

**Analisi dei Dati per il
Marketing Digitale II**

HO COMMESSO UN ***ERRORE DI COMUNICAZIONE***

**L'effetto delle controversie di marketing di Chiara
Ferragni sulla percezione del brand e sulle
intenzioni di acquisto: il caso "Pandoro-Gate."**

GRUPPO 6

Irene Marchetti

Matricola 187634

332909@studenti.unimore.it

INDICE

| | |
|------------------------------|----|
| 1. Introduzione | 4 |
| 2. Obiettivi Finali | 5 |
| 3. Metodi | 5 |
| 3.1 Raccolta dei Dati | 5 |
| 3.2 Analisi Statistiche..... | 6 |
| 4. Risultati..... | 25 |
| 5. Conclusioni | 26 |
| 6. Appendice..... | 27 |

1. INTRODUZIONE

In Italia, se dovessimo associare il ruolo dell'Influencer Marketing ad una specifica persona, il primo nome che ci verrebbe in mente sarebbe sicuramente il suo: **Chiara Ferragni**.

Dividiamo la sua vita, principalmente in **3 fasi**:

1. Nascita

*È il 2009: Nasce il blog "**The Blonde Salad**", creato da Chiara Ferragni e Riccardo Pozzoli. Il blog non parla solo di moda, ma di esperienze della protagonista.*

*Da lì una salita vertiginosa, che la portano nel 2013 alla realizzazione del vero brand: "**Chiara Ferragni Collection**".*

*Bilanci annuali folli, con fatturati altissimi che la portano nel 2015 ad essere la protagonista della copertina di Vogue Spagna, diventando la prima "**Imprenditrice Digitale**" al mondo a ottenere questo importante riconoscimento.*

2. Crescita

Il periodo tra il 2018 e il 2021 rappresenta sicuramente gli anni d'oro dell'Influencer, Fashion Blogger, Imprenditrice e chi più ne ha, più ne metta.

La famiglia Ferragnez diventa una delle famiglie più apprezzate in Italia dai social media e anche più seguite grazie:

a. Alla serie di Amazon Prime Video

*b. Al documentario "**Chiara Ferragni Unposted**"*

*Chiara Ferragni è alla conquista di tutto: collaborazioni con brand, fama, apprezzamenti e sicuramente non rinuncia alla proposta, nel 2021, di far parte del consiglio di amministrazione di **TOD'S**.*

3. Declino

*Nel 2022 arriva il sogno più seguito dagli Italiani: **Sanremo**.*

Un sogno che diventa realtà, ma che allo stesso tempo diventa un incubo per la co-presentatrice: Rosa Chemical e i baci non richiesti iniziano a far crollare quel "Palazzo di Cristallo" che la maggior parte di noi aveva pensato fosse "Casa Ferragnez".

*Nel 2023, con un documento di 25 pagine, l'**Antitrust** decide che l'operazione commerciale del pandoro "**Pink Christmas**", con relativa operazione benefica verso un ospedale pediatrico, "**costituisce, una pratica commerciale scorretta e ne vieta la reiterazione**". Le società di Chiara vengono sanzionate con una multa superiore un milione di euro. Sicuramente un'importante spesa monetaria, che comunque la "**Ferry**" si è potuta permettere.*

Il problema è stata la **perdita di followers**. Dallo scandalo ad oggi il profilo è diminuito di **800K** followers, che hanno portato ad una diminuzione:

- **di vendite del suo brand**
- **di collaborazioni con altri brand**

e quindi di **meno guadagni**.

Dal punto di vista imprenditoriale, sicuramente non siamo in una delle fasi migliori.

Ma c'è qualcosa che potrebbe riportare Chiara, il brand a rinascere e ricominciare a crescere?

Da qui ha inizio la nostra analisi, concentrata su come le vicende abbiano cambiato la percezione del brand agli occhi degli utenti, partendo dall'ipotesi che ci sia stato un impatto negativo.

Dopo aver realizzato e somministrato il questionario a un ampio campione di utenti, ci siamo chiesti se fosse possibile identificare un potenziale target di clienti su cui implementare una strategia di marketing mirata ad incrementare il tasso di conversione.

2. OBIETTIVI FINALI

L'obiettivo finale della nostra analisi è confermare che gli scandali recentemente avvenuti legati al personaggio dell'influencer Chiara Ferragni hanno influito negativamente sulla percezione e fiducia del brand da parte degli utenti, individuando in seguito un target di clienti potenziali sulla quale costruire una strategia di marketing.

Per fare questo abbiamo delineato dei sotto-obiettivi per raggiungere quello finale:

1. **Descrizione delle caratteristiche generali del dataset**
2. **Analisi delle distribuzioni di alcune variabili di interesse**
3. **Individuare delle componenti che racchiudessero le informazioni della Prima e Seconda sezione del questionario**
4. **Individuare diversi profili di clienti all'interno del campione**
5. **Analisi e descrizione di ogni profilo ottenuto, facendo attenzione sulla propensione di acquisto dopo gli scandali e i fattori che potrebbero influenzare il comportamento di acquisto del gruppo, così da poter individuare il target di potenziali clienti migliore**
6. **Individuare quali variabili aumentano la probabilità della propensione di acquisto dopo gli scandali all'interno del gruppo.**

3. METODI

3.1 RACCOLTA DEI DATI

Per la creazione del dataset su cui si basa la nostra analisi abbiamo raccolto le risposte del campione tramite un questionario realizzato da noi relativo a Chiara Ferragni e la percezione del Brand e dei prodotti.

Il questionario si compone di **25 domande** diviso in **5 batterie**:

1. **Io non volevo diventare famosa, non era questo il mio obiettivo**
La sezione è dedicata a indagare la conoscenza del soggetto su Chiara Ferragni
2. **Chiara dice che mi adora**
La sezione si concentra sul valutare la differenza tra l'opinione prima degli scandali e l'opinione dopo relativa a Chiara Ferragni, e quanto sia stata influenzata dagli scandali
3. **Questo è il pezzo mio preferito!**
La sezione riguarda la propensione di acquisto del soggetto nel momento prima e dopo gli scandali
4. **La dichiarazione più forte, per me, è la personalità**
La sezione contiene domande grazie a cui il soggetto valuta quanto sette fattori potrebbero influenzare la decisione di acquisto.
5. **Sezione Anagrafica**
La sezione riguarda la parte anagrafica e abitudini d'uso dei social.

Abbiamo ottenuto 25 variabili, sia quantitative che qualitative, che abbiamo deciso di considerare all'interno del report. Le variabili qualitative sono state prima codificate in numeri.

Il questionario è stato somministrato **dal 24/10/2024 fino al 12/11/2024**, diffuso tramite Whatsapp, Instagram, LinkedIn e Moodle raggiungendo un totale di 255 risposte.

3.2 ANALISI STATISTICA

In questa sezione andremo a descrivere dettagliatamente i metodi utilizzati nella nostra analisi per il raggiungimento dell'obiettivo finale.

Analisi descrittive univariate del dataset (3.2.1)

Analizziamo e rappresentiamo graficamente alcune variabili per descrivere il campione in esame

Analisi delle frequenze relative delle variabili di interesse (3.2.2)

Analisi delle Componenti Principali (3.2.3)

Riduciamo le dimensioni delle variabili, al fine di ottenere delle componenti principali che descrivano la conoscenza dell'influencer in questione, l'influenza degli scandali nella formulazione dell'opinione personale e l'utilizzo dei social.

Analisi Cluster (3.2.4)

Per individuare i diversi profili di clienti presenti all'interno del campione abbiamo utilizzato un'analisi cluster, utilizzando le componenti estratte precedentemente dall'ACP. Abbiamo testato sia un Metodo Gerarchico che un Metodo Non Gerarchico, e in seguito, dal confronto dei due abbiamo deciso di proseguire con il metodo non gerarchico

Analisi descrittive per ogni cluster (3.2.5)

Per la descrizione di ogni cluster ottenuto abbiamo utilizzato delle analisi univariate per il genere, l'età, altre variabili di interesse che impattano maggiormente la divisione in gruppi e la propensione di acquisto prima e dopo gli scandali per individuare il miglior target di clienti potenziali.

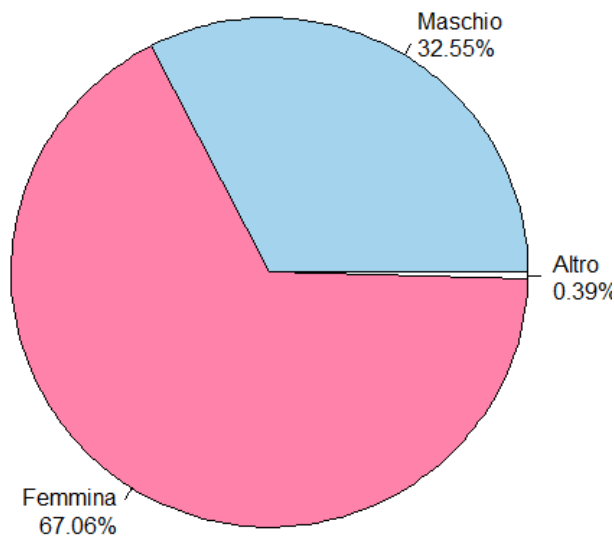
Inoltre, abbiamo creato dei boxplot che rappresentano quali fattori potrebbero influenzare maggiormente il comportamento di acquisto, importante per poter sviluppare una strategia di marketing futura.

Regressione Logistica (3.2.6)

Tramite la regressione logistica, abbiamo individuato quali sono le variabili che influenzano la probabilità di avere una propensione di acquisto più alta dopo gli scandali sul target selezionato.

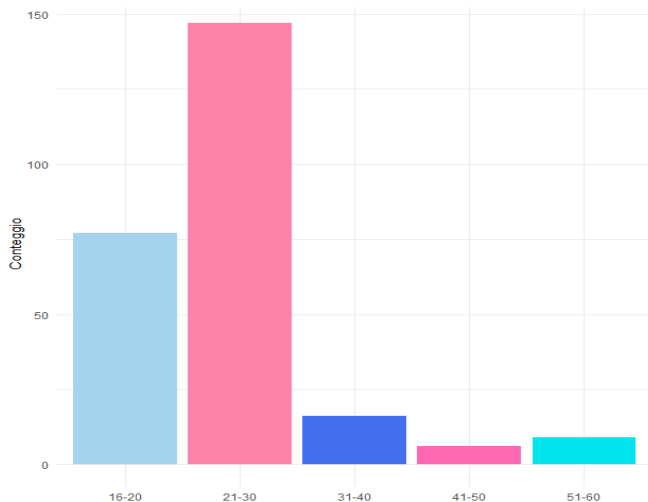
3.2.1 Analisi Descrittive Univariata

Figura 1: Percentuali Genere



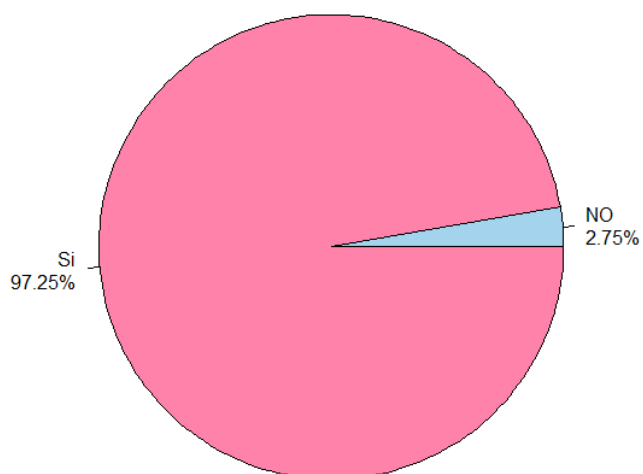
Il campione formato da 255 unità statistiche risulta essere composto per il **67,06% da femmine**, per il **32,55% da maschi** e per un **0,39%** da soggetti che hanno selezionato l'opzione "Altro" (Figura 1).

Figura 2: Distribuzioni delle età



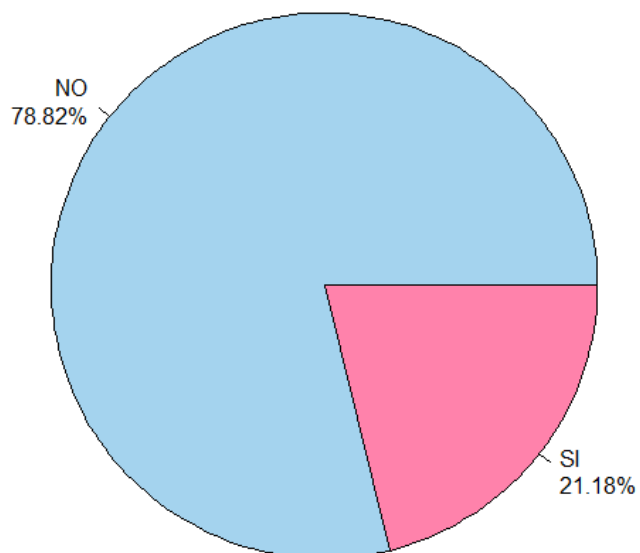
Presenta un'età media di circa **25 anni (24.75)** e una **deviazione standard di 8.10**, ciò significa che la maggior parte delle età nel dataset si trova entro un intervallo di circa 8 anni sopra o sotto la media. In altre parole, la maggior parte delle età è compresa tra circa **17 e 33 anni (25 ± 8.10)** (Figura 2).

Figura 3: Percentuali Conoscenza Scandali



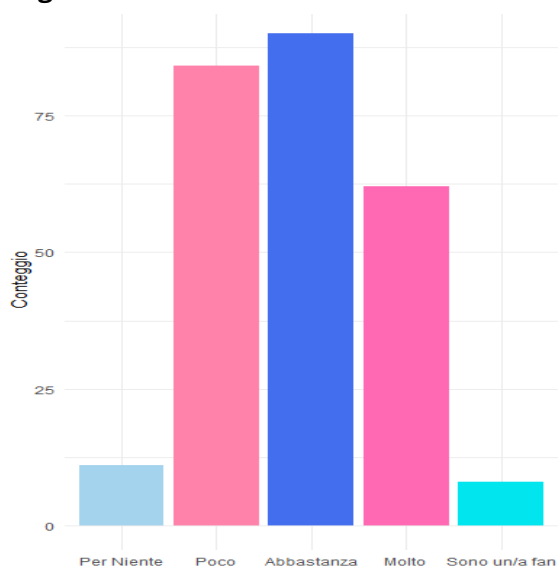
Possiamo affermare che quasi tutte le unità statistiche sono a conoscenza degli scandali legati a Chiara Ferragni. Infatti il **97,25%** del campione ha risposto "Si" alla domanda "Sei a conoscenza degli scandali legati a Chiara Ferragni?", mentre solo il **2,75%** ha selezionato "No" (Figura 3).

Figura 4: Percentuali Acquisto Fatto



Il **21,18%** dichiara di aver acquistato almeno un prodotto promosso da Chiara Ferragni (*Figura 4*).

Figura 5: Distribuzione dei livelli di conoscenza



Per quanto riguarda i livelli di conoscenza su Chiara Ferragni, la *Figura 5* mostra che **90** soggetti hanno dichiarato di conoscere *“Abbastanza”* l’influencer, **84** *“Poco”*, **62** *“Molto”*, **11** *“Per Niente”*, **8** dichiarano di considerarsi un **fan**.

