**Tesina di Analisi di mercato**

**TIM**

**Chap 1: I DATI**

**1.1 Il Questionario**

Per la stesura del nostro questionario ci siamo basati su un questionario che va ad indagare la customer satisfaction, ed abbiamo indagato sia la percezione, sia le attese riguardo il servizio fornito da TIM mobile ai clienti. Inizialmente abbiamo chiesto, con una domanda filtro iniziale se l’intervistato fosse o meno cliente TIM in modo da fare domande specifiche in base alla risposta.

Oltre a questa grande domanda filtro abbiamo indagato altri aspetti per una eventuale segmentazione del mercato come ad esempio da che piattaforma vengono monitorati i consumi, per andare successivamente a vedere se ci fossero differenze nette tra le varie categorie in termini di risposte. Per quelli che si sono dichiarati clienti TIM abbiamo indagato se avessero mai usufruito del servizio di call center così da vedere anche la soddisfazione dei clienti riguardo a quel servizio TIM.

Abbiamo poi chiesto le aspettative e le percezioni su tutti gli argomenti che ritenevamo utili per ricreare successivamente le dimensioni di indagine. Avevamo inizialmente inserito anche una parte riguardante la soddisfazione dei clienti rispetto ai negozi TIM, per valutare la dimensione di indagine Aspetti Tangibili, ma che poi abbiamo omesso nel questionario dato che portava lo stesso ad essere troppo lungo e dopo una prima somministrazione del questionario risultavano pochi i clienti che avevano usufruito anche di questo servizio.

Abbiamo poi fatto delle domande generiche agli intervistati come chiedere un giudizio generale sulla soddisfazione TIM per vedere se le risposte precedenti fossero coerenti.

Agli intervistati che non si sono dichiarati clienti TIM, abbiamo fatto qualche domanda riguardo ai motivi per cui non fossero clienti, per far sì che in una possibile campagna di marketing di TIM fossero tenuti in considerazione questi aspetti e l’azienda possa così provare a guadagnare altre “fette” di mercato. Infine abbiamo indagato l’anagrafica di tutti quelli che hanno compilato il questionario per avere delle informazioni sul campione e sulla sua distribuzione riguardo a variabili come Sesso, Età e Professione.

E’ stato detto nell’introduzione del questionario che i dati sarebbero stati trattati in modo da non violare le leggi della privacy. Il questionario è stato somministrato via WEB con il metodo CAWI ed è stato somministrato a chiunque si è offerto di rispondere, con la sola limitazione che gli intervistati avessero almeno 15 anni, in modo tale che abbiano coscienza di gestire una scheda TIM. Dopo che il questionario è stato diffuso abbiamo ricevuto 360 risposte, e non abbiamo avuto grossi problemi riguardo ad eventuali errori di compilazione del questionario stesso. I problemi maggiori si sono verificati quando indagavamo di che provincia fosse l’intervistato chiedendo la sigla della provincia di appartenenza. A quella domanda molte persone hanno risposto con l’intero nome della città oppure con altre imperfezioni, ma che non si sono rivelate significative dato che abbiamo aggiustato tutte le sigle errate, correggendole.

Un’altra domanda che ha dato risposte non soddisfacenti è stata quella dove chiedevamo la percentuale di consumi ideale fatto 100 il totale; molti non hanno fatto una suddivisione che portava alla somma 100 e quindi per l’analisi eventuale di quella domanda siamo costretti ad eliminare 109 osservazioni. Per tutte le altre risposte non abbiamo avuto problemi, quindi non è stata esclusa nessuna osservazione dal campione.

**1.2 Presentazione dei dati raccolti**

In questo capitolo andremo a presentare i dati raccolti da chi scrive tramite il questionario presentato nel capitolo 1.1 e somministrato secondo i metodi lì descritti.

Verranno presentate diverse statistiche descrittive al fine di delineare il campione e la sua composizione.

**I Numeri e Limiti**

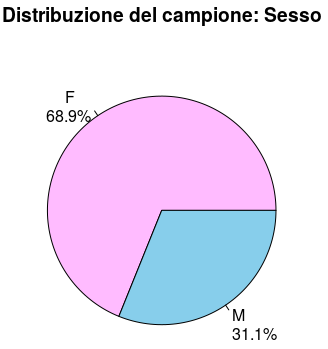
Analizzando i dati nel complesso notiamo subito la loro “grande mole”; inatteso era infatti il risultato, a nostro avviso ottimo, di 360 intervistati.

Possiamo imputare questo risultato forse al tipo di somministrazione (CAWI).

In fase di analisi però è stato possibile notare la lama a doppio taglio che è la somministrazione esclusiva del questionario via internet, sono molti infatti i limiti del nostro studio e cercheremo di farli presente tempestivamente ogni volta questi sopraggiungeranno.

**Distribuzione del campione**

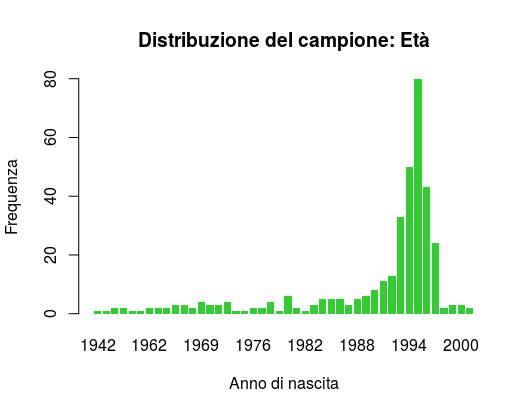
Il primo grafico che andiamo a presentare è quello della distribuzione del campione stratificato per sesso.



Basta una semplice occhiata per identificare il primo limite del campione, questo infatti è composto al 68.9% di femmine ed al 31.1% di maschi, dato distorto rispetto alla media nazionale di 52.2% e 47.8% rispettivamente di femmine e maschi con età superiore ai 15 anni ( fonte “CIA World Factbook”, dati relativi a 07/2014).

Questa e le distorsioni che tratteremo in seguito non ci permetteranno di considerare il campione rappresentativo della popolazione italiana, procederemo comunque con l’analisi a fini didattici.

Il secondo grafico che presenteremo è quello della distribuzione del campione stratificata per età.



Dal grafico riportato possiamo notare uno sbilanciamento del campione verso destra, ad indicare la prevalenza di intervistati molto giovani.

La prima informazione che possiamo estrapolare quindi da questo grafico è l’alto apporto di giovani, se considerassimo la fascia di età che va dal 1994 al 1996 avremmo inserito in quest’unica categoria il 48% degli intervistati.

Procedendo dalla moda delle osservazioni (anno 1995) verso sinistra osserviamo una graduale diminuzione di osservazioni, fino ad arrivare alla fascia più estrema, quella che va dal 1942 al 1961, dove abbiamo diverse annate mancanti e molte con una sola osservazione; è confortante invece il dato che ricaviamo osservando i dati dal 1961 al 2001, qui ritroviamo almeno una osservazione per ogni anno ( ad eccezione del 1963 e 1973).

*Frequenze delle osservazioni per anno:*

*1942 1951 1956 1959 1960 1961 1962 1964 1965 1966 1967 1968 1969 1970 1971*

*1 1 2 2 1 1 2 2 2 3 3 2 4 3 3*

*1972 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987*

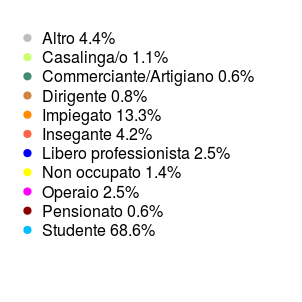
*4 1 1 2 2 4 1 6 2 1 3 5 5 5 3*

*1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001*

*5 6 8 11 13 33 50 80 43 24 2 3 3 2*

Procederemo ora con l’analisi della distribuzione del campione stratificato per professione.





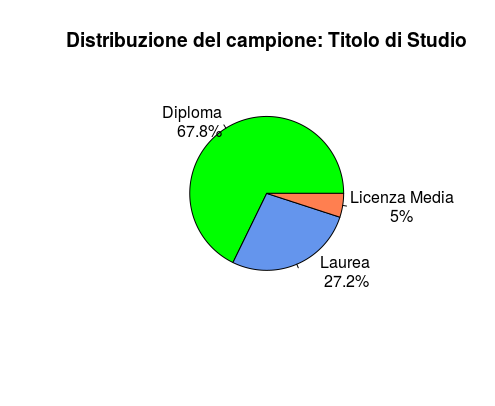
Quello riportato sopra è il grafico a torta rispetto alle professioni svolte dagli intervistati.

Una conseguenza logica a quanto detto per la distribuzione del campione rispetto all’età è che qui ritroviamo un eccesso di studenti, col 68.6% di presenza nel campione. Anche questo bias, come gli altri individuati fino ad ora, è imputabile al tipo di somministrazione del questionario; era attesa infatti una componente di studenti forte dal momento che il questionario è stato diffuso via web e più in particolare tramite i

profili Facebook di chi scrive. La gran parte della rete di persone raggiunte in questa maniera ha all’incirca l’età degli analisti ed uno spaccato di età e professione dimostrano come il tipo di diffusione abbia influito sul tipo di soggetti raggiunti.

Fatta eccezione per gli studenti, i soggetti intervistati si distribuiscono equamente tra le varie categorie.

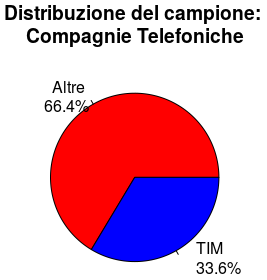
Il prossimo grafico descriverà e dividerà il nostro campione in base al titolo di studio.



Parlando di titolo di studio la situazione nel nostro campione è quella riportata nel grafico.

Confrontando i dati da noi raccolti con i dati ISTAT del 4° trimestre 2016 notiamo delle differenze significative: mentre nessuno dei nostri intervistati afferma di non essere in possesso di alcun titolo di studio o della sola licenza elementare, ISTAT riporta un dato nazionale del 18.4% per questa categoria; contro il nostro 5% di intervistati in possesso della licenza media come massimo titolo di studio vediamo il dato di ISTAT differenziarsi di diversi punti percentuali, il 32.5% della popolazione appartiene a questa categoria; di conseguenza registriamo un sovrannumero di diplomati e laureati, 67.8% e 27.8% nel campione contro il 35.7% e 13.5% di ISTAT.

Indagheremo ora il grado di penetrazione di TIM nel mercato.

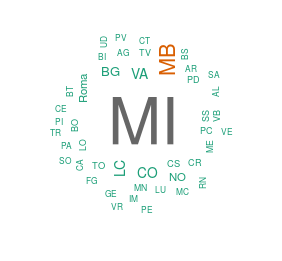


Secondo il rapporto Agcom TIM è leader del mercato della telefonia mobile con il 32.3% di quota di mercato. Il nostro studio ritrova una situazione analoga a quanto riportato da Agcom, si registra infatti una quota del 33.6%.

Riportiamo ora i dati riguardanti la provincia di residenza dei soggetti intervistati.

**Distribuzione del campione:**

**Provincia di residenza**



*Frequenza provincie di residenza:*

*AG AL AR BG BI BO BS BT CA*

*1 1 1 15 1 3 1 1 1*

*CE CO CR CS CT FG GE IM LC*

*1 17 3 4 1 1 1 1 18*

*LO LU MB MC ME MI MN NO PA*

*2 1 41 1 2 166 2 10 1*

*PC PD PE PI PV RN Roma SA SO*

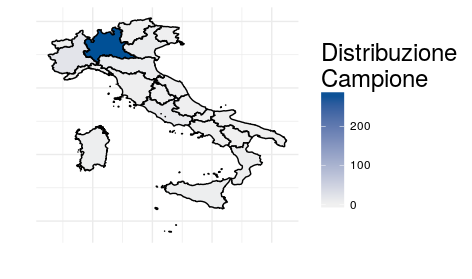
*3 3 1 1 1 1 10 1 1*

*SS TO TR TV UD VA VB VE VR*

*3 4 1 3 1 17 3 1 1*

Il wordcloud in alto denota la grande centralità di Milano; sono infatti 166 le osservazioni che riportano come provincia di residenza MI, il 46.1% del totale.

In basso riportiamo tutte le frequenze delle varie provincie di residenza degli intervistati.



Dalla cartina riportata sopra invece osserviamo come la gran parte dei soggetti intervistati venga dalla Lombardia ed in generale la quasi totalità del campione ha residenza nel nord Italia, entrambe le osservazioni sottolineano un altro bias, che come gli altri è riconducibile al tipo di diffusione del questionario.

**Chap 2: LE PRIME ANALISI**

**2.1 Dimensioni di Indagine**

In questo capitolo parleremo delle variabili del nostro dataset, derivante appunto dal questionario presentato nel capitolo 1.1, e di come queste verranno suddivise per facilitare le analisi.

Il questionario ci restituisce un dataset composto da 360 osservazioni e 44 variabili, che noi andremo a suddividere in 4 macroclassi.

Sono infatti state individuate da noi 4 delle 5 canoniche dimensioni di indagine, in particolare possiamo ritrovare domande che indagano l’affidabilità di TIM, la capacità di risposta, l’empatia e la capacità di rassicurazione. Come dicevamo manca l’indagine sugli aspetti tangibili che in un primo momento abbiamo pensato di inserire con domande riguardanti i negozi TIM, ma che dopo un breve periodo di test del questionario è risultata una parte controproducente: poche persone erano state in negozio e chi doveva rispondere alle domande in questione lamentava una durata del questionario eccessiva, che portava poi alla poca attenzione nelle domande successive.

Per quanto riguarda le osservazioni invece, considereremo di interesse solo i soggetti che riportano di essere clienti TIM e che hanno usufruito del servizio di call center, questo perché la dimensione “Capacità di risposta” è indagabile solo tramite le domande fatte esclusivamente ai fruitori del servizio di supporto telefonico.

**La suddivisione**

Il nostro obiettivo a questo punto dell’analisi era raggruppare le diverse domande del questionario per formare le dimensioni “Affidabilità”,”Capacità di risposta”,”Empatia”,”Capacità di rassicurazione”.

Innanzitutto dobbiamo distinguere tra percezioni ed aspettative, per ognuna delle categorie creeremo quindi una coppia di nuove variabili.

* Per rappresentare l’***affidabilità*** di TIM *percepita* dai clienti useremo 3 variabili:

“Come valuti l’uso di dispositivi con sim TIM?”

·Stabilità della rete 1 2 3 4 5

·Trasparenza tariffe 1 2 3 4 5

·Precisione riguardo ai consumi 1 2 3 4 5

La media tra il valore assegnato ad ognuno di questi aspetti per l’ i-esimo intervistato sarà l’i-esimo valore della macroclasse “Affidabilità percezione”.

Per rappresentare l’*affidabilità* di TIM nelle *aspettative* dei clienti useremo 3 variabili:

“Quanto sono importanti i seguenti per te i seguenti aspetti riguardo all’uso di cellulari con sim TIM?”

·Stabilità della rete 1 2 3 4 5

·Trasparenza tariffe 1 2 3 4 5

·Precisione riguardo ai consumi 1 2 3 4 5

La media tra il valore assegnato ad ognuno di questi aspetti per l’ i-esimo intervistato sarà l’i-esimo valore della macroclasse “Affidabilità aspettativa”.

* Per rappresentare la ***capacità di risposta*** di TIM *percepita* dai clienti useremo 2 variabili:

“Come valuti il servizio di call center declinato nei seguenti aspetti?”

·Competenza dell’operatore 1 2 3 4 5

·Efficacia nella risoluzione dei problemi 1 2 3 4 5

La media tra il valore assegnato ad ognuno di questi aspetti per l’ i-esimo intervistato sarà l’i-esimo valore della macroclasse “Capacità di risposta percezione”.

Per rappresentare la *capacità di risposta* di TIM nelle *aspettative* dei clienti useremo 2 variabili:

“Quanto sono importanti per te i seguenti aspetti in un servizio di call center?”

·Competenza dell’operatore 1 2 3 4 5

·Efficacia nella risoluzione dei problemi 1 2 3 4 5

La media tra il valore assegnato ad ognuno di questi aspetti per l’ i-esimo intervistato sarà l’i-esimo valore della macroclasse “Capacità di risposta aspettativa”.

* Per rappresentare l’***empatia***di TIM *percepita* dai clienti useremo 3 variabili:

“Come valuti il servizio di call center declinato nei seguenti aspetti?”

·Cortesia del personale 1 2 3 4 5

“Come valuti l’uso di dispositivi con sim TIM?”

·Servizi paralleli (es. agevolazioni cinema ecc.) 1 2 3 4 5

“Quanto sei d’accordo con le seguenti affermazioni?”

·Mi sento parte della famiglia TIM 1 2 3 4 5

La media tra il valore assegnato ad ognuno di questi aspetti per l’ i-esimo intervistato sarà l’i-esimo valore della macroclasse “Empatia percezione”.

Per rappresentare l’*empatia* di TIM nelle *aspettative* dei clienti useremo 3 variabili:

“Quanto sono importanti per te i seguenti aspetti in un servizio di call center?”

·Cortesia del personale 1 2 3 4 5

“Quanto sono importanti i seguenti per te i seguenti aspetti riguardo all’uso di cellulari con sim TIM?”

·Servizi paralleli (es. agevolazioni cinema ecc.) 1 2 3 4 5

“Quanto sei d’accordo con le seguenti affermazioni?”

·E’ importante parte della famiglia TIM 1 2 3 4 5

La media tra il valore assegnato ad ognuno di questi aspetti per l’ i-esimo intervistato sarà l’i-esimo valore della macroclasse “Empatia aspettativa”.

* Per rappresentare la ***capacità di rassicurazione***di TIM *percepita* dai clienti useremo 3 variabili:

“Come valuti il servizio di call center declinato nei seguenti aspetti?”

·Semplicità di utilizzo 1 2 3 4 5

“Quanto sei d’accordo con le seguenti affermazioni?”

·Mi ritengo perfettamente informato delle iniziative di TIM 1 2 3 4 5

·Mi sento sicuro affidandomi a TIM 1 2 3 4 5

La media tra il valore assegnato ad ognuno di questi aspetti per l’ i-esimo intervistato sarà l’i-esimo valore della macroclasse “Capacità di rassicurazione percezione”.

Per rappresentare l*a capacità di rassicuazione* di TIM nelle *aspettative* dei clienti useremo 3 variabili:

“Quanto sono importanti per te i seguenti aspetti in un servizio di call center?”

·Semplicità di utilizzo 1 2 3 4 5

“Quanto sei d’accordo con le seguenti affermazioni?”

·E’ importante essere perfettamente informato delle iniziative di TIM 1 2 3 4 5

·E’ importante sentirsi sicuri con TIM 1 2 3 4 5

La media tra il valore assegnato ad ognuno di questi aspetti per l’ i-esimo intervistato sarà l’i-esimo valore della macroclasse “Capacità di rassicurazione aspettativa”.

**2.2 Gap analysis**

Una volta create le dimensioni di indagine distinte in aspettative e percezioni possiamo condurre una gap analysis, in particolare andremo ora a studiare il gap 5, quello tra percezioni ed aspettative del cliente, che è anche l’unico possibile soggetto di analisi non essendo in possesso di dati aziendali.

*Medie per dimensione*

*affidabilità\_P risposta\_P empatia\_P rassicurazione\_P*

*3.184783 3.016304 2.547101 2.684783*

*affidabilità\_A risposta\_A empatia\_A rassicurazione\_A*

*4.289855 4.423913 2.971014 3.702899*

Quelle riportate sopra sono le medie per ogni dimensione individuata, per ottenere il gap 5 basterà fare la differenza tra la media delle percezioni e le aspettative per ogni macrogruppo.

*GAP 5*

*affidabilità* *risposta empatia rassicurazione*

*-1.105072 -1.407609* *-0.423913*  *-1.018116*

Dai risultati della gap analysis riportati sopra non otteniamo informazioni incoraggianti, infatti la percezione di ogni dimensione indagata delude le aspettative in media. Questo è un segnale molto forte per l’azienda, infatti questi risultati sottolineano un deciso disallineamento tra percezioni ed aspettative soprattutto per quanto riguarda la capacità di risposta e l’affidabilità, ma in generale nessuna percezione supera le aspettative.

Questo scostamento è indice di una bassa qualità percepita dal cliente e per questo ognuno di questi gap andrebbe colmato con politiche mirate per ognuna di queste categorie.

**Chap 3: ANALISI DELLE COMPONENTI PRINCIPALI**

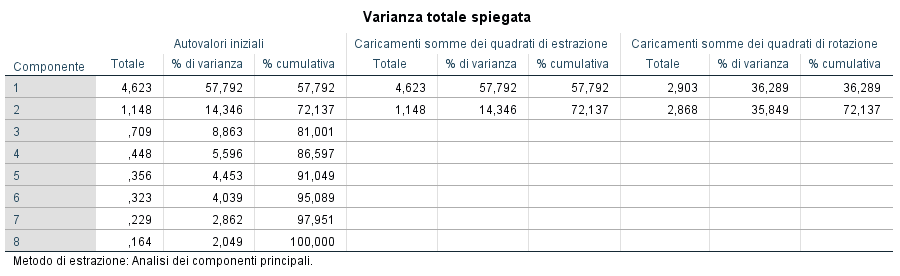
Per procedere con l’analisi delle componenti principali, innanzitutto abbiamo individuato le variabili su cui eseguire l’analisi: le 4 dimensioni di indagine descritte nel capitolo 2.1. Di queste 4 categorie abbiamo fatto sia le percezioni che le attese, in base alle risposte del nostro questionario. Abbiamo così ottenuto 8 variabili con cui abbiamo iniziato a fare l’analisi delle CP. La prima tabella che andiamo a visualizzare è quella che contiene i due test di Bartlett e di KMO, per vedere se ha senso procedere con l’analisi delle CP.

Il test di Bartlett ha come ipotesi nulla che le variabili siano incorrelate tra loro, ma come si vede dal p-value del nostro test andiamo a rifiutare l’ipotesi nulla e quindi diciamo che le variabili sono correlate tra loro.

|  |  |  |
| --- | --- | --- |
| **Test di KMO e Bartlett** | | |
| Misura di Kaiser-Meyer-Olkin di adeguatezza del campionamento. | | ,850 |
| Test della sfericità di Bartlett | Appross. Chi-quadrato | 430,767 |
| gl | 28 |
| Sign. | ,000 |

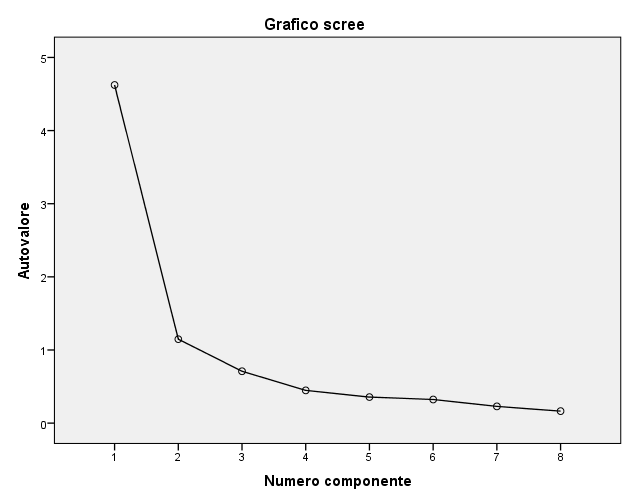
Poi osserviamo il risultato in tabella del test di KMO, che misura l’adeguatezza campionaria. Dato che la soglia del test di KMO è 0.700 e nel nostro caso risulta di 0.850, allora sia per questo risultato che per quello precedente, diciamo che è significativo procedere con un’analisi delle componenti principali.

Quindi iniziando con l’analisi, visualizzando la tabella della varianza spiegata che ci guida nella scelta delle CP.



Come si può vedere dalla tabella, abbiamo deciso di estrarre solamente 2 componenti principali, dato che solamente queste due spiegano più del 70% della varianza totale. In questo modo le nostre 8 variabili prese in considerazione per l’analisi siamo riusciti a riassumerle in due dimensioni che spiegano una buona parte di variabilità.

La scelta di prendere proprio due componenti principali è dovuta in parte allo scree plot, che ci mostra in maniera ben precisa che le CP ottimali da estrarre sono 2. Un altro motivo è legato al fatto che, solitamente si estraggono tutte le CP con auto valore maggiore di 1, e come possiamo vedere nella tabella sopra, le prime due componenti sono le uniche con un auto valore maggiore di 1.



Procediamo ora con la vera e propria analisi delle CP, dato che abbiamo verificato la presenza di tutti prerequisiti necessari.

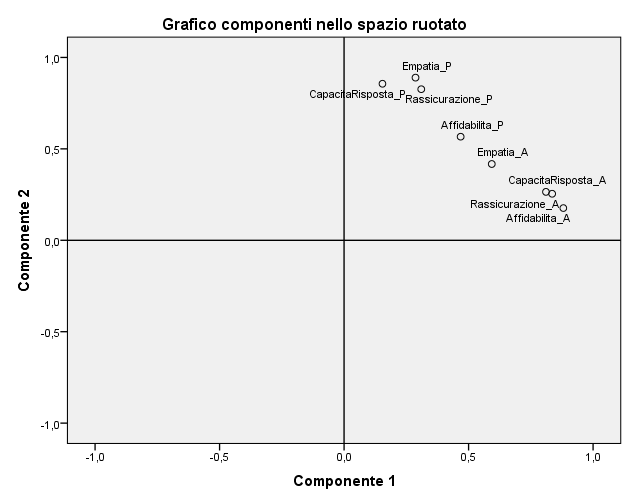
Per commentare i risultati delle componenti principali utilizziamo come output la matrice dei dati ruotati riportata qua sotto, perché nella nostra analisi abbiamo eseguito una rotazione con metodo varimax, che quindi ci restituisce i dati ruotati. Dalla tabella possiamo subito notare che c’è una forte distinzione tra le due componenti principali. Nella prima componente osserviamo che i valori più prossimi al valore 1 sono tutti quelli che fanno riferimento alle attese, nelle 4 dimensioni da indagare; se dovessimo dare allora un nome alla prima componente principale, potremmo dire che questa spiega molto bene le attese dei consumatori. Per quanto riguarda la seconda componente principale, si può subito notare che questa è caratterizzata principalmente dalle percezioni dei consumatori nel nostro campione. Infatti si nota che i valori più prossimi all’ uno sono tutti quelli che fanno riferimento alle percezioni e quindi se dovessimo nominare questa seconda CP, potremmo dire che spiega molto bene le percezioni dei clienti. Se vogliamo essere precisi in tutte e due le componenti c’è una variabile che è spiegata non all’80% come le altre ma a meno del 60%, nella prima componente sarebbe la variabile Empatia e nella seconda Affidabilità. Questo non cambia però ai fini del risultato statistico in quanto gli altri valori delle altre variabili nella CP sono comunque tutti inferiori al 60%, cosa che porta pieno sostegno alla nostra analisi delle componenti principali.

|  |  |  |
| --- | --- | --- |
| **Matrice dei componenti ruotatia** | | |
|  | Componente | |
| 1 | 2 |
| CapacitaRisposta\_P | 0,154 | 0,856 |
| Empatia\_P | 0,287 | 0,890 |
| Rassicurazione\_P | 0,310 | 0,826 |
| Affidabilita\_P | 0,468 | 0,566 |
| CapacitaRisposta\_A | 0,811 | 0,265 |
| Empatia\_A | 0,593 | 0,417 |
| Rassicurazione\_A | 0,835 | 0,254 |
| Affidabilita\_A | 0,880 | 0,176 |
| Metodo di estrazione: Analisi dei componenti principali.   Metodo di rotazione: Varimax con normalizzazione Kaiser. | | |
| a. Convergenza per la rotazione eseguita in 3 iterazioni. | | |

Possiamo ora procedere a calcolare i contributi di ogni variabile alla spiegazione delle due componenti principali. Per calcolare i contributi bisogna dividere i contributi assoluti nella tabella sopra per l’autovalore della CP corrispondente.

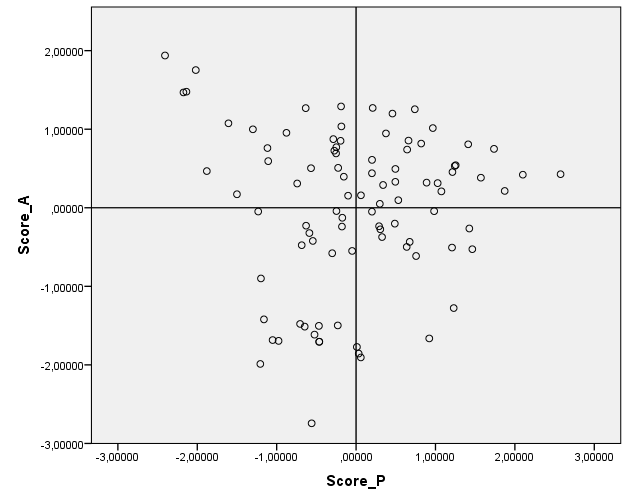
|  |  |  |
| --- | --- | --- |
| **Matrice dei contributi** | | |
|  | Contributi in % | |
| 1 | 2 |
| CapacitaRisposta\_P | 0,513392 | 63,84596 |
| Empatia\_P | 1,77876 | 68,94472 |
| Rassicurazione\_P | 2,077068 | 59,47409 |
| Affidabilita\_P | 4,739807 | 27,95448 |
| CapacitaRisposta\_A | 14,21426 | 6,101808 |
| Empatia\_A | 7,607044 | 15,15487 |
| Rassicurazione\_A | 15,09738 | 5,641214 |
| Affidabilita\_A | 16,76911 | 2,700093 |

Come si può vedere dalla tabella a sinistra, i contributi confermano le nostre proposte sulle componenti principali infatti vediamo come le variabili della percezione contribuiscono maggiormente a spiegare la seconda CP mentre le variabili delle attese contribuiscono maggiormente a spiegare la prima componente principale. In questo caso ci viene detto per esempio che la variabile Rassicurazione\_A contribuisce per il 15% alla spiegazione della prima componente principale.



Il grafico sopra riportato mette in evidenza la distinzione tra i due gruppi di variabili: le attese e le percezioni; si vede chiaramente dal grafico che i gruppi sono tra di loro ben distinti, segno che conferma la nostra analisi delle CP. Questo grafico rappresenta le variabili ruotate, infatti lo abbiamo plottato dopo aver eseguito la trasformazione varimax alle variabili.

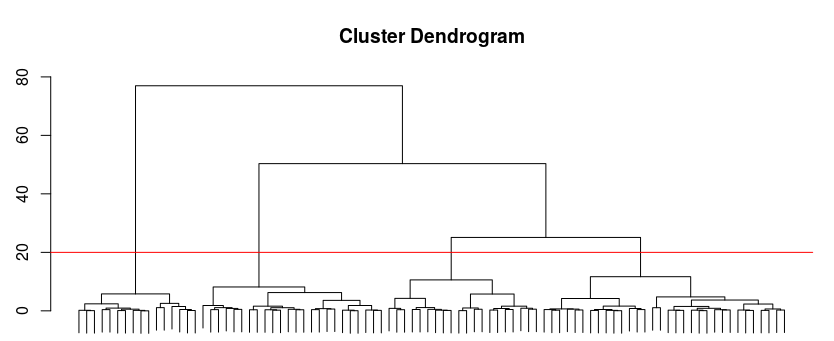
Nel grafico sottostante, invece, abbiamo plottato i due score delle due componenti principali per vedere come sono distribuiti gli intervistati all’interno del nostro campione e per vedere le loro valutazioni tra attese e percezioni. Ci sono parecchie persone che sono situate nel I quadrante dove si trovano le persone molto soddisfatte e che avevano delle grandi aspettative dal servizio TIM. Sono quindi i clienti più soddisfatti e con una visione dell’azienda più positiva e appagante. Ci sono anche delle persone che sono situate nel II e nel IV quadrante del grafico che sono, le prime le persone che avevano grandi aