006354e-02
44	datiFine\$others_personal.disturbo	-4.923544e-02	-2.725413	7.181168e-03
45	datiFine\$others_personal.NonLoSo	-1.073770e+00	-3.832439	1.857041e-04
46	datiFine\$others_personal.boh	3.839397e-02	3.229546	1.521516e-03
47	datiFine\$others_personal.ansia	-1.112922e-01	-3.383194	9.127473e-04
48	datiFine\$mess_activity_time_summary.we	-6.070318e-01	-2.118612	3.575999e-02

Multiple R-squared: 0.5008, **Adjusted R-squared: 0.3455**
F-statistic: 3.223 on 47 and 151 DF, p-value: 3.3e-08

Il valore dell' R^2 corretto è 0.3455. Questo significa che la combinazione di variabili di questo modello riesce a spiegare il 34.55% della variabilità della variabile dipendente, cioè del valore della scala di Personalità-Stabilità Emotiva.

Analisi dei residui:

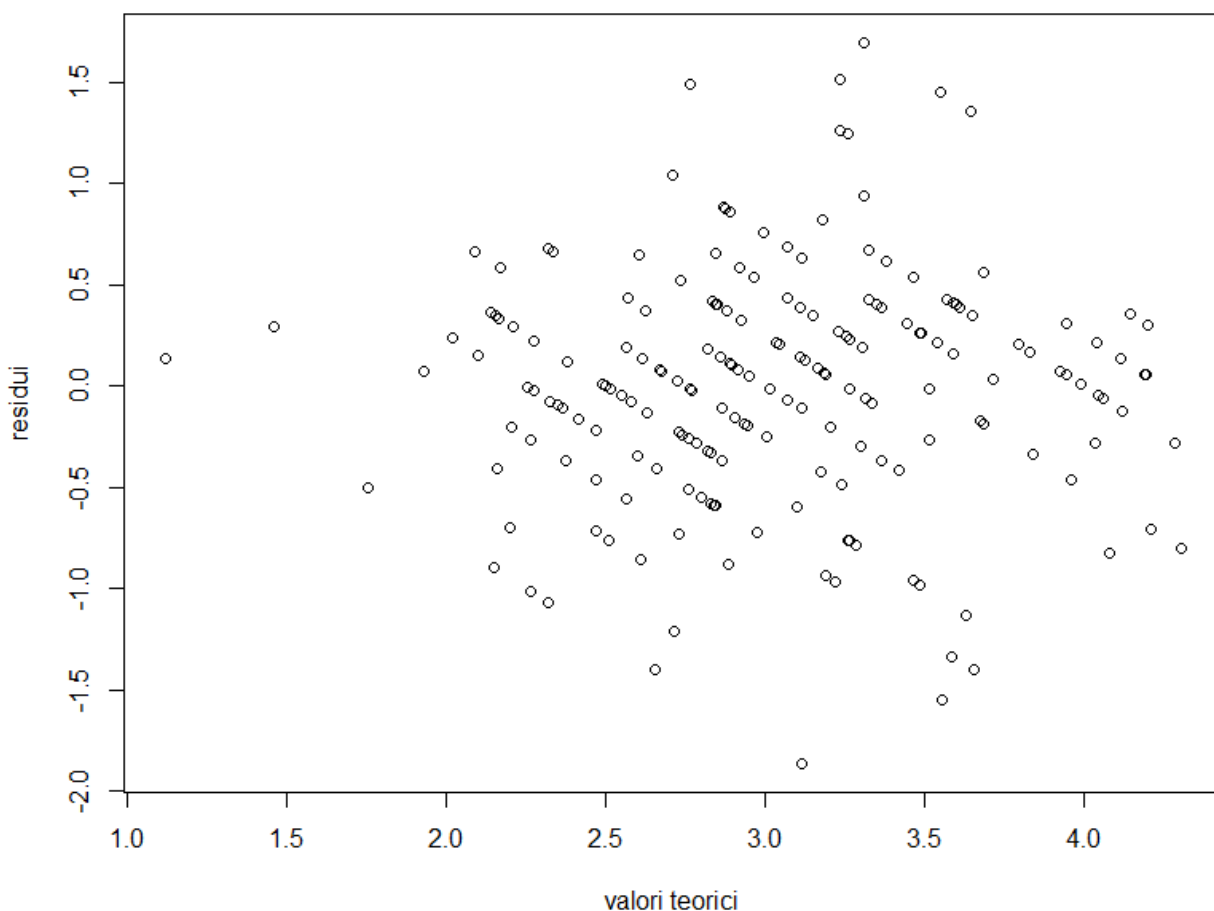


Figura 8: residui del modello per la scala di Personalità-Stabilità Emotiva con tutte le variabili