3.2.2 Analisi delle Frequenze Relative

| FREQUENZE RELATIVE | | | | | |
|-----------------------------|------------------|------------|-------------------|-------------------|------------------|
| PRIMA SEZIONE | | | | | |
| | Per Niente | Poco | Abbastanza | | Completamente |
| Importanza Ruolo Influencer | 28,50% | 36,70% | 23,80% | 10,50% | 0,40% |
| Conoscenza Chiara | 4,30% | 33,20% | 35,20% | 24,20% | 3,10% |
| Frequenza Contenuti Chiara | 24,20% | 22,70% | 13,70% | 23,40% | 16,00% |
| | No | Sì | | | |
| Vita Pubblica | 24% | 76% | | | |
| Attività Imprenditoriali | 25% | 75% | | | |
| Conoscenza Scandali | 3% | 97% | | | |
| SECONDA SEZIONE | | | | | |
| | Per Niente | Poco | Abbastanza | Molto | Completamente |
| Influenza Scandali | 12,10% | 29,30% | 28,50% | 22,30% | 7,80% |
| | Molto Negativa | Negativa | Indifferente | Positiva | Molto Positiva |
| Opinione Prima | 4,30% | 8,20% | 45,30% | 34,00% | 8,20% |
| Opinione Dopo | 14,10% | 35,90% | 42,60% | 7,00% | 0,40% |
| | Molto Peggiorata | Peggiorata | Indifferente | Migliorata | Molto Migliorata |
| Notizia Uova | 15,20% | 41,40% | 43,00% | 0,00% | 0,40% |
| | No | Sì | | | |
| Umanità | 55,90% | 44,10% | | | |
| TERZA SEZIONE | | | | | |
| | No | Sì | | | |
| Acquisto Fatto Prima | 78,90% | 21,10% | | | |
| Propensione Acquisto Prima | 70,70% | 29,30% | | | |
| | Per Niente | Poco | Abbastanza | Molto | Completamente |
| Propensione Acquisto Dopo | 54,70% | 16,40% | 21,10% | 7,40% | 0,40% |
| QUARTA SEZIONE | | | | | |
| | Per niente | Poco | Abbastanza | Molto | Completamente |
| Fattore Prezzo | 36,00% | 17,00% | 31,00% | 11,00% | 5,00% |
| Fattore Tendenza | 46,00% | 29,00% | 21,00% | 3,00% | 1,00% |
| Fattore Trasparenza | 34,00% | 23,00% | 24,00% | 15,00% | 4,00% |
| Design e Stile | 27,00% | 17,00% | 30,00% | 18,00% | 7,00% |
| Fattore Qualità | 24,00% | 14,00% | 30,00% | 24,00% | 9,00% |
| Fattore Experience | 36,00% | 24,00% | 26,00% | 10,00% | 4,00% |
| Fattore Innocenza | 51,00% | 20,00% | 20,00% | 7,00% | 2,00% |
| QUINTA SEZIONE | | | | | |
| | Mai | Raramente | < 1 volta a sett. | > 1 volta a sett. | Quotidianamente |
| Uso Social | 1,20% | 2,70% | 1,60% | 7,00% | 87,50% |
| | No | Sì | | | |
| Segui altri Influencer | 18,00% | 82,00% | | | |

La *Tabella 1* mostra che la maggior parte della popolazione conosce abbastanza bene Chiara Ferragni, le sue attività imprenditoriali, la sua vita pubblica e gli scandali in cui è stata coinvolta.

La maggioranza dei soggetti dichiara che gli scandali hanno influenzato “*Abbastanza*” (29%) o “*Poco*” (29%) sulla formulazione della propria opinione personale riguardo l’influencer, ma si nota un cambiamento in senso **negativo** dell’opinione degli utenti in seguito agli scandali: se prima degli eventi il 42% degli intervistati aveva un’opinione “*Positiva*” o migliore, con l’uscita degli scandali **solo il 7% ha dato un punteggio positivo**.

Inoltre, inizialmente l’8% ha dichiarato di aver avuto un’opinione “**Molto Positiva**” prima degli scandali, cosa che **non si mantiene** nella rilevazione dell’opinione dopo (0% “*Molto Positiva*”).

Una buona percentuale, 71%, dichiara di essere stato propenso all’acquisto prima degli scandali.

Percentuale che si modifica in seguito agli scandali, infatti il 55% ha selezionato “*Per Niente*” alla

domanda 2: “*Quanto sei propenso ad acquistare suoi prodotti dopo lo scandalo?*”.

La **qualità** si classifica al primo posto per essere il fattore più importante che determina il comportamento di acquisto (9% “*Completamente*”, 24% “*Molto*”).

Una buona percentuale dei soggetti (88%) usa **quotidianamente** i social e l’82% segue influencer.

3.2.3 Analisi delle Componenti Principali (ACP)

L'ACP è un'analisi statistica che permette di ridurre la dimensionalità del dataset pur mantenendo la maggior parte della variabilità presente nei dati. Trasforma le variabili in nuove variabili non correlate tra loro, ordinate in modo tale che la prima componente cattura la massima varianza possibile, la seconda cattura la massima varianza residua, e così via. Effettuiamo l'ACP solo su alcune variabili interessanti:

1. ruolo_influencer

Valuta quanto il soggetto ritiene importante il ruolo dell'influencer nelle sue decisioni di acquisto da 1 a 5

2. conoscenza_chiara

Valuta quanto il soggetto afferma di conoscere Chiara Ferragni su una scala da 1 a 5

3. frequenza_contenuti

Indica con quale frequenza il soggetto entra in contatto con i contenuti dell'influencer di Chiara Ferragni, su una scala da 1 a 5

4. attività_imprenditoriali

È una variabile dicotomica del tipo Si/No. Essa valuta se il soggetto conosce o meno le attività imprenditoriali dell'influencer

5. vita_pubblica

È una variabile dicotomica del tipo Si/No. Essa valuta se il soggetto conosce o meno aspetti della vita pubblica di Chiara Ferragni

6. uso_social

Indica con quale frequenza il soggetto utilizza i social, su una scala da 1 a 5

7. segui_influencer,

È una variabile dicotomica del tipo Si/No, valuta se il soggetto segue altri influencer

8. influenza_scandali,

Indica qual è stata l'influenza degli scandali nella formulazione dell'opinione personale del soggetto, su una scala da 1 a 5

9. opinione_prima

Indica il tipo di opinione che il soggetto aveva nei confronti di Chiara Ferragni prima degli scandali, su una scala da 1 a 5 dove 1 è uguale a "Molto Negativa" e 5 corrisponde a "Molto Positiva"

10. opinione_dopo

Indica il tipo di opinione che il soggetto ha nei confronti di Chiara Ferragni dopo degli scandali, su una scala da 1 a 5, dove 1 è uguale a "Molto Negativa" e 5 corrisponde a "Molto Positiva"

11. notizia_uova

Indica quanto il venire a conoscenza del secondo scandalo legato alle uova di Pasqua promosse dall'influencer, ha influito sulla formulazione dell'opinione personale. In una scala da 1 a 5, dove 1 è uguale a "Molto Peggiorata" e 5 corrisponde a "Molto Migliorata"

12. umanità

È una variabile dicotomica del tipo Si/No, valuta se secondo il soggetto essere venuto a conoscenza degli scandali abbia contribuito a vedere la figura di Chiara Ferragni più umana e fallibile

Ci aspettiamo che le variabili con la deviazione standard maggiore, siano quelle che pesano maggiormente nella formazione e all'interno delle Componenti Principali. Le variabili con deviazione standard maggiore sono le variabili **1,2,3,8,9 e 10**.

È possibile fare un'analisi ACP se le correlazioni tra le variabili non sono omogenee. Per descrivere questo aspetto si crea una **matrice di correlazione** (*Figura 1, App.*).

Per decidere se proseguire l'ACP si applica il test **KMO**.

Se $KMO > 0.6$, vale la pena indagare oltre. **Il nostro KMO è di 0.7390472**, quindi decidiamo di proseguire con l'analisi.

Come risultato dell'analisi abbiamo 12 Componenti Principali e decidiamo di estrarre le **prime 4**, quelle con la **deviazione standard maggiore di 1**, che insieme spiegano circa **il 61% della variabilità totale**.

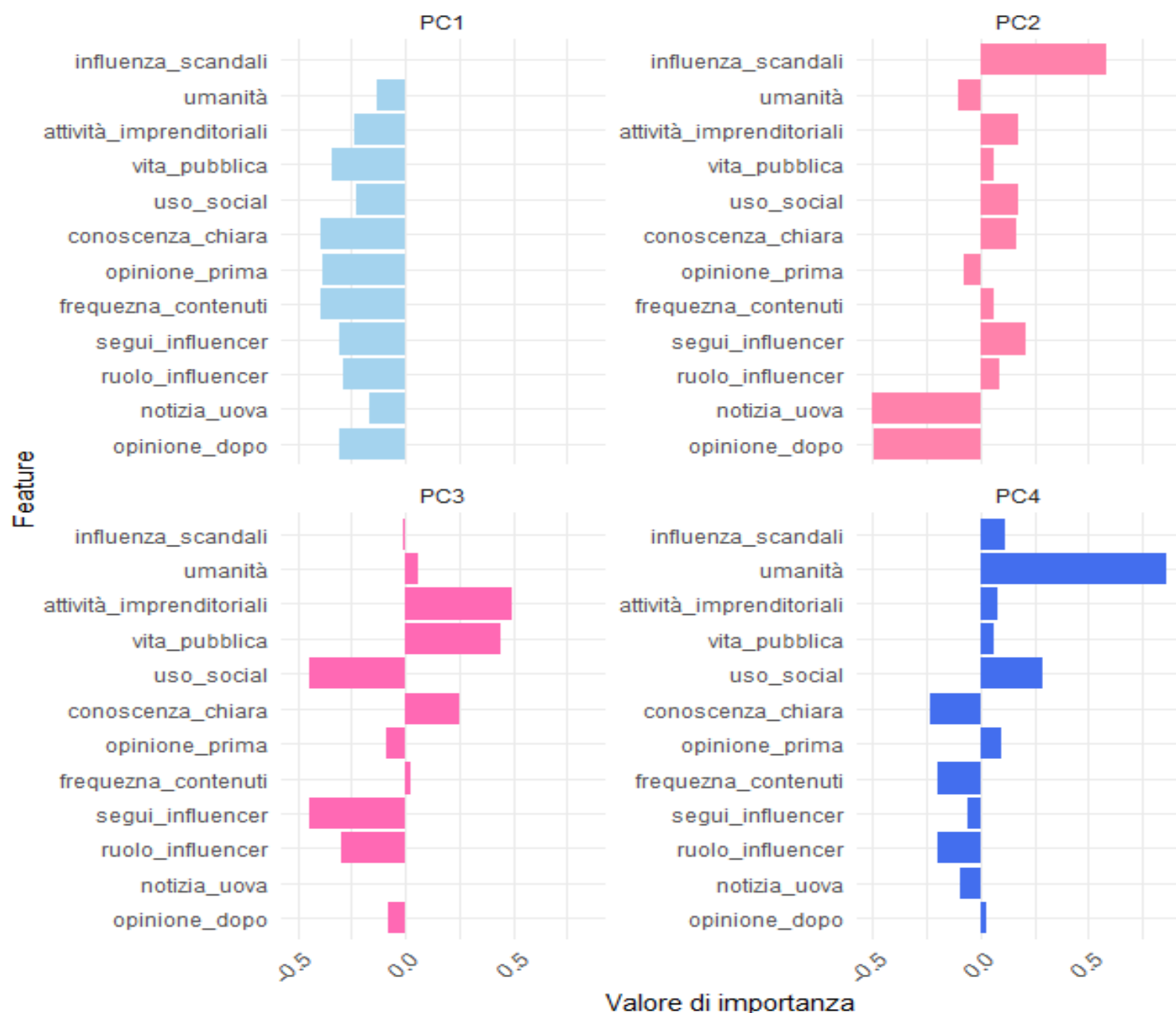


Figura 6: Importanza delle variabili per PC

La Prima componente principale (PC1) rappresenta una misura della "Conoscenza di Chiara"

Questa componente è influenzata dalle variabili legate alla conoscenza di Chiara, alla conoscenza della sua vita pubblica, alla conoscenza delle sue attività imprenditoriali e alla frequenza dei suoi contenuti.

Un valore negativo di PC1 indica un alto livello di conoscenza dell'influencer.

La Seconda componente principale (PC2) rappresenta una misura del "Influenza degli scandali"

Questa componente è influenzata dalle variabili "influenza_scandali" e "opinione_dopo".

Un valore positivo di PC2 indica un alto livello di influenza degli scandali.

Un valore negativo di PC2 indica un'opinione positiva dopo gli scandali.

La Terza componente principale (PC3) rappresenta un indicatore del "Uso dei Social"

Questa componente è influenzata dalle variabili "uso_social" e "seguì_influencer".

Un valore negativo di PC3 indica che il soggetto utilizza molto i social e segue altri influencer, ma non Chiara Ferragni. Un valore positivo di PC3 conosce la vita e le attività imprenditoriali di Chiara ma usa poco i social.

La Quarta componente principale (PC4) rappresenta una misura della "Percezione dell'influencer"

Questa componente è influenzata dalle variabili "umanità".

Un valore positivo di PC4 indica che il soggetto percepisce la figura di Chiara più umana dopo gli scandali.

3.2.4 Analisi Cluster

L'**Analisi Cluster** è una tecnica statistica utilizzata per raggruppare un insieme di soggetti in gruppi (o cluster), in modo che gli soggetti all'interno di ciascun cluster siano più simili tra loro rispetto agli soggetti appartenenti a gruppi diversi. Tra i metodi principali di clusterizzazione troviamo:

- **Metodo Gerarchico**: costruisce una struttura ad albero (**dendrogramma**) che rappresenta le relazioni gerarchiche tra i dati
- **Metodo Non Gerarchico**: richiede di specificare il **numero di cluster a priori**. L'algoritmo inizia con una selezione casuale di centri di cluster (**centroidi**) e assegna ogni punto dati al cluster con il centroide più vicino (es. *K-means*)

Nella nostra analisi abbiamo deciso di testare entrambi i metodi.

METODO GERARCHICO

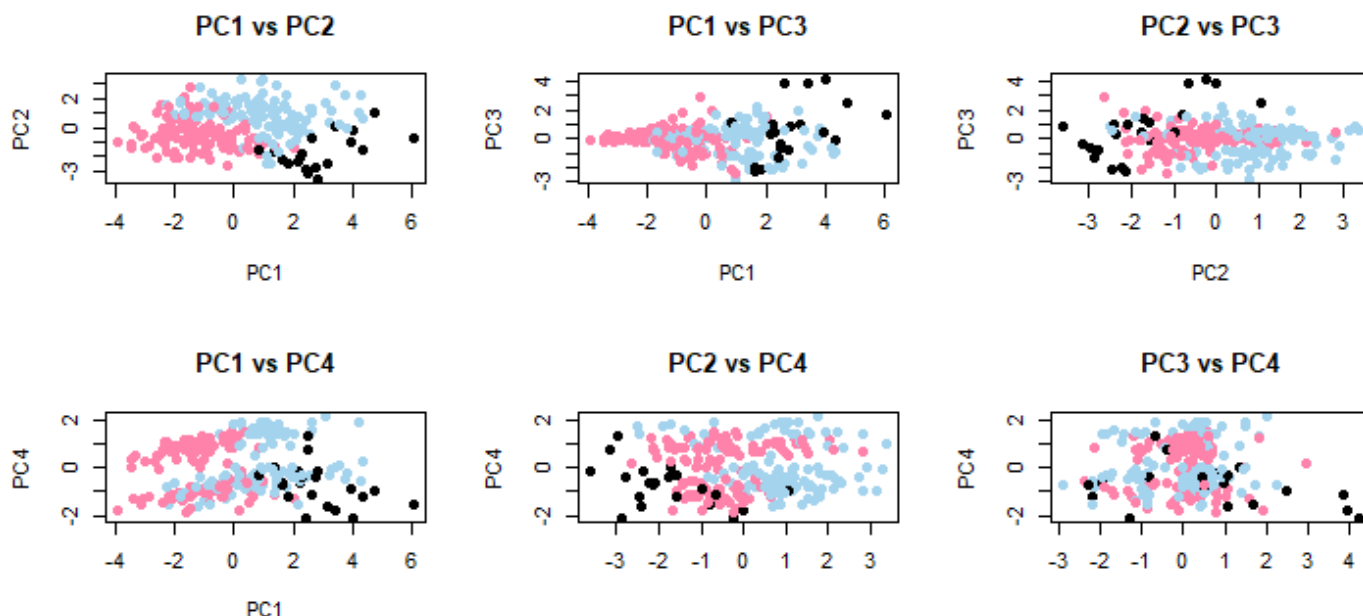
Abbiamo applicato la tecnica di clusterizzazione utilizzando il **Metodo Gerarchico** per le Componenti Principali estratte precedentemente, con lo scopo di creare dei cluster basati sulle caratteristiche principali identificate dall'Analisi delle Componenti Principali (ACP).

Per individuare il metodo più idoneo alla clusterizzazione, abbiamo confrontato i dendrogrammi ottenuti con legami differenti (*Figura 2 App.*): il metodo del "**Legame Completo**", con la metrica di distanza **Gower** si è rivelato il più adatto per rappresentare le relazioni tra le osservazioni, suggerendo una suddivisione in tre cluster (*Figura 2bis App.*). Per validare la scelta del numero di cluster, valutiamo tre tipi di suddivisioni in cluster (*Figura 3 App.*).

La prima valutazione "ad occhio", suggerisce una divisione in 4 (*Figura 3bis App.*); La funzione **NBClust**, che combina diversi criteri matematici per suggerire il numero ottimale di gruppi, ha indicato una soluzione con 2 cluster (*Figura 3ter App.*); tuttavia, abbiamo optato per mantenere la suddivisione in **tre cluster** (*Figura 3quater App.*) per i seguenti motivi:

- **Dimensioni dei cluster**: la soluzione con due cluster presentava gruppi di dimensioni troppo sbilanciate, mentre la suddivisione in tre cluster risultava più equilibrata
- **Analisi qualitativa**: ulteriori approfondimenti hanno evidenziato che i tre cluster individuati riflettono gruppi ben distinti e significativi nel dataset

Per visualizzare graficamente la distanza, e quindi la diversità, tra i vari cluster abbiamo creato degli **scatter-plot** dai confronti di ogni componente principale con le altre (*Figura 7*)



Nota: il rosa corrisponde al gruppo 1, il celeste al 2, il nero al 3.

Cluster formati con Legame Completo, metrica: "Gower". N. unità: 255

Figura 7: Scatter-Plot con Metodo Gerarchico

Il grafico del confronto “**PC1 vs PC2**” mette in evidenza, l’esistenza di tre gruppi abbastanza diversi tra di loro, anche se lievemente sovrapposti e con numerosità diverse: il gruppo di colorazione nera è meno numeroso rispetto agli altri due. Per questo, non soddisfatti del risultato, procediamo con l’analisi cluster utilizzando il **metodo non gerarchico**.

METODO NON GERARCHICO

Utilizzando lo stesso procedimento dell’analisi precedente, e selezionando come numero a priori di cluster 3, abbiamo ottenuto altri 6 diversi scatter-plot. (Figura 8)

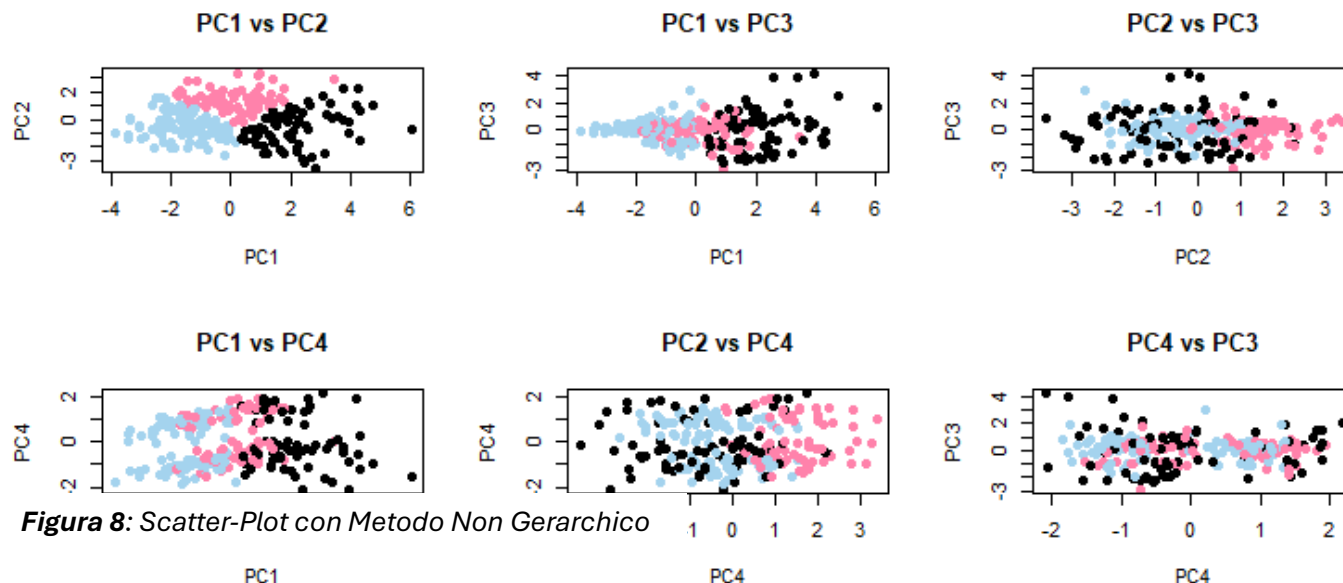


Figura 8: Scatter-Plot con Metodo Non Gerarchico

Confrontiamo ora gli scatter-plot “**PC1 vs PC2**” di entrambi i metodi (Figura 9).

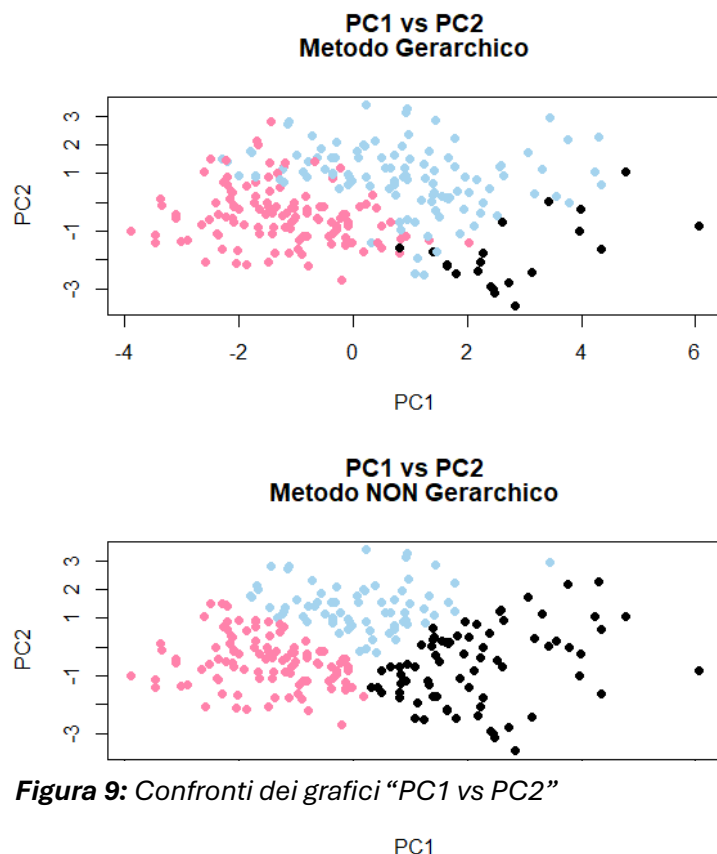


Figura 9: Confronti dei grafici “PC1 vs PC2”

Per la nostra analisi riteniamo che il **Metodo Non Gerarchico** sia il più idoneo in quanto:

- **Individua tre gruppi ben distinti senza sovrapposizione**, al contrario dello scatterplot gerarchico
- **Il gruppo nero nello scatterplot non gerarchico è più numeroso dello scatterplot gerarchico.**

GRUPPO ROSA

Corrisponde al cluster 1.

Ha una PC1 negativa, significa che i soggetti del gruppo conoscono bene Chiara Ferragni, la sua vita pubblica e le sue attività imprenditoriali e inoltre avevano un'opinione positiva prima degli scandali.

Ha una PC2 per la maggior parte negativa, ciò dimostra esserci stata solo una leggera variazione nell'opinione dopo che comunque è migliore rispetto agli altri gruppi.

Da adesso in poi li chiameremo “Fan Di Chiara”.

GRUPPO CELESTE

Corrisponde al cluster 2

Ha una PC1 che va da circa -2 a circa 4, dimostra che alcuni dei soggetti conoscono abbastanza bene l'influencer, mentre altri non la conoscono.

Ha una PC2 positiva, questo significa che hanno un 'opinione abbastanza negativa che è stata molto influenzata dagli scandali.

Da adesso in poi li chiameremo “Indignati”.

GRUPPO NERO

Corrisponde al gruppo 3

Ha una PC1 positiva, quindi i soggetti del gruppo non conoscono Chiara.

Ha una PC2 che va da circa -3 a circa 2, questo significa che non sono stati molto influenzati dagli scandali e hanno un'opinione dopo abbastanza indifferente o negativa.

Da adesso in poi li chiameremo “Indifferenti”.

Ipotizziamo, dunque, che probabilmente il gruppo “**Fan di Chiara**” avrà una propensione di acquisto dopo gli scandali maggiore rispetto agli altri. Inoltre, hanno una conoscenza discreta di Chiara, e vengono a contatto con i suoi contenuti più frequentemente degli altri gruppi; infatti, è il gruppo con PC1 più negativa. Dai valori della PC2, il gruppo sembra comunque avere un'opinione dopo non del tutto negativa, considerando che gli scandali non hanno avuto molta influenza.

Il gruppo “**Indignati**” non va escluso a priori perché:

- **C'è una buona parte che conosce abbastanza bene Chiara Ferragni**
- **E' il gruppo su cui gli scandali hanno avuto una maggiore influenza: si denota quindi un “attaccamento” e interessamento all'influencer e alle sue attività**

Se il gruppo non fosse composto da persone particolarmente interessate a Chiara, sarebbe difficile spiegare perché lo scandalo possa avere un'influenza così forte. Quindi questo potrebbe suggerire che anche se non sono fan stretti potrebbero comunque avere un certo interessamento o sensibilità nei suoi confronti. Se lo scandalo che li ha influenzati negativamente non è stato molto grave da compromettere la loro opinione e se il team di Chiara Ferragni riuscisse a gestire in maniera positiva la situazione potrebbero tornare a considerarla favorevolmente.

Pensiamo che il gruppo “**Indifferenti**” fosse lontano dall'influencer e dal suo mondo anche prima degli scandali, in quanto:

- **La loro conoscenza sembra essere bassa**
- **L'influenza degli scandali è minore rispetto agli altri gruppi**

Questi due aspetti portano a pensare che non ci sia stato, né c'è tutt'ora un interessamento da parte di questo gruppo. Ciò renderebbe ancora più difficile il fatto che possano acquistare suoi prodotti.

3.2.5 Analisi Descrittive Univariata dei gruppi

Per ogni gruppo abbiamo eseguito delle analisi su variabili di tipo anagrafico, per una descrizione generale del gruppo, e sui valori delle mediane delle variabili che caratterizzano e differenziano ogni gruppo. Infine, abbiamo analizzato le variabili obiettivo, ovvero la propensione di acquisto prima e dopo gli scandali per vedere come cambiasse e quale gruppo possiede la propensione di acquisto dopo più alta.

Tramite dei boxplot abbiamo individuato quali fattori influenzino di più il gruppo nel suo comportamento di acquisto. I fattori presi in esame sono: Prezzo, Qualità, Tendenza, Design e Stile, Trasparenza nella Comunicazione, Experience di Acquisto e Dichiarazione di Innocenza di Chiara Ferragni.

Fan di Chiara

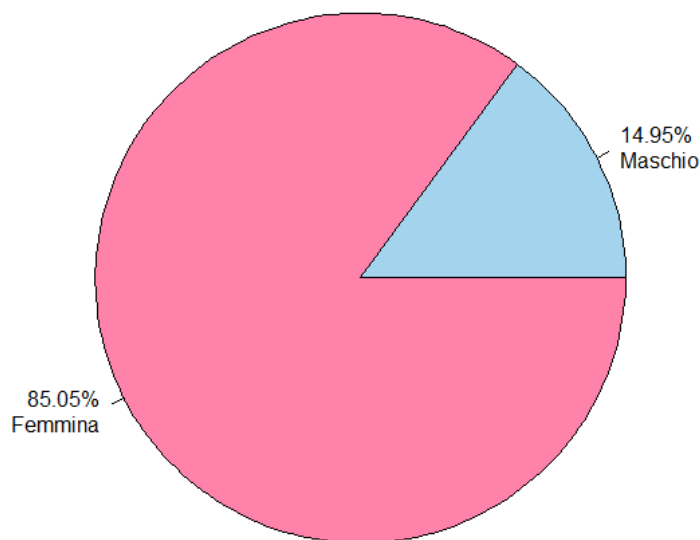


Figura 10: Percentuali Genere Fan Di Chiara

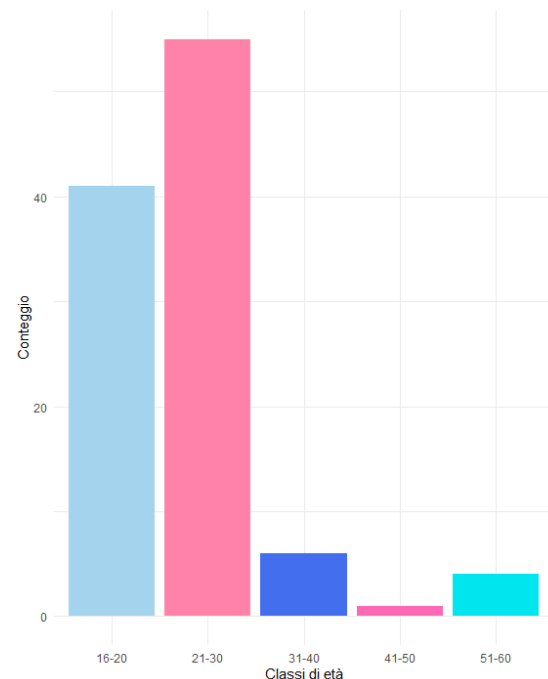


Figura 11: Distribuzioni di età Fan di Chiara

Il gruppo è formato da **107 unità statistiche**, composto per l'**85,05% da femmine** e per il **14,95% da maschi** (Figura 10).

Presenta un'età media di circa **25 anni (24.26)** e una deviazione standard di **8.04**, ciò significa che la maggior parte delle età nel gruppo si trova entro un intervallo di circa **8 anni sopra o sotto la media**. In altre parole, la maggior parte delle età è compresa tra circa **17 e 33 anni (25 ± 8.04)**. (Figura 11).

| | Fan Chiara |
|--------------------|------------|
| Conoscenza Chiara | 3 |
| Uso Social | 5 |
| Opinione Prima | 4 |
| Opinione Dopo | 3 |
| Influenza Scandali | 2 |

Tabella 2: Mediane dei Fan di Chiara

Il gruppo si caratterizza per avere una buona conoscenza di Chiara. La variabile `opinione_prima` ha come mediana 4 che corrisponde a “Positiva” nella scala posta all’interno del questionario. C’è stato un cambiamento in negativo nell’opinione passando da mediana 4 a 3, “Indifferente”, ma comunque gli scandali hanno influenzato poco il gruppo, questo lo dimostra la mediana della variabile `influenza_scandali` che è pari a due, ossia “Poco”.

Nota: la tabella riporta solo le mediane delle variabile misurate all’interno del gruppo (107 unità)

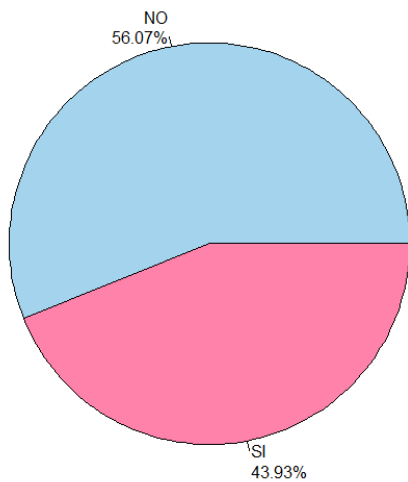


Tabella 3: Propensione di Acquisto Dopo Fan Di Chiara

| <i>Min</i> | <i>1st Qu.</i> | <i>Median</i> | <i>Mean</i> | <i>3rd Qu.</i> | <i>Max.</i> |
|------------|----------------|---------------|-------------|----------------|-------------|
| 1,000 | 1,000 | 3,000 | 2,355 | 3,000 | 4,000 |

Nota: la tabella riporta i valori relativi alla variabile "propensione_acquisto_dopo" del gruppo fan di chiara

Figura 12: Percentuali Propensione di Acquisto Prima

Dal grafico a torta (*Figura 12*) risulta che abbiamo il **44%** del campione **era propenso ad un acquisto prima**, dopo gli scandali la propensione agli acquisti ha mediana pari a 3 e una media di 2,355 che corrisponderebbe al 33,88%, perciò possiamo dire che **il cambiamento è stato minimo**. Ragione per cui crediamo che questo cluster sia il **migliore** per costruire una strategia di mercato volta all'acquisto.

Dai **Boxplot** in *Figura 13*, i fattori più importanti per il gruppo che condizionerebbero la decisione di acquisto sono **prezzo, qualità, trasparenza, experience e design/stile**.

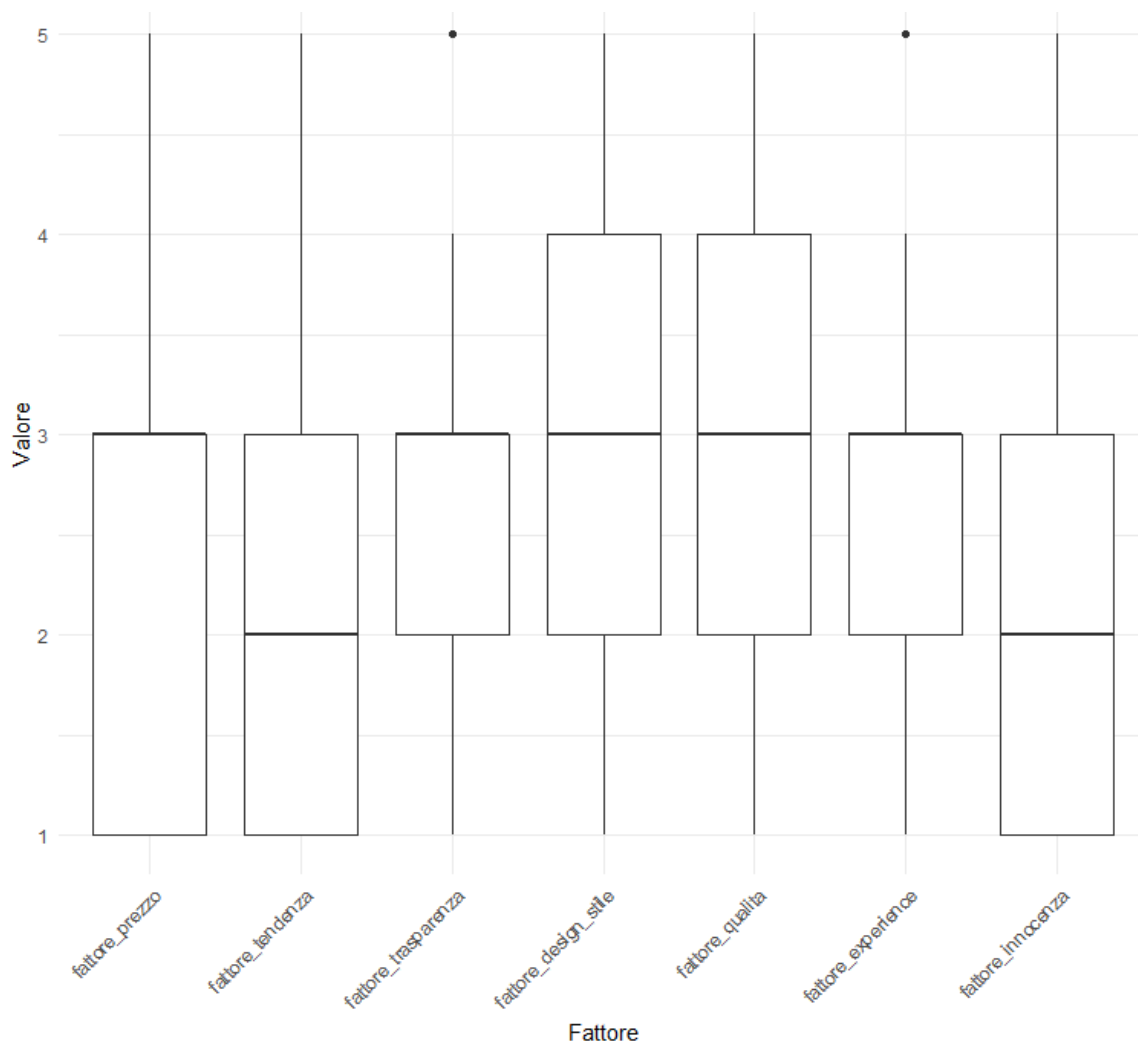


Figura 13: BoxPlot Fattori Fan Chiara

Indignati

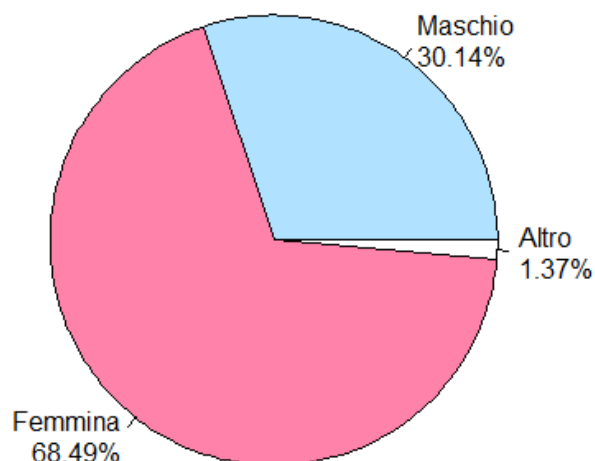


Figura 14: Percentuali Genere Indignati

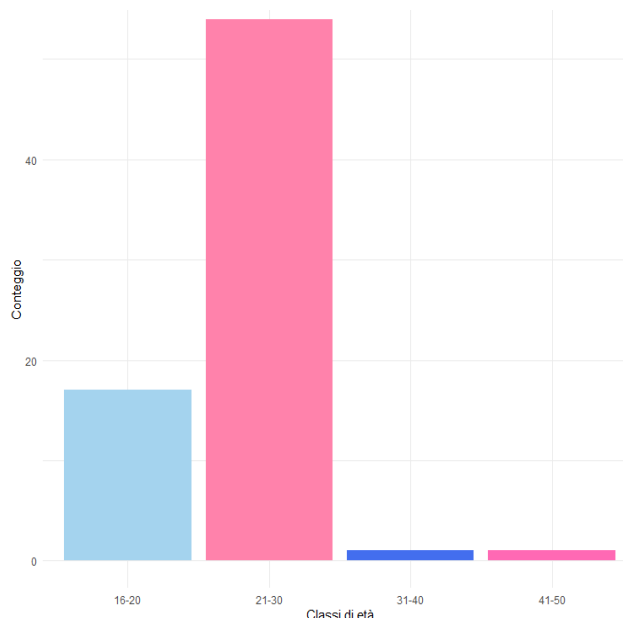


Figura 15: Distribuzione Età Indignati

Il gruppo è formato da **73 unità statistiche**, composto per il **68,49%** da **femmine**, per il **30,14%** da **maschi**, e da un **1,37%** che ha selezionato “**Altro**” (Figura 14).

Presenta un’età media di circa **22 anni (22.56)** e una **deviazione standard di 3.71**, ciò significa che la maggior parte delle età nel gruppo si trova entro un intervallo di circa **4 anni sopra o sotto la media**. In altre parole, la maggior parte delle età è compresa tra **circa 18 e 26 anni (22 ± 3.71)**. (Figura 15)

| | <i>Fan Chiara</i> | <i>Indignati</i> |
|---------------------------|-------------------|------------------|
| Conoscenza Chiara | 3 | 3 |
| Uso Social | 5 | 5 |
| Opinione Prima | 4 | 3 |
| Opinione Dopo | 3 | 2 |
| Influenza Scandali | 2 | 4 |

Tabella 4: Mediane dei Fan di Chiara e Indignati

Nota: la tabella riporta i valori relativi alla variabile “propensione_acquisto_dopo” del gruppo Indignati

Il gruppo si caratterizza per avere una buona conoscenza di Chiara. La mediana della variabile *opinione_prima* è uguale a 3 che tra le modalità corrisponde a “Indifferente”. In seguito agli scandali l’opinione peggiora e passa ad avere una mediana pari a 2 cioè “Negativa”. Gli scandali su questo gruppo hanno avuto molta influenza, infatti la mediana della variabile che misura quanto gli eventi successi abbiano influenzato l’opinione è pari a 4 che corrisponde a “Molto”, a differenza del gruppo Fan Di Chiara che ha mediana pari a 2, “Poco”.

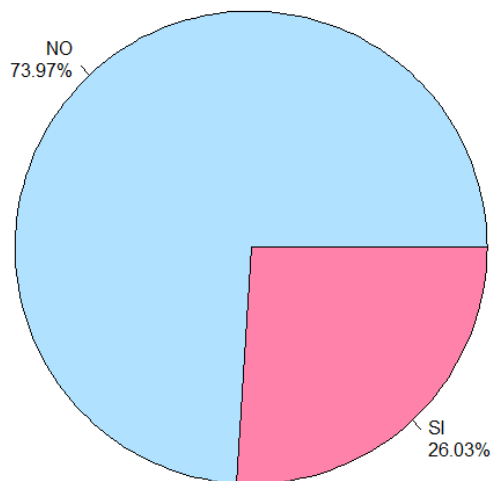


Tabella 5: Propensione di Acquisto Dopo Indignati

| Min | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
|-------|---------|--------|-------|---------|-------|
| 1,000 | 1,000 | 1,000 | 1,384 | 2,000 | 3,000 |

Nota: la tabella riporta solo le mediane delle variabili misurate all'interno dei gruppi (107 e 73 unità)

Figura 16: Percentuali Propensione di Acquisto Prima Indignati

Dal grafico a torta risulta che solo il **26%** sarebbe stato propenso ad acquistare prima degli scandali (*Figura 16*). Questa percentuale è **diminuita in seguito**, infatti la variabile propensione_acquisto_dopo ha mediana 1 e media 1,384 (*Tabella 5*) che in uno scenario binario corrisponde a circa il 10%.

Dai **boxplot** in *Figura 17* i fattori che potrebbero influenzare la decisione di acquisto di un prodotto Chiara Ferragni sono il **prezzo, la qualità e il design/stile**.

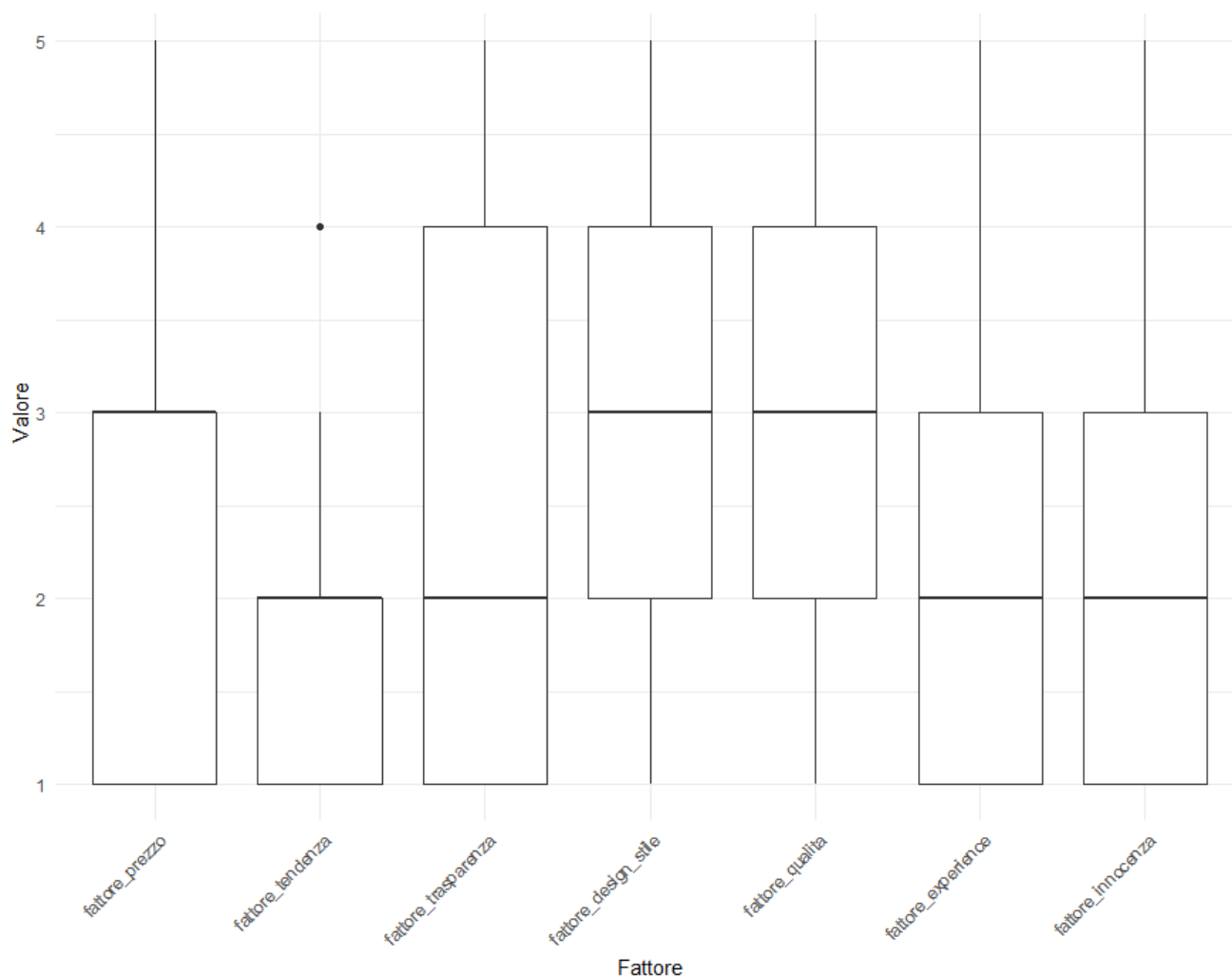


Figura 17: BoxPlot Fattori Indignati

Indifferenti

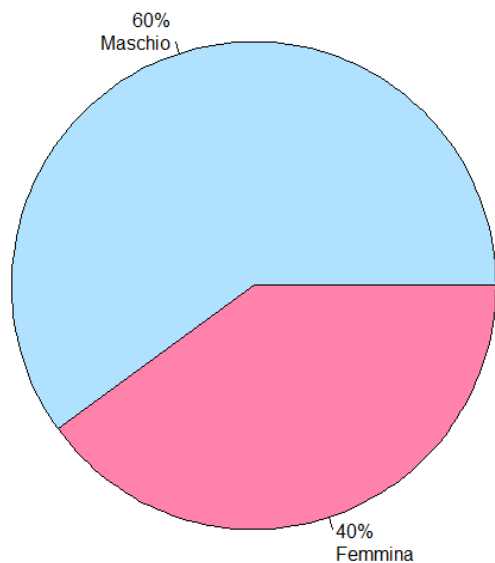


Figura 18: Percentuali Genere Indifferenti

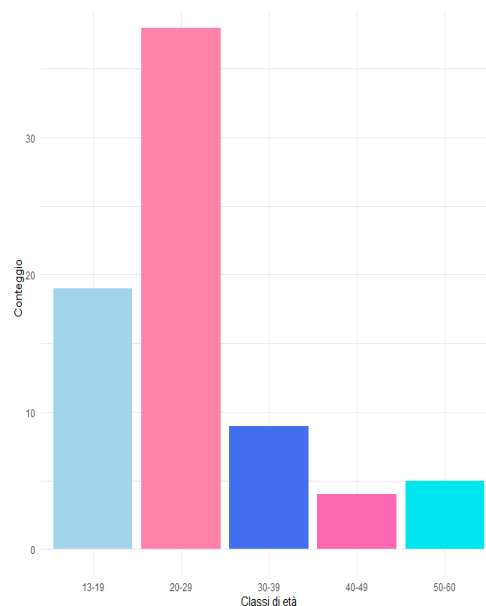


Figura 19: Distribuzione età Indifferenti

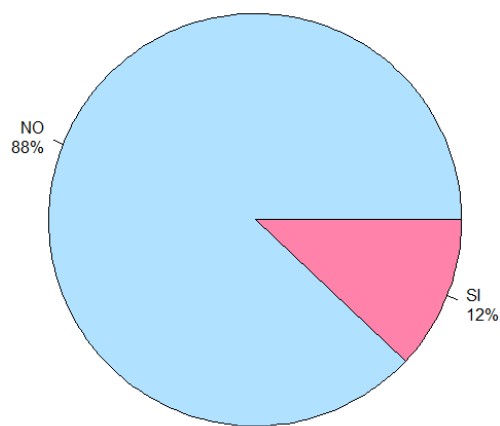
Il gruppo è formato da **75 unità statistiche**, composto per il **40% da femmine** e per il **60% da maschi**. Presenta un'età media di circa **27 anni (27.58)** e una **deviazione standard di 10.30**, ciò significa che la maggior parte delle età nel gruppo si trova entro un intervallo di circa **10 anni sopra o sotto la media**. In altre parole, la maggior parte delle età è compresa tra circa **17 e 37 anni (27 ± 10)**. È il gruppo **più eterogeneo** per quanto riguarda l'età; infatti, gli altri due gruppi hanno come deviazione standard 8 e 4: la deviazione standard misura la dispersione dei dati rispetto alla media: una deviazione standard più alta indica una maggiore variabilità tra le età dei membri del gruppo.

Tabella 6: Mediane dei Fan di Chiara, Indignati e Indifferenti

| | <i>Fan Chiara</i> | <i>Indignati</i> | <i>Indifferenti</i> |
|---------------------------|-------------------|------------------|---------------------|
| Conoscenza Chiara | 3 | 3 | 2 |
| Uso Social | 5 | 5 | 5 |
| Opinione Prima | 4 | 3 | 3 |
| Opinione Dopo | 3 | 2 | 2 |
| Influenza Scandali | 2 | 4 | 2 |

*Nota: la tabella riporta solo le mediane delle variabile misurate all'interno del gruppo (107, 73 e 75 unità).
Cluster ottenuti con legame Completo, Metrica Gower e metodo non gerarchico*

Il gruppo si caratterizza per avere una conoscenza dell'influencer minore rispetto agli altri due gruppi. Non conoscendo molto Chiara, è coerente il fatto che la mediana della variabile `opinione_prima` sia pari a 3, ovvero "Indifferente". Anche in questo caso si può osservare un peggioramento riguardo l'opinione che passa da avere una mediana pari a 3, ad averla uguale a 2 ("Negativa"), anche se gli scandali non hanno avuto un grande impatto su quest'ultima infatti la mediana pari a 2 corrisponde alla modalità "Poco":



| Min | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
|-------|---------|--------|-------|---------|-------|
| 1,000 | 1,000 | 1,000 | 1,507 | 2,000 | 5,000 |

Tabella 7: Propensione Acquisto Dopo Indifferenti

Nota: la tabella riporta i valori relativi alla variabile "propensione_acquisto_dopo" del gruppo Indifferenti

Figura 20: Percentuali Propensione di Acquisto Prima Indifferenti

Dal grafico a torta risulta che **solo il 12% sarebbe stato propenso ad acquistare prima degli scandali**. Questa percentuale è rimasta **invariata** in seguito agli scandali, infatti la mediana della variabile propensione_acquisto_dopo è 1, con una media pari a 1,507, rapportata ad uno scenario binario è del 12,6 %.

La *Figura 21* mostra quanto i vari fattori proposti possano influenzare la decisione di acquisto dei prodotti Chiara Ferragni del gruppo Indifferenti.

La maggior parte hanno una mediana pari a 1 tranne per la qualità che ha mediana pari a 2. Si può affermare che il gruppo non dà grande importanza ai fattori, probabilmente perché non acquisterebbe un prodotto Chiara Ferragni in nessun caso.

Dunque è verosimile pensare che **nessun prodotto/strategia potrebbe far cambiare idea al gruppo**.

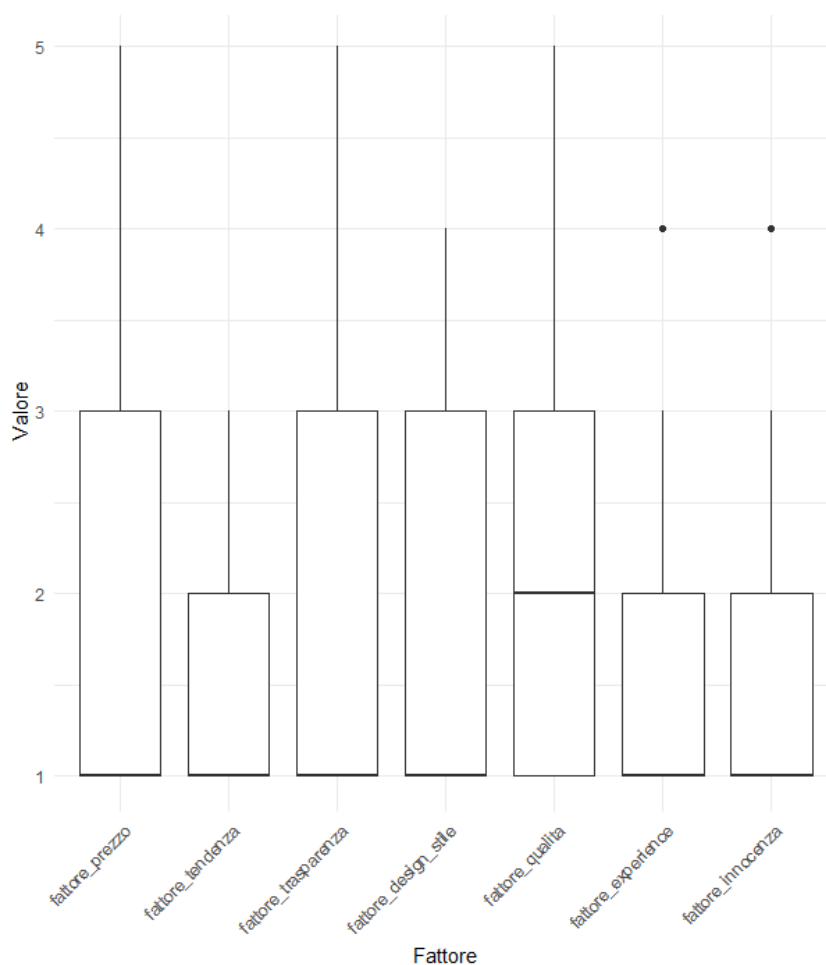


Figura 21: BoxPlot fattori Indifferenti

Alla luce di questa analisi, il gruppo “*Fan di Chiara*” risulta essere interessante per poter sviluppare una strategia di marketing volta all'acquisto perché si delinea come il target migliore in quanto ha un'opinione dopo migliore e una propensione di acquisto dopo maggiore rispetto agli altri gruppi (*Tabella 3 e Tabella 6*).

Per questo proseguiamo applicando la Regressione Logistica su questo gruppo per poter individuare quali variabili influenzano sulla probabilità di avere una propensione di acquisto alta.

3.2.6 Regressione Logistica

La **regressione logistica** è una tecnica statistica utilizzata per modellare la probabilità di un evento binario. È un modello di regressione utilizzato per predire l'esito di una variabile dipendente binaria basata su una o più variabili indipendenti. La funzione logistica, o sigmoide, è utilizzata per mappare qualsiasi valore reale in un intervallo compreso tra 0 e 1, che rappresenta la probabilità dell'evento di interesse.

Nella nostra analisi l'evento di interesse, e quindi la variabile dipendente, è “**Avere una propensione di acquisto alta**”. Nel nostro questionario la variabile “*propensione_acquisto_dopo*” viene misurata tramite una scala che va da **1** a **5**, questa tecnica statistica viene utilizzata per predire l'esito di una variabile dipendente binaria; perciò, la nostra variabile dipendente binaria sarà:

- **1, PROPENSIONE DI ACQUISTO DOPO ALTA,**
se il valore della variabile “*propensione_acquisto_dopo*” è maggiore o uguale a quattro
- **0, PROPENSIONE DI ACQUISTO DOPO BASSA,**
se il valore della variabile “*propensione_acquisto_dopo*” è minore a quattro

La formula della nostra regressione logistica (*propensione_acquisto_dopo*) è:

$$\left(\sum_{i=1}^4 (PC_i \times w_i) \right) + \text{acquisto_fatto_prima_fc21} \times w_5 + \text{propensione_acquisto_prima_fc21} \times w_6$$

Dove:

- PC_i sono le componenti principali estratte durante l'ACP (3.2.3)
- “*acquisto_fatto_prima_fc2*” e “*propensione_acquisto_prima_fc2*” sono variabili dicotomiche dummy trasformate in fattori. Rispettivamente valutano:
 - o Se il soggetto ha mai acquistato prodotti Chiara Ferragni
 - o Se il soggetto era propenso ad acquistare prodotti Chiara Ferragni prima dei scandali
- w è il peso di ogni variabile

Dalla Regressione Logistica risulta che:

- Al diminuire del valore di PC2 (alta opinione dopo), la probabilità di avere una propensione di acquisto dopo aumenta. Si tratta quindi in rapporto inverso.
- Chi ha fatto un acquisto prima ha più probabilità di rifarlo rispetto a chi non lo ha fatto

| | <i>Estimate</i> | <i>Std.Error</i> | <i>Z value</i> | <i>P value</i> |
|--|-----------------|------------------|----------------|-------------------|
| <i>Intercept</i> | -4.967211 | 1.047880 | -4.740 | $2,13e^{-0,6***}$ |
| <i>PC1</i> | -0.378164 | 0.353839 | -1.069 | 0.28518 |
| <i>PC2</i> | -1.216386 | 0.431745 | -2.817 | 0.00484 ** |
| <i>PC3</i> | 0.065101 | 0.459701 | 0.142 | 0.88738 |
| <i>PC4</i> | 0.005322 | 0.347813 | 0.015 | 0.98779 |
| <i>acquisto_fatto_prima_fc21</i> | 0.501116 | 0.822863 | 0.609 | 0.54253 |
| <i>propensione_acquisto_prima_fc21</i> | 2.246876 | 0.999079 | 2.249 | 0.02452 * |

Tabella 8: Regressione Logistica

Interpretazione OR

Intercept= 0.006962543, è un valore molto basso significa che senza l'intervento delle variabili l'evento non si verificherebbe

PC2 = 0.296299072, è inferiore a 1, indicando che un aumento di PC2 è associato a una diminuzione della probabilità dell'evento. Le probabilità che l'evento di interesse si verifichi diminuiscono del 31,5% ad ogni aumento unitario della PC2 ($1 - 0.296 = 0.704$).

propensione_acquisto_prima_fc21= 9,458, è maggiore di 1 indica un effetto positivo. In questo caso, l'aumento unitario della variabile provoca un aumento della probabilità che l'evento di interesse si verifichi di circa 845,8% ($9.458 - 1 = 8.458$).

| | OR |
|---------------------------------------|-------------|
| <i>Intercept</i> | 0.006962543 |
| <i>PC1</i> | 0.685117951 |
| <i>PC2</i> | 0.296299072 |
| <i>PC3</i> | 1.067267126 |
| <i>PC4</i> | 1.005335895 |
| <i>acquisto_fatto_prima_fc2</i> | 1.650561742 |
| <i>propensione_acquisto_prima_fc2</i> | 9.458139370 |

Tabella 9: Valori OR

4. Risultati

Tramite le **analisi descrittive** (3.2.1) abbiamo potuto delineare un profilo della nostra popolazione composta da **255 individui**.

Il **67,06%** è composto da **femmine**, il **32,55%** da **maschi**, e lo **0,39%** di individui che ha selezionato l'opzione "**Altro**" (*Figura 1*).

La maggior parte dei soggetti rientra nella fascia di età **21-30**, con una **media di 25 anni** circa e una **deviazione standard di 8,10**, ciò significa che l'età della maggioranza degli individui rientra nell'intervallo di 8 anni sopra o sotto la media, quindi tra i **17 e i 33 anni** (*Figura 2*). Più nello specifico:

- **77** rientrano nella fascia **16-20**
- **147** rientrano nella fascia **21-30**
- **16** rientrano nella fascia **31-40**
- **6** rientrano nella fascia **41-50**
- **9** rientrano nella fascia **51-60**

Dalla **Tabelle delle Frequenze relative** (*Tabella 1*) emerge esserci una conoscenza discreta di Chiara Ferragni, dove la moda corrisponde alla modalità "**Abbastanza**" (**35%**), anche se la risposta più comune alla richiesta di quanto spesso il soggetto è esposto ai contenuti dell'influencer è "**Molto**" (**24%**) e i $\frac{3}{4}$ della popolazione affermano di conoscere la sua vita pubblica e le sue attività imprenditoriali, rispettivamente il **76%** e il **75%** hanno risposto "**Sì**" alle domande. Comunque possiamo affermare che quasi tutti sono a conoscenza degli scandali legati a Chiara Ferragni (il **97,25%** ha risposto "**Sì**").

Per quanto riguarda l'influenza degli scandali le risposte sono distribuite con "**Abbastanza**" e "**Poco**" entrambe al **29%**, suggerendo una moderata influenza degli scandali.

Il **45%** dichiara di aver avuto un'opinione neutra riguardo Chiara Ferragni prima degli scandali, seguita subito dopo dalla modalità "**Positiva**" con il **34%**.

La situazione **dopo gli scandali** mostra un **36%** per un'opinione **Negativa**, ma la moda rimane la modalità "**Indifferente**" con il **43%**.

Le risposte alla domanda su quanto la notizia che anche le uova di Pasqua fossero oggetto di una truffa ha influenzato l'opinione, sono distribuite con un:

- **15%** per la modalità "**Molto Peggiorata**"
- **42%** per la modalità "**Peggiorata**"
- **43%** per la modalità "**Indifferente**"
- **0%** per la modalità "**Migliorata**"
- **0%** per la modalità "**Molto Migliorata**"

Il **56%** non percepisce l'influencer più "**umana**" dopo gli scandali.

Il **79%** della nostra popolazione non ha effettuato un acquisto dal brand, anche se il **71%** ha dichiarato di non essere stato propenso ad acquistare suoi prodotti prima degli scandali.

Pertanto, una piccola percentuale di soggetti che sarebbe stato propenso ad acquistare, non lo ha mai fatto.

Per quanto riguarda la propensione di acquisto dopo gli scandali:

- il **55%** della popolazione ha dichiarato di non essere affatto propenso ad acquistare ("**Per Niente**")
- **16%** per "**Poco**"
- **21%** per "**Abbastanza**"
- **7%** per "**Molto**"
- **0%** per "**Completamente**"

Dalle frequenze relative dei fattori che influenzano il comportamento di acquisto, emerge che quello più importante per la popolazione è la **qualità**.

L'88% dichiara di utilizzare **quotidianamente i social**, un dato che ci aspettavamo in quanto la maggior parte dei soggetti della popolazione è **under 30**.

Tramite un'**Analisi delle Componenti Principali** (3.2.3) abbiamo potuto individuare **4 componenti principali** (*Figura 6*) che andassero a spiegare 12 variabili di nostro interesse, dove:

- La prima componente principale (**PC1**) rappresenta una misura della "**Conoscenza di Chiara**"
- La seconda componente principale (**PC2**) rappresenta una misura del "**Influenza degli scandali**"
- La terza componente principale (**PC3**) rappresenta un indicatore del "**Uso dei Social**"
- La quarta componente principale (**PC4**) rappresenta una misura della "**Percezione dell'influencer**"

Successivamente abbiamo svolto un'Analisi Cluster (3.2.4) con lo scopo di creare dei cluster basati sulle caratteristiche principali identificate dalle componenti principali precedentemente estratte.

Abbiamo prima clusterizzato seguendo un **Metodo Gerarchico** con un **legame completo** e la **metrica Gower**.

Abbiamo individuato la presenza di **tre gruppi** all'interno della popolazione (Figura 7), ma non essendo completamente soddisfatti della distanza tra i gruppi abbiamo applicato una seconda clusterizzazione, questa volta scegliendo un **Metodo Non Gerarchico**, con un **legame completo** e la **metrica Gower**. Dal confronto dei risultati di questi due metodi (Figura 8) siamo giunti alla conclusione che il secondo, fosse il metodo **più idoneo** per la nostra analisi.

Il primo cluster, da noi chiamato "**Fan di Chiara**", è formato da **107 individui**, composto per l'**85.05%** da **femmine**, **14.95%** da **maschi** (Figura 10).

La maggior parte delle età degli individui rientra nella fascia **21-30**, e presenta un'età **media di 25 anni** con una **deviazione standard di 8.04** (Figura 11).

Il gruppo si caratterizza per (tabella 2):

- **Mediana 3** (conoscenza discreta dell'influencer sulla variabile "*conoscenza_chiara*")
- **Mediana 4** (opinione positiva, nella variabile "*opinione_prima*")
- **Mediana 3** (valutazione indifferente, nella variabile "*opinione_dopo*")
- **Mediana 2** (valutazione bassa, nella variabile "*influenza_scandali*")

Il gruppo si divide quasi a metà tra chi era propenso a fare un acquisto precedentemente gli scandali (43,93%) e chi no (56,07%) (Figura 12).

La mediana e la media della variabile "Propensione di Acquisto Dopo" sono 3 e 2,355 che rapportata in percentuale in uno scenario binario, corrisponderebbe al 33,88%; possiamo quindi affermare che il cambiamento è di lieve entità.

La Figura 13 mostra i fattori che condizionerebbero il comportamento di acquisto. Dai box-plot appare che quelli più rilevanti sono il Prezzo, la Qualità, la Trasparenza, l'Experience e il Design (tutte con mediana=3).

Il secondo cluster, da noi chiamato "**Indignati**", è formato da **73 unità**, composto per il **68,49%** da **femmine**, per il **30,14%** da **maschi**, e da un **1,37%** che ha selezionato "**Altro**" (Figura 14).

È il gruppo più giovane, con un'età **media di 22 anni** e **deviazione standard di 3.7**. La maggior parte degli individui rientra nella fascia **21-30** (Figura 15).

Il gruppo si caratterizza per (Tabella 4):

- **Mediana 3** (conoscenza discreta dell'influencer sulla variabile "*conoscenza_chiara*")
- **Mediana 3** (opinione neutrale, nella variabile "*opinione_prima*")
- **Mediana 2** (opinione negativa, nella variabile "*opinione_dopo*")
- **Mediana 4** (valutazione alta, nella variabile "*conoscenza_scandali*")

Solo il **26,03%** dei membri era stato propenso ad un acquisto prima degli scandali, percentuale diminuita in seguito: infatti, la mediana e la media della variabile "*propensione_acquisto_dopo*", rispettivamente 1,000 e 1,384, in percentuale corrisponderebbe a circa il **10%**.

La Figura 17 mostra i fattori più importante per il gruppo che condizionerebbero il comportamento di acquisto. Questi sono il **Prezzo, la Qualità e il Design**.

Il terzo cluster, da noi chiamato "**Indifferenti**", è composto da un **60% di maschi** e un **40% da femmine** (Figura 18).

È il gruppo **più anziano**, con una media di 27 anni e deviazione standard di 10,30: si dimostra essere il gruppo **più eterogeneo** per quanto riguarda l'età.

Si caratterizza per:

- **Mediana 2** (conoscenza bassa dell'influencer sulla variabile "*conoscenza_chiara*")
- **Mediana 3** (opinione neutrale, nella variabile "*opinione_prima*")
- **Mediana 2** (opinione negativa, nella variabile "*opinione_dopo*")
- **Mediana 2** (valutazione bassa, nella variabile "*conoscenza_scandali*")

Applichiamo la **Regressione Logistica (3.2.6)** al primo cluster, in quanto è il target più idoneo sul quale costruire una strategia di marketing, avendo la propensione di acquisto dopo più alta.

L'evento di interesse, ovvero la variabile dipendente binaria, è la variabile "*propensione_acquisto_dopo*" trasformata in 1, (**ALTA**), se il valore della variabile "*propensione_acquisto_dopo*" è maggiore o uguale a quattro; 0, (**BASSA**), se il valore della variabile "*propensione_acquisto_dopo*" è minore a quattro.

5. Conclusioni

In conclusione, il gruppo denominato “Fan di Chiara” si è dimostrato essere il profilo di cliente ideale per essere un target potenziale sul quale costruire una strategia di mercato volta all’acquisto.

Questo perché è il gruppo che si è dimostrato più “fedele” all’influencer: aveva un’opinione positiva prima degli scandali, ha un’opinione dopo neutra e una maggior propensione all’acquisto in seguito a quest’ultimi.

Un’efficiente strategia di marketing potrebbe essere basata sui fattori ritenuti importanti dal gruppo che condizionerebbero il comportamento di acquisto, ovvero:

- **Il Prezzo e la Qualità**
incentivando l’acquisto tramite offerte e promozioni, cose che il brand ha già utilizzato durante occasioni specifiche come il Black Friday
- **La trasparenza nelle comunicazioni**
*Si potrebbe decidere, di adottare una politica di garanzia e resi semplici per rassicurare i clienti sulla qualità del prodotto.
Sarebbe essenziale utilizzare una comunicazione trasparente, dando ai clienti delle informazioni chiare e dettagliate in merito alla produzione, l’origine dei prodotti e materiali.*
- **L’esperienza di Acquisto**
Si potrebbe offrire un servizio clienti reattivo e disponibile tramite vari canali e personalizzare l’esperienza al massimo offrendo raccomandazioni basate sugli acquisti precedenti.
- **Il Design**
La cura del packaging per renderlo attraente e funzionale potrebbe migliorare l’esperienza complessiva del cliente. Infatti, il brand possiede già un packaging e un design distintivo, ovvero l’iconico occhio stilizzato che appare in ogni confezione firmata Chiara Ferragni. Incorporare l’iconico logo in modo prominente sul packaging renderebbe immediatamente riconoscibile il prodotto, aiutando a rafforzare l’identità del brand. Creando inoltre un’esperienza di unboxing unica e coinvolgente con l’aggiunta di messaggi personalizzati oppure adesivi e piccoli omaggi, si sorprenderebbe i clienti all’apertura del pacco e sarebbero incentivati a condividere il momento dell’apertura sui vari social, dando visibilità al brand.

Appendice

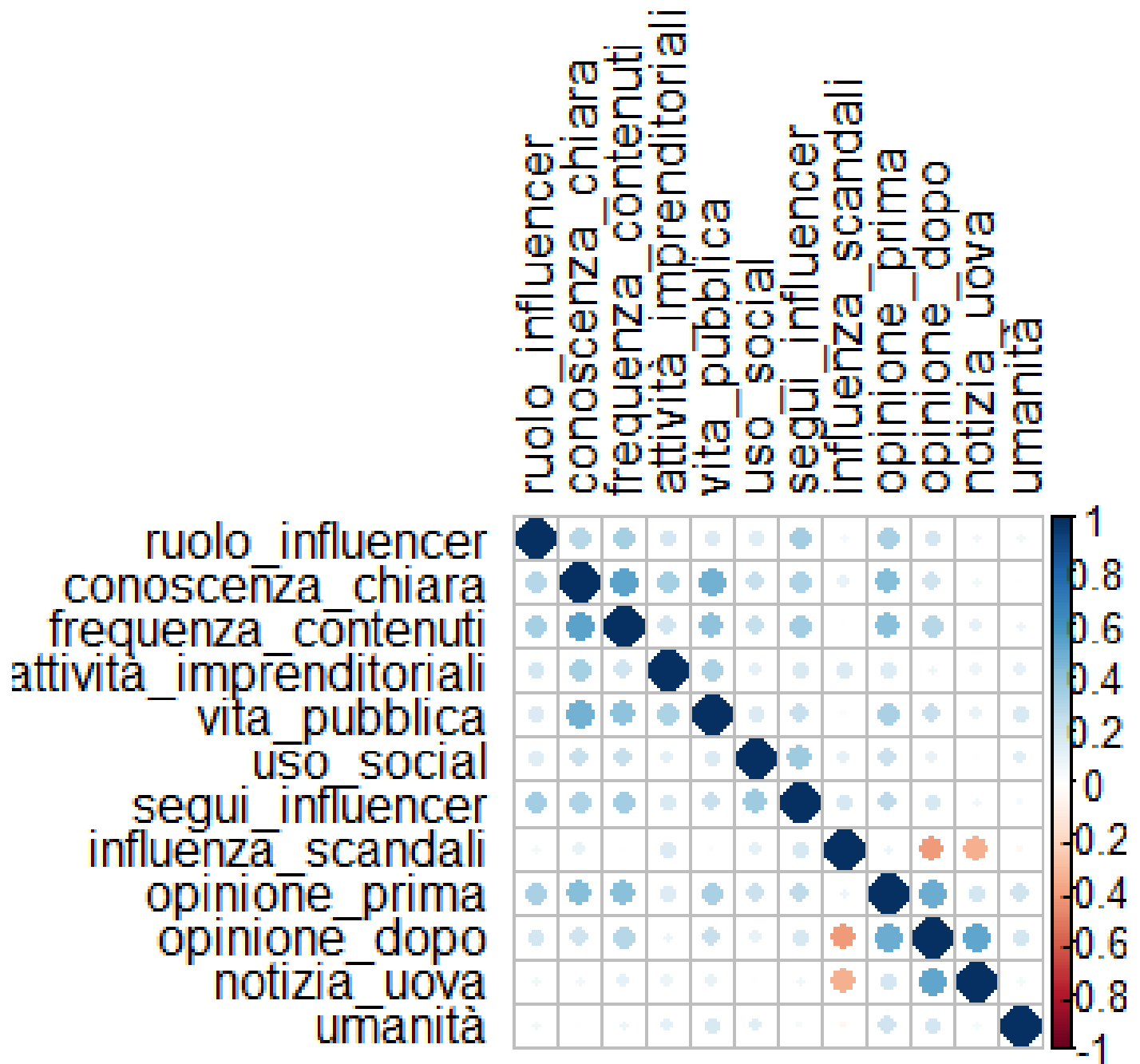


Figura 1: Matrice di Correlazione

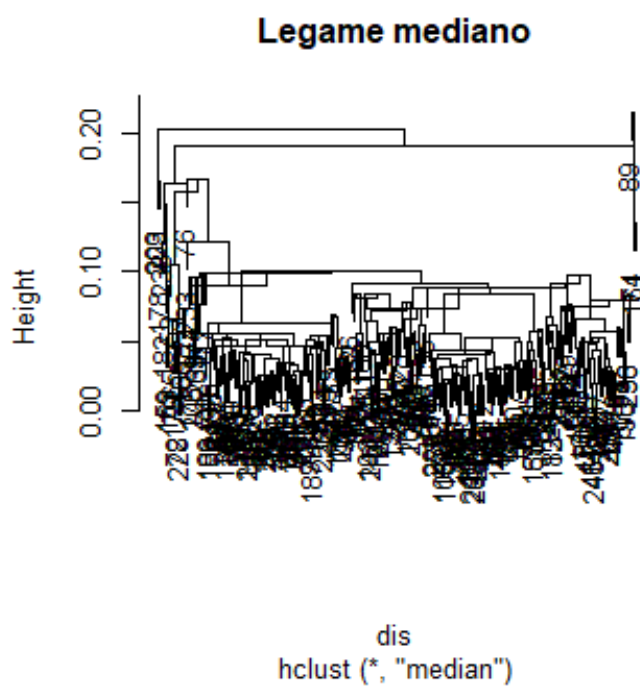
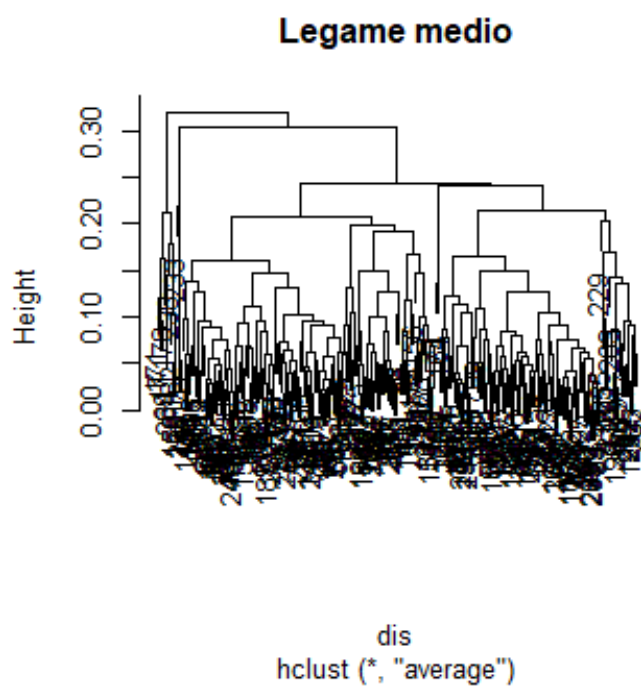
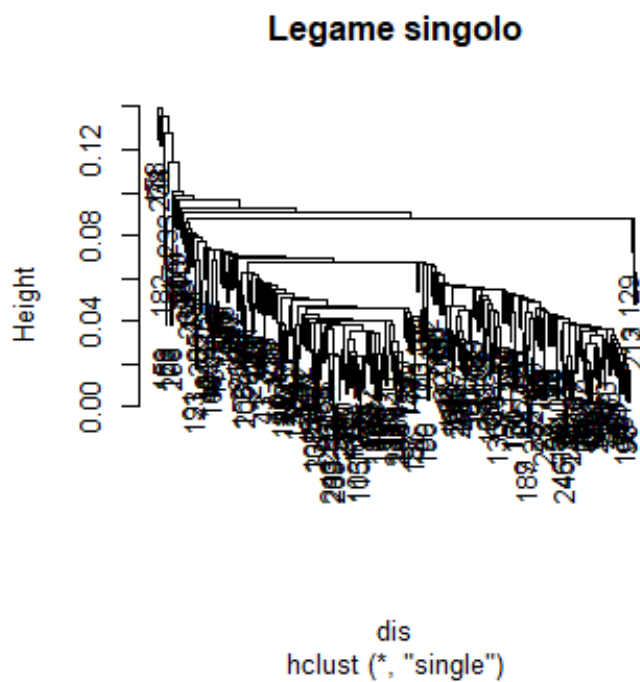
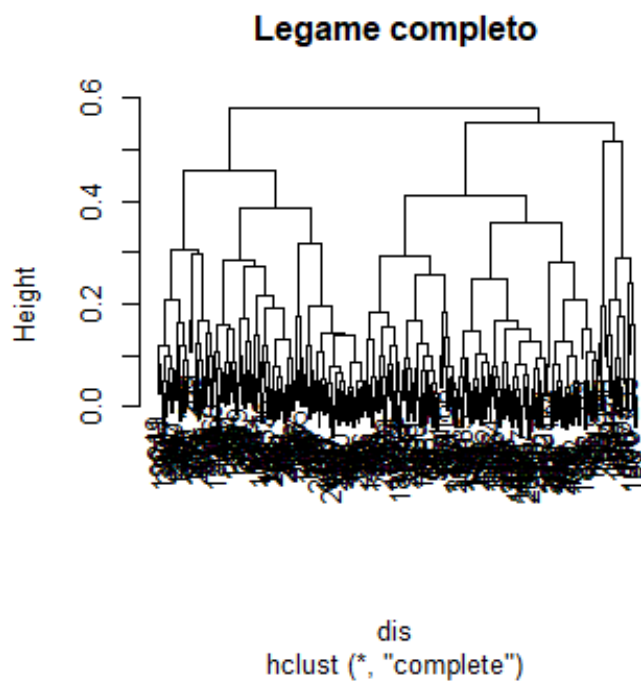


Figura 2: Confronto Dendogrammi

Legame completo

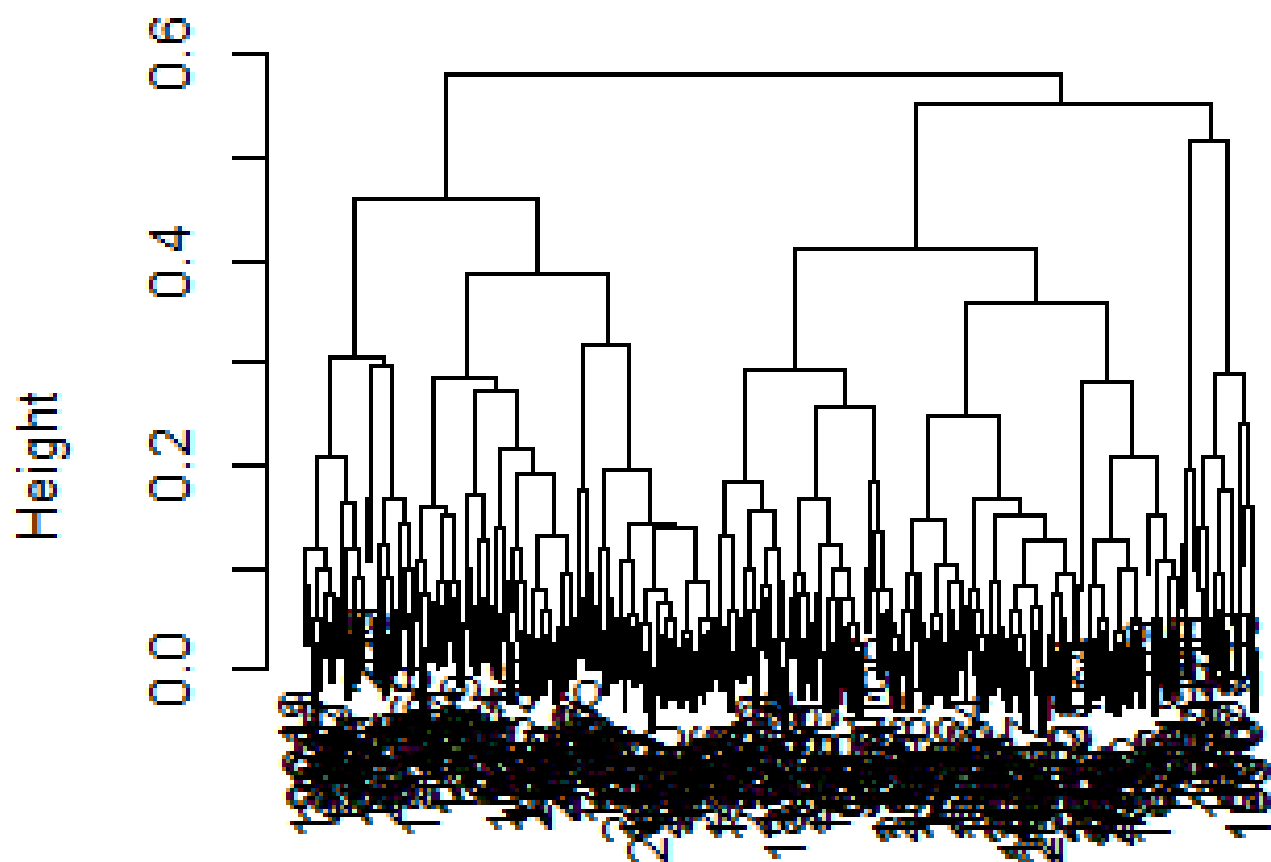


Figura 2bis: Legame Completo

Legame singolo

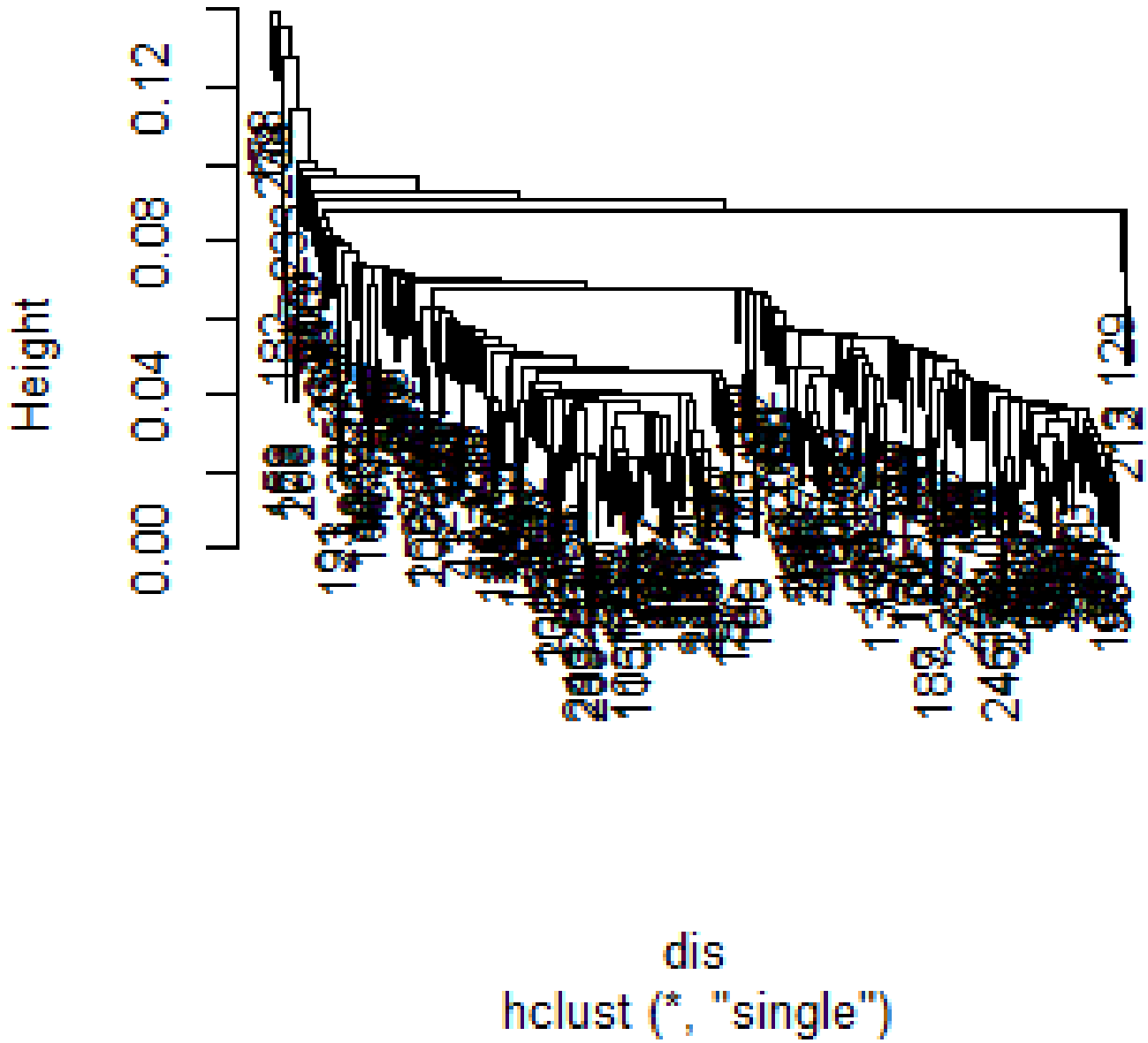


Figura 2ter: Legame Singolo

Legame medio

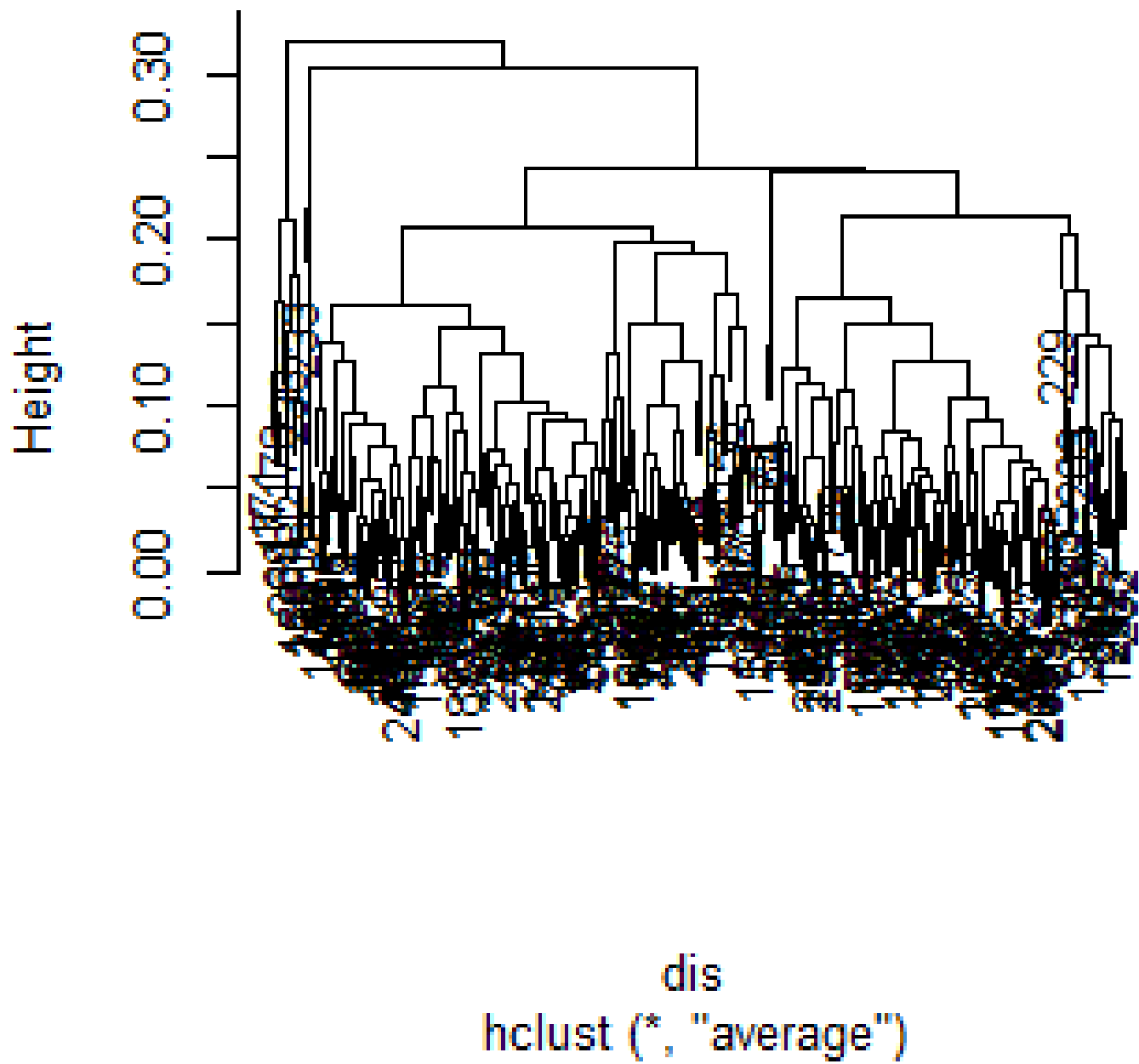


Figura 2quater: Legame Medio

Legame mediano

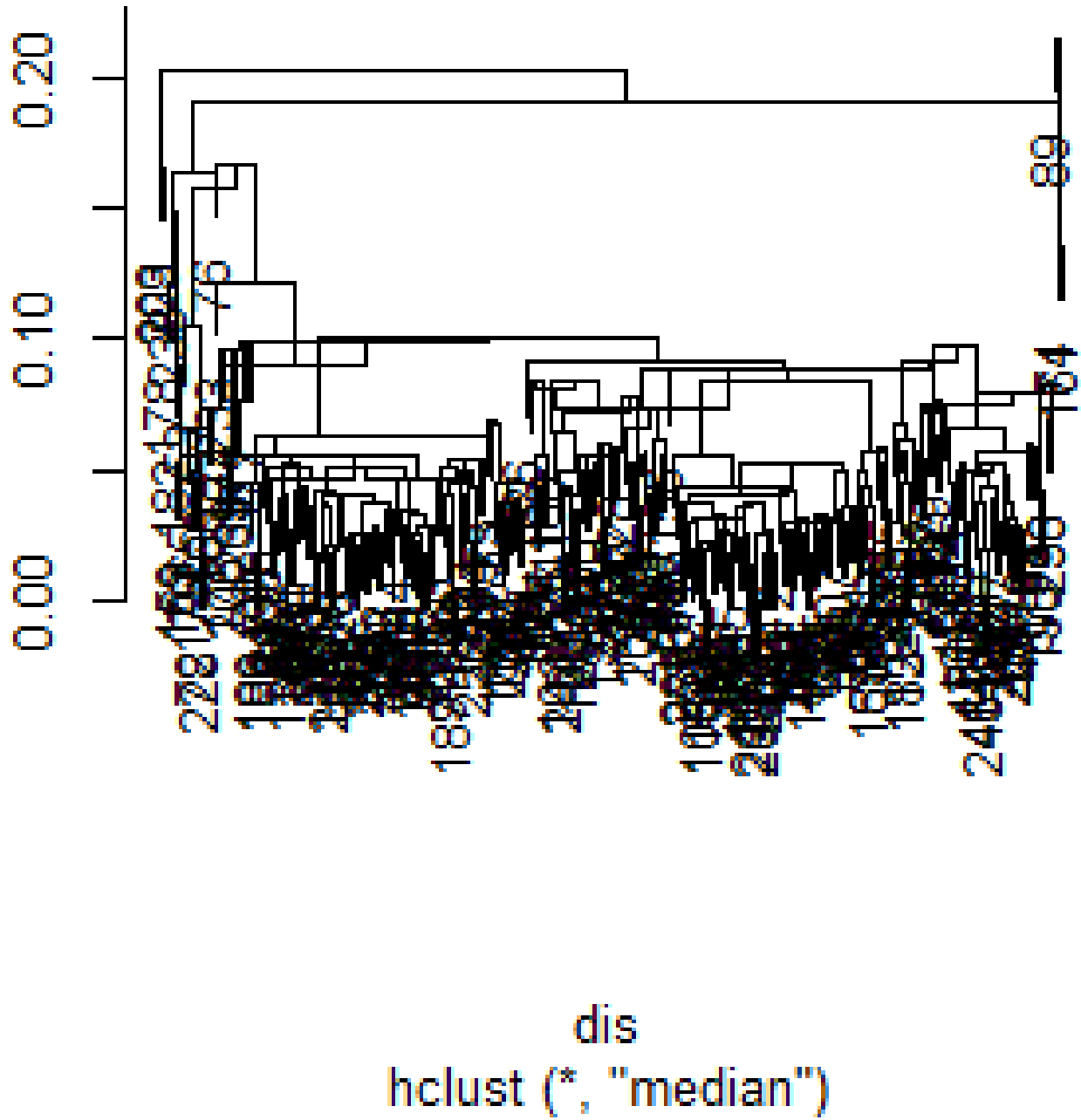


Figura 2quinquies: Legame Mediano

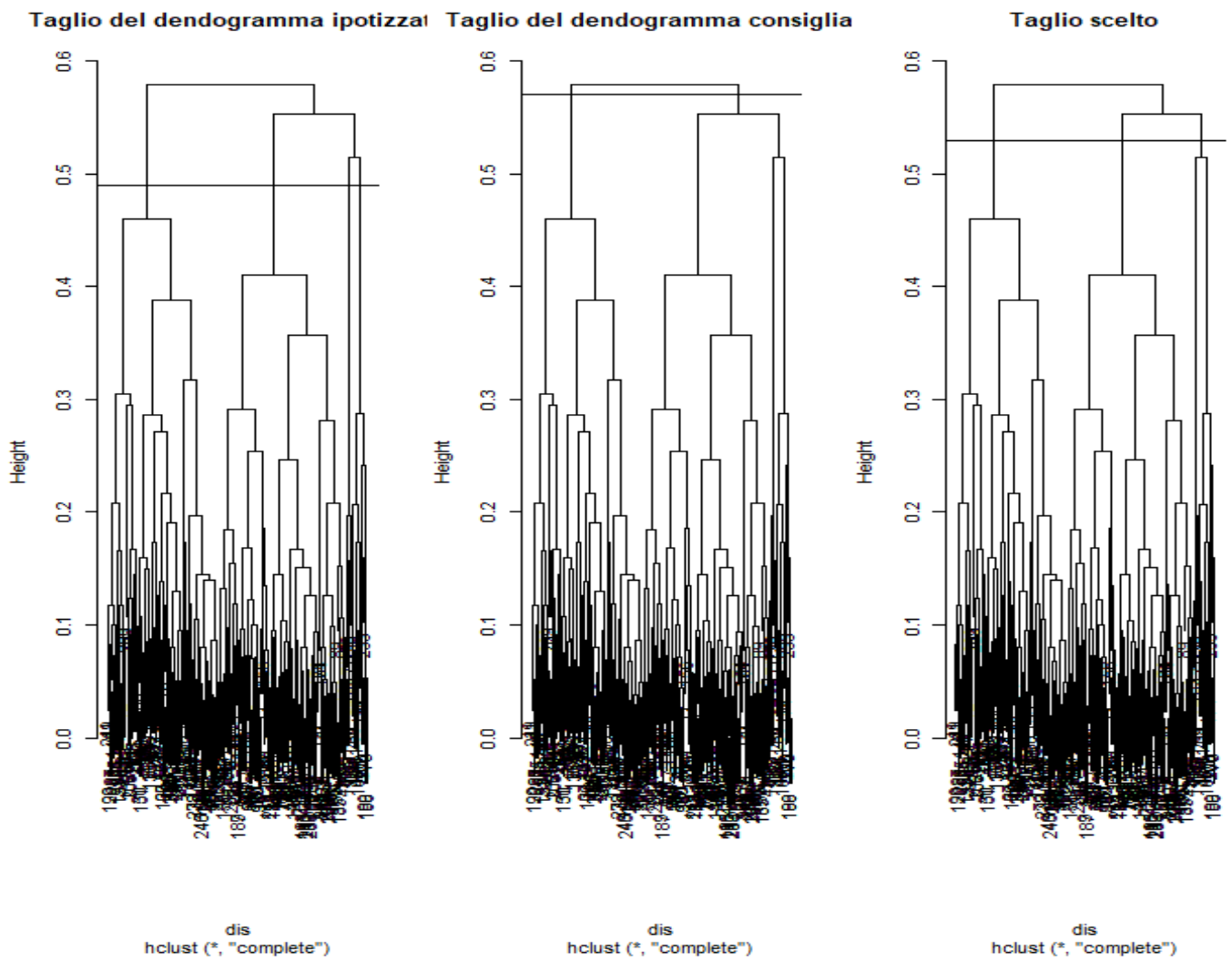


Figura 3: Scelta del Numero dei cluster

Taglio del dendrogramma ipotizzato

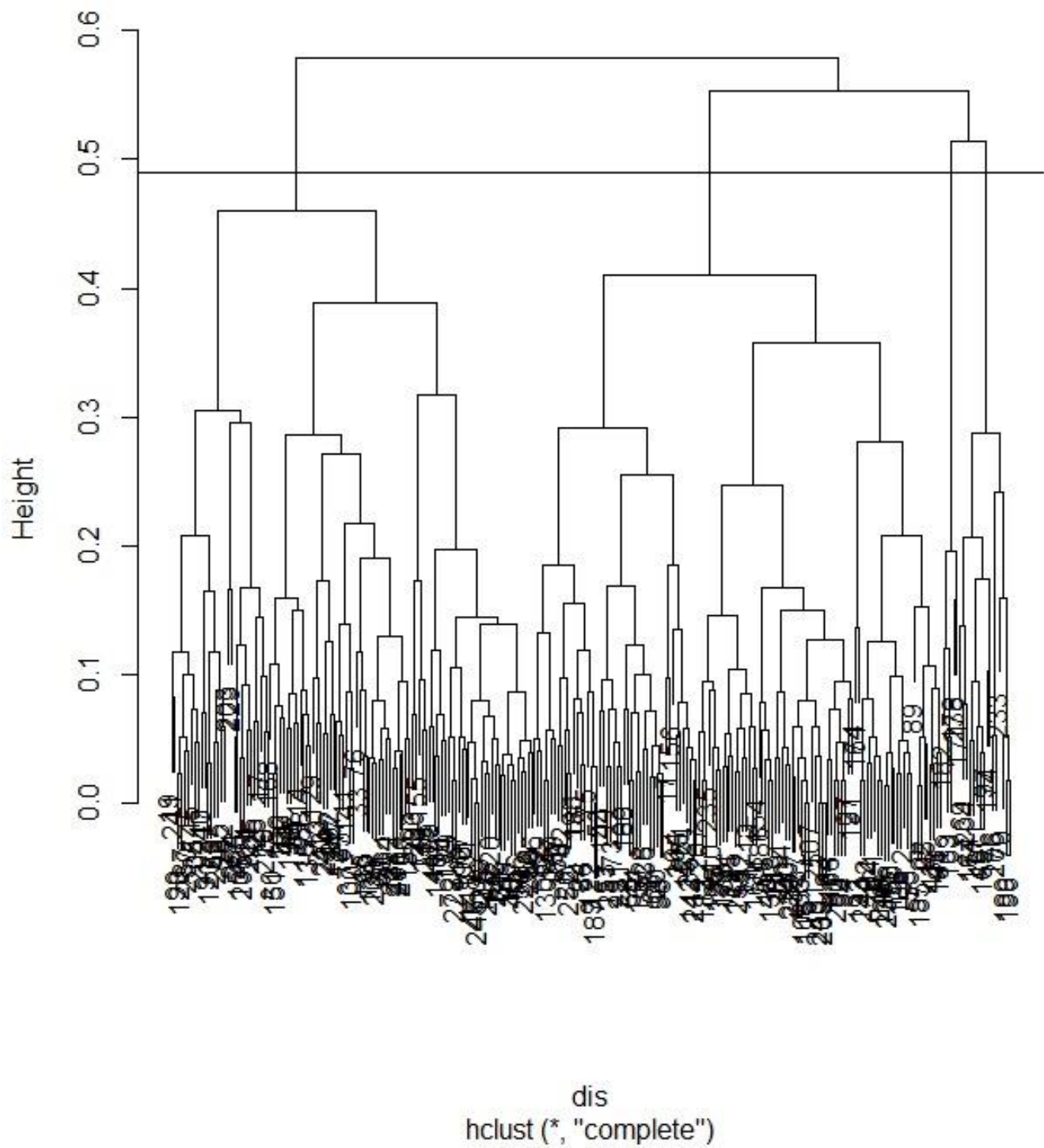


Figura 3bis: taglio “ad occhio”

Taglio del dendrogramma consigliato

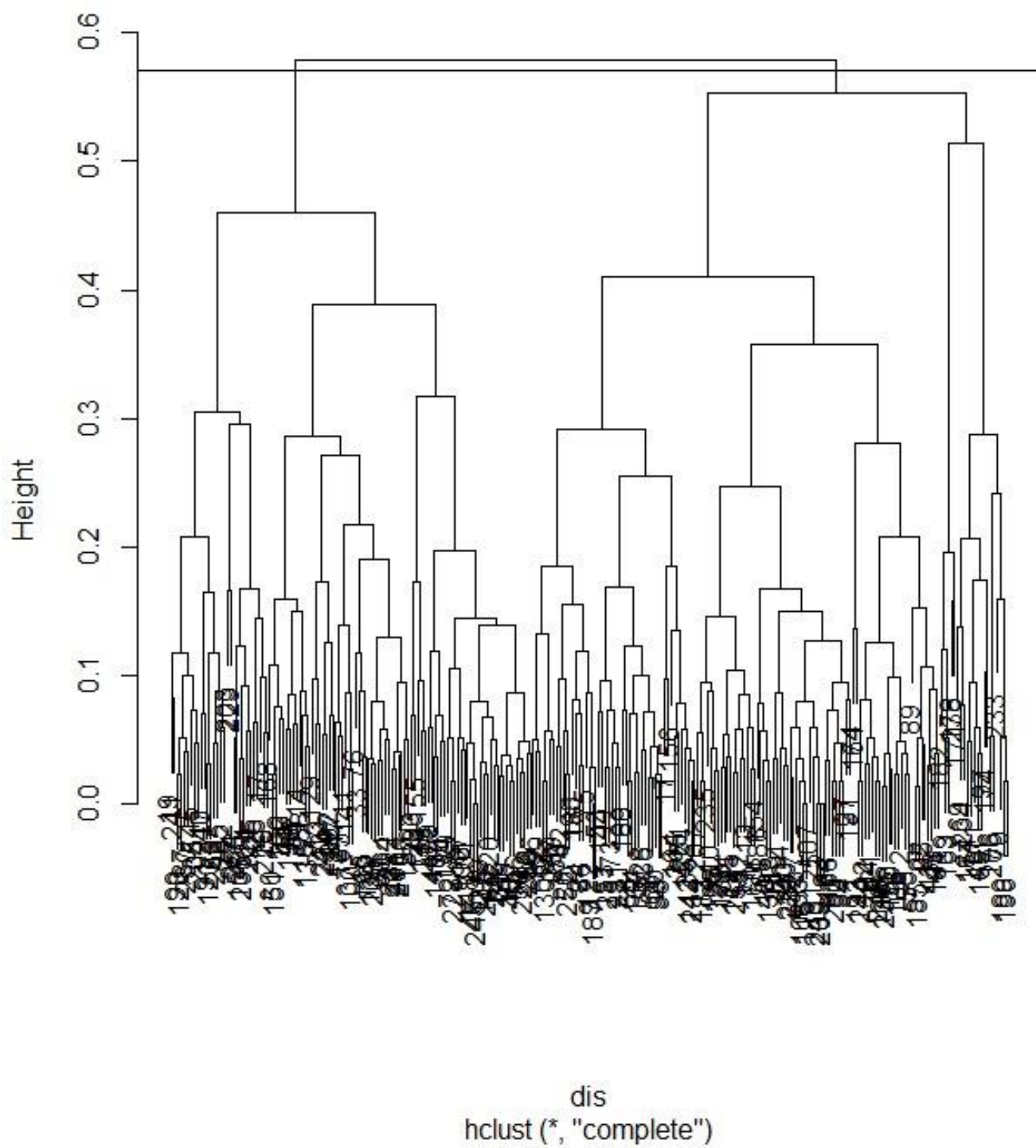


Figura 3ter: taglio consigliato (NBClust)

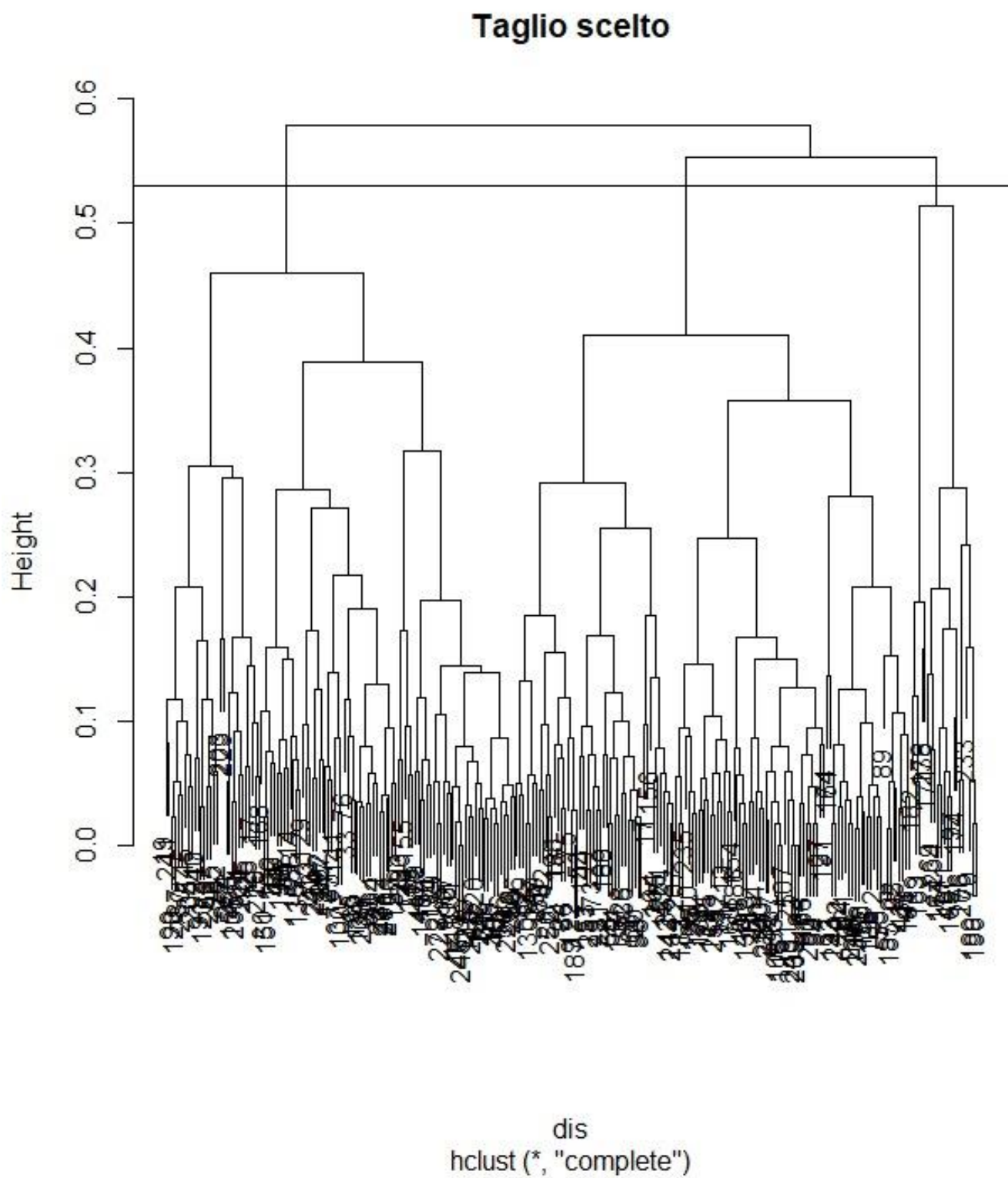


Figura 3ter: taglio scelto