Questo grafico presenta in ascissa i valori teorici e in ordinata i residui di questa regressione. Eventuali strutture o trend presenti nel grafico indicano che i residui non sono casuali e quindi che i regressori non hanno colto tutta la variabilità della variabile risposta. In questo caso non si osservano particolari problemi, perchè i residui sembrano disperdersi in maniera casuale sopra e sotto la loro media (lo zero). Questo sottolinea il fatto che non c'è una tendenza nella distribuzione dei punti, quindi si esclude l'ipotesi che nel modello manchino variabili che spieghino l'andamento dei dati. Non si rilevano nemmeno molti valori anomali, infatti vediamo che non ci sono osservazioni che si discostano notevolmente dallo zero.

4.1.2 Modello con solo le variabili di analisi dei messaggi

Successivamente è stata effettuata un'altra analisi di stepwise regression da un modello con solamente le variabili conteggio di parole ed emoticons scritte dall'utente in chat personali e in chat di gruppo.

Questo per studiare la bontà di un modello in cui le scale psicologiche sono spiegate solamente da ciò che l'utente scrive nei messaggi.

Il modello trovato è il seguente:

```
modello 2 = lm(scala di Personalità-Stabilità Emotiva ~
  datiFine$user_personal.EMOTEGOOD +
  datiFine$user_personal.EMOTELOVE +
  datiFine$user_personal.EMOTEWINK +
  datiFine$user_personal.sicuro +
  datiFine$user_personal.ragione +
  datiFine$user_personal.boh +
  datiFine$user_personal.bho +
  datiFine$user_personal.aiuto +
  datiFine$user_personal.agitat +
  datiFine$user_personal.male +
  datiFine$user_grp.EMOTEWINK +
  datiFine$user_grp.sicuro +
  datiFine$user_grp.giusto +
  datiFine$user_grp.tu +
  datiFine$user_grp.bene)
```

Sono state selezionate 15 variabili tra le 52 variabili totali presenti prima di eseguire la stepwise regression.

Il valore stimato dei coefficienti, il valore del test t e il valore del p-value dei coefficienti di ciascuna variabile sono descritti nella seguente tabella:

	Nome delle variabili	Valore stimato	Test t	p-value
1	(Intercept)	3.000627186	48.149486	8.224932e-119
2	datiFine\$user_personal.EMOTEGOOD	0.001991096	4.200164	3.875387e-05
3	datiFine\$user_personal.EMOTELOVE	-0.003915127	-3.954143	1.035663e-04
4	datiFine\$user_personal.EMOTEWINK	-0.001829043	-1.524108	1.289180e-01
5	datiFine\$user_personal.sicuro	-0.021775637	-2.720474	7.040001e-03
6	datiFine\$user_personal.ragione	-0.074798930	-2.148460	3.276966e-02
7	datiFine\$user_personal.boh	0.012520128	1.789525	7.490567e-02
8	datiFine\$user_personal.bho	-0.018288350	-2.124535	3.474131e-02
9	datiFine\$user_personal.aiuto	-0.022808783	-1.388150	1.664954e-01
10	datiFine\$user_personal.agitat	0.122499555	1.835146	6.783413e-02
11	datiFine\$user_personal.male	0.010752947	2.609824	9.680935e-03
12	datiFine\$user_grp.EMOTEWINK	0.045524255	2.737493	6.697408e-03
13	datiFine\$user_grp.sicuro	0.367462917	3.975971	9.509142e-05
14	datiFine\$user_grp.giusto	0.248314498	2.694045	7.603508e-03
15	datiFine\$user_grp.tu	-0.023166205	-3.046946	2.594450e-03
16	datiFine\$user_grp.bene	-0.060656872	-2.108839	3.608951e-02

Multiple R-squared: 0.1975, **Adjusted R-squared: 0.1428**

F-statistic: 3.611 on 15 and 220 DF, p-value: 1.292e-05

Il valore dell' R^2 corretto è 0.1428. Questo significa che la combinazione di variabili di questo modello riesce a spiegare il 14.28% della variabilità della variabile dipendente, cioè del valore della scala di Personalità-Stabilità Emotiva.

Analisi dei residui:

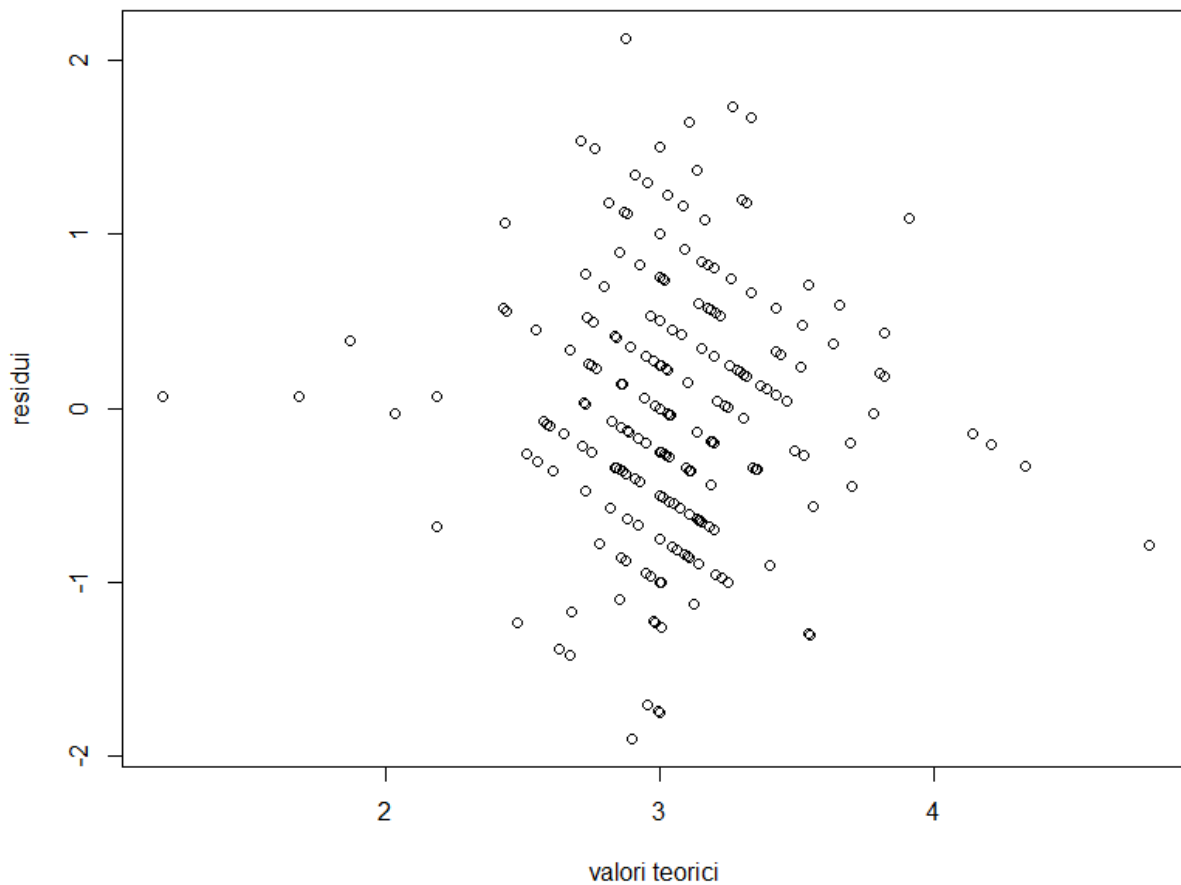


Figura 9: residui del modello per la scala di Personalità-Stabilità Emotiva con solo le variabili di analisi dei messaggi

Anche in questo caso non si osservano particolari problemi, perchè i residui sembrano disperdersi in maniera casuale sopra e sotto la loro media (lo zero).

Non si rilevano nemmeno molti valori anomali.

Quindi si esclude l'ipotesi che nel modello manchino variabili che spieghino l'andatura delle osservazioni.

4.2 Modelli scala di Benessere Emotivo

4.2.1 Modello con tutte le variabili

Per la scala di Benessere Emotivo è stato individuato inizialmente un modello tramite l'applicazione della stepwise regression a tutte le variabili descrittive dei profili Facebook, tutta la parte relativa al conteggio di parole e di emoticons e la frequenza di attività in Facebook, escludendo quelle scritte da altre persone (diverse dall'utente del profilo Facebook) in chat di gruppo, e le variabili che descrivono la percentuale di attività e inattività in Facebook.

Il modello trovato è il seguente:

```
modello 1 = lm(scala di Benessere Emotivo ~
  datiFine$sex +
  datiFine$richEff +
  datiFine$likes +
  datiFine$shared +
  datiFine$linkTuo +
  datiFine$partecipato +
  datiFine$nuovaFoto +
  datiFine$eventiConf +
  datiFine$eventiForse +
  datiFine$eventiNoRisp +
  datiFine$nPhoto +
  datiFine$nPosti +
  datiFine$user_personal.EMOTELOVE +
  datiFine$user_personal.EMOTEBAD +
  datiFine$user_personal.EMOTEAMAZE.1 +
  datiFine$user_personal.bello +
  datiFine$user_personal.sicuro +
  datiFine$user_personal.capisco +
  datiFine$user_personal.NonLoSo +
  datiFine$user_personal.bho +
  datiFine$user_personal.scusa +
  datiFine$user_personal.male +
  datiFine$user_grp.EMOTEWINK +
  datiFine$user_grp.giusto +
  datiFine$user_grp.bene +
  datiFine$user_grp.ansia +
  datiFine$others_personal.EMOTEBAD +
  datiFine$others_personal.ascolto +
  datiFine$others_personal.disturbo +
  datiFine$others_personal.bene +
  datiFine$others_personal.boh +
  datiFine$others_personal.scusa +
  datiFine$wall_activity_time_summary.days_off.prop_inactive)
```


Sono state selezionate 33 variabili tra le 106 variabili iniziali, precedenti all'analisi tramite la stepwise regression.

Il valore stimato dei coefficienti, il valore del test t e il valore del p-value dei coefficienti di ciascuna variabile sono descritti nella seguente tabella:

	Nome delle variabili	Valore stimato	Test t	p-value
1	(Intercept)	3.4145509825	24.519272	2.189632e-57
2	datiFine\$ sessoF	0.2171310304	3.317391	1.113804e-03
3	datiFine\$ richEff	-0.0059228458	-3.160048	1.871286e-03
4	datiFine\$ piace	0.0006708922	2.601398	1.011304e-02
5	datiFine\$ condiviso	0.0036085844	2.863307	4.726911e-03
6	datiFine\$ linkTuo	-0.0044713556	-1.437801	1.523507e-01
7	datiFine\$ partecipato	-0.0086641096	-1.854176	6.546711e-02
8	datiFine\$ nuovaFoto	-0.0036871462	-3.215207	1.563434e-03
9	datiFine\$ eventiConf	0.0095798413	1.921682	5.633842e-02
10	datiFine\$ eventiForse	0.0089724493	1.749879	8.196499e-02
11	datiFine\$ eventiNoRisp	-0.0007659373	-2.636724	9.156050e-03
12	datiFine\$ nPhoto	0.0028844651	2.454898	1.511165e-02
13	datiFine\$ nPosti	-0.2652937435	-1.320594	1.884328e-01
14	datiFine\$ user_personal.EMOTELOVE	-0.0092268885	-3.042057	2.727155e-03
15	datiFine\$ user_personal.EMOTEBAD	-0.0036212349	-3.544375	5.098516e-04
16	datiFine\$ user_personal.EMOTEAMAZE.1	0.0010346518	2.495855	1.352998e-02
17	datiFine\$ user_personal.bello	0.0050287296	2.937776	3.770315e-03
18	datiFine\$ user_personal.sicuro	0.0210282226	4.790619	3.634061e-06
19	datiFine\$ user_personal.capisco	0.0626614965	2.364044	1.921912e-02
20	datiFine\$ user_personal.NonLoSo	-0.7910183510	-3.609614	4.044122e-04
21	datiFine\$ user_personal.bho	0.0153355480	3.560093	4.823091e-04
22	datiFine\$ user_personal.scusa	0.0025355216	1.705831	8.988763e-02
23	datiFine\$ user_personal.male	-0.0090215472	-3.601753	4.159304e-04
24	datiFine\$ user_grp.EMOTEWINK	-0.0194450793	-3.124348	2.099531e-03
25	datiFine\$ user_grp.giusto	-0.1104860304	-2.533373	1.221243e-02
26	datiFine\$ user_grp.bene	0.0395126879	3.182115	1.741936e-03
27	datiFine\$ user_grp.ansia	-0.1490220448	-1.569438	1.184275e-01
28	datiFine\$ others_personal.EMOTEBAD	0.0039808859	3.236786	1.456353e-03
29	datiFine\$ others_personal.ascolto	-0.3321122160	-2.113838	3.600543e-02
30	datiFine\$ others_personal.disturbo	-0.0197469587	-2.126157	3.495101e-02
31	datiFine\$ others_personal.bene	-0.0048729535	-3.957555	1.115498e-04
32	datiFine\$ others_personal.boh	-0.0189073224	-3.619512	3.903366e-04

33	datiFine\$others_personal.scusa	0.0138511843	2.794098	5.809615e-03
34	datiFine\$wall_activity_time_summary.days_off. prop_inactive	-0.4977209116	-2.835134	5.143228e-03

Multiple R-squared: 0.4209, **Adjusted R-squared: 0.3071**
 F-statistic: 3.7 on 33 and 168 DF, p-value: 1.266e-08

Il valore dell' R^2 corretto è 0.3071. Questo significa che la combinazione di variabili di questo modello riesce a spiegare il 30.71% della variabilità della variabile dipendente, cioè del valore della scala di Benessere Emotivo.

Analisi dei residui:

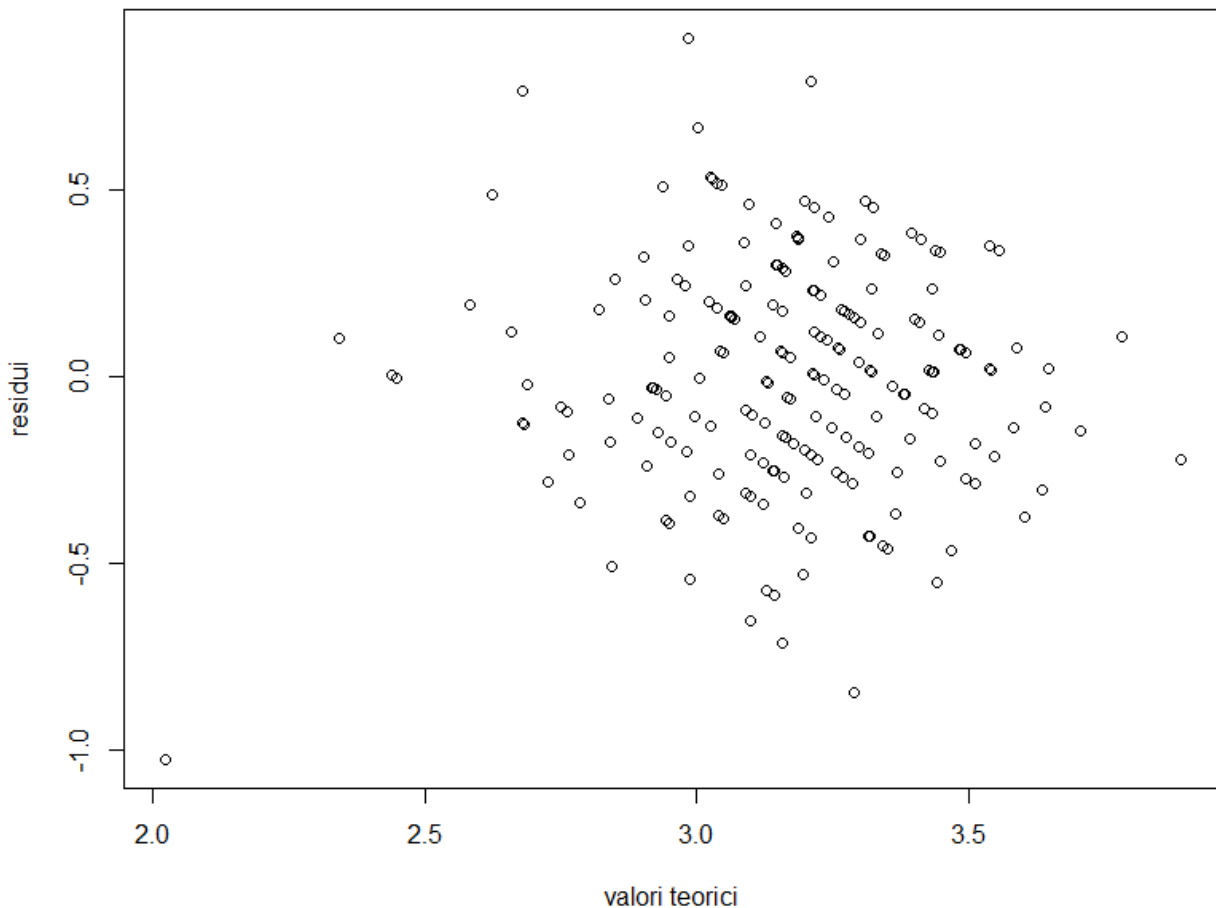


Figura 10: residui del modello per la scala di Benessere Emotivo con tutte le variabili

Da questo grafico possiamo notare che non ci sono particolari problemi, perchè i residui sembrano disperdersi in maniera casuale sopra e sotto la loro media (lo zero).

Non si rilevano nemmeno molti valori anomali.

Quindi, a causa dell'assenza di particolari andamenti o trend, si può affermare che nel modello non mancano variabili che spieghino l'andamento dei dati.

4.2.2 Modello con solo le variabili di analisi dei messaggi

Anche per questa scala è stata effettuata un'altra analisi di stepwise regression da un modello con solamente le variabili conteggio di parole ed emoticons scritte dall'utente in chat personali e in chat di gruppo.

Questo per studiare la bontà di un modello in cui le scale psicologiche sono spiegate solamente da ciò che l'utente scrive nei messaggi.

```
Modello 2 = lm (scala di Benessere Emotivo ~
    datiFine$user_personal.EMOTEBAD +
    datiFine$user_personal.sicuro +
    datiFine$user_personal.capisco +
    datiFine$user_personal.male +
    datiFine$user_grp.EMOTELove +
    datiFine$user_grp.EMOTESHOCK.1 +
    datiFine$user_grp.EMOTEAMAZE.1 +
    datiFine$user_grp.sicuro +
    datiFine$user_grp.aiuto)
```

Sono state selezionate 9 variabili tra le 52 considerate all'inizio, prima della stepwise regression.

Il valore stimato dei coefficienti, il valore del test t e il valore del p-value dei coefficienti di ciascuna variabile sono descritti nella seguente tabella:

	Nome delle variabili	Valore stimato	Test t	p-value
1	(Intercept)	3.137548353	100.073710	3.766906e-189
2	datiFine\$user_personal.EMOTEBAD	-0.001433630	-2.012489	4.535493e-02
3	datiFine\$user_personal.sicuro	0.005263910	1.515029	1.311625e-01

4	datiFine\$user_personal.capisco	0.050336245	1.882572	6.104317e-02
5	datiFine\$user_personal.male	-0.003452724	-1.866805	6.322372e-02
6	datiFine\$user_grp.EMOTELOVE	-0.078732205	-2.195145	2.917106e-02
7	datiFine\$user_grp.EMOTESHOCK.1	0.169858309	2.766396	6.136972e-03
8	datiFine\$user_grp.EMOTEAMAZE.1	-0.007220743	-3.284397	1.184394e-03
9	datiFine\$user_grp.sicuro	0.068204686	2.021068	4.445208e-02
10	datiFine\$user_grp.aiuto	-0.096066818	-1.852176	6.530464e-02

Multiple R-squared: 0.08687, **Adjusted R-squared: 0.05051**
F-statistic: 2.389 on 9 and 226 DF, p-value: 0.01332

Il valore dell'R quadro corretto è 0.05051. Questo significa che la combinazione di variabili di questo modello riesce a spiegare il 5.051% della variabilità della variabile dipendente, cioè del valore della scala di Benessere Emotivo.

Analisi dei residui:

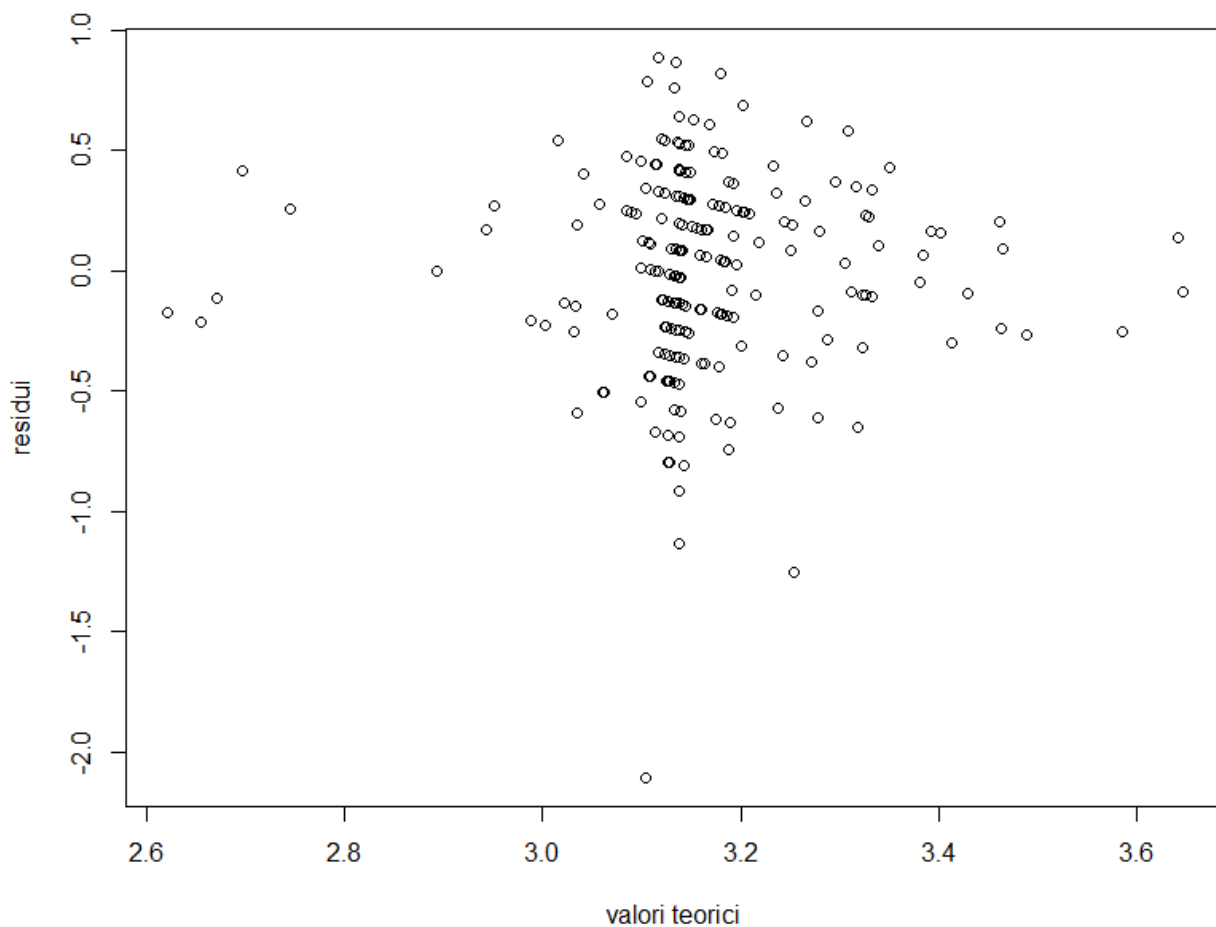


Figura 11: residui del modello per la scala di Benessere Emotivo con solo le variabili di analisi dei messaggi

Da questo grafico possiamo notare che non ci sono particolari problemi, perchè i residui sembrano disperdersi in maniera casuale sopra e sotto la loro media (lo zero).

Sono presenti pochi valori anomali al di sotto dello zero. Questo però non altera i risultati del modello perchè si tratta di pochi casi.

Capitolo 5

Conclusioni

Dall'analisi dei modelli descritti precedentemente si possono trovare delle relazioni tra le scale psicologiche relative alla regolazione delle emozioni e le variabili relative all'uso di Facebook.

Dai modelli che comprendono tutte le variabili, si nota che le variabili di Facebook che descrivono l'utente e le azioni principali di questo in Facebook sono molto significative per entrambe le scale. Infatti in entrambi i modelli che comprendono tutte le variabili sono presenti come regressori variabili quali le richieste di amicizia effettuate, il numero di foto inserite nella bacheca, il numero di eventi a cui l'utente ha dato conferma o ha partecipato e altre.

Più nello specifico, per il modello della scala di Personalità-Stabilità Emotiva si nota un valore del coefficiente positivo per le variabili come per esempio richieste di amicizia effettuate, eventi confermati e numero di foto pubblicate nella bacheca. Questo indica che un numero elevato di queste azioni contribuisce all'aumento del valore della scala psicologica di Personalità-Stabilità Emotiva, che corrisponde a una tendenza alla vulnerabilità, insicurezza, instabilità emotiva e agitazione.

Per il modello della scala di Benessere Emotivo, ci sono alcune variabili come il numero di richieste di amicizia effettuate, il numero di eventi a cui l'utente ha partecipato ed il numero di nuove foto pubblicate che presentano un valore del coefficiente negativo. Questo sta ad indicare che un aumento del valore di queste variabili contribuisce ad una diminuzione del valore della scala psicologica di Benessere Emotivo, che corrisponde cioè ad una tendenza ad avere un comportamento poco responsabile, poca capacità di capire gli altri e poca capacità di saper pazientare. Invece altre variabili come il numero di "mi piace" inseriti e il numero totale di foto presenti nel profilo presentano un coefficiente positivo, quindi un aumento del valore di queste contribuisce all'aumento del valore della scala psicologica, che corrisponde cioè al polo positivo di questa.

In questi modelli si nota la presenza di variabili che contano le emoticons e le parole scritte dall'utente in chat personali e di gruppo, ma anche molte variabili che contano emoticons e parole che la persona con cui l'utente sta scambiando messaggi scrive. Quindi sembra che anche la risposta a ciò che l'utente scrive possa essere legata alla personalità e al benessere di un individuo dal punto di vista emotivo.

Analizzando invece i modelli che prendono in considerazione esclusivamente le variabili che contano le parole che l'utente del profilo Facebook scrive in chat personali e di gruppo, possiamo fare le seguenti osservazioni.

Per la scala di Personalità-Stabilità Emotiva, si nota che le emoticons che esprimono amore e felicità e alcune parole quali "sicuro", "ragione", "bene", "tu", presentano un coefficiente con valore negativo. Questo indica che l'aumentare dell'utilizzo di queste parole è collegato alla tendenza ad una personalità sicura, capace di dominare gli stati d'ansia, rilassata e stabile emotivamente. Invece ci sono altre parole come "boh", "agitato/a", "male" che presentano invece un valore positivo del coefficiente, e quindi un aumento di scrittura di queste potrebbe

essere collegato ad una personalità poco sicura di sé poco capace di dominare gli stati d'ansia, poco rilassata e instabile emotivamente.

Per la scala di Benessere Emotivo si nota che l'emoticon che esprime tristezza e parole come “male” e “aiuto” presentano un coefficiente negativo. Questo indica che un aumento dell'utilizzo di queste parole contribuisce ad una diminuzione del valore di questa scala, cioè ad un comportamento poco responsabile e poco maturo emotivamente, alla poca capacità di comprendere gli altri e alla poca capacità di saper pazientare.

Invece parole come “sicuro” e “capisco” contribuiscono all'aumento del valore di questa scala e quindi a caratteristiche opposte a quelle elencate precedentemente, cioè comportamento responsabile, maturità emotiva, capacità di comprendere gli altri e di saper pazientare.

In conclusione, i risultati evidenziano uno stretto legame tra il comportamento reale degli user su Facebook e le loro caratteristiche individuali legate a tratti e competenze emotive.

BIBLIOGRAFIA

McCrae R., Costa P., *Validation of the five-factor model of personality across instruments and observers*, Journal of Personality and Social Psychology, 1987 Jan, vol 52, 81:90.

Caprara G.V., Barbaranelli C., Livi S., *Mapping Personality Dimensions in the Big Five Model*, European Review of Applied Psychology, 1994, vol. 44, 9:15.

Caprara G.V., Barbaranelli C., Borgogni L., Perugini M., *The “Big Five Questionnaire”: a new questionnaire to assess the five factor model*, Personality and Individual Differences, 1993, vol. 15, 281:288.

Furlong M. J., You S., Renshaw T. L., Smith D. C., O’Malley M. D., *Preliminary Development and Validation of the Social and Emotional Health Survey for Secondary School Students*, Social Indications Research, 2014, 117:1011.

Shaffer E., Suldo S., Loker T., March A., *How adolescents’ mental health predicts their physical health: unique contributions of indicators of subjective well-being and psychopathology*, Applied Research in Quality of Life, 2010 March, vol. 5, 203:217.

Meneghetti D., Righetto., S, Finos L. (2015), *GitHub-livioivil/myFBr: Collection of functions for text mining, specially devoted to the italian language*.

URL: <https://github.com/livioivil/myFBr>, agosto 2016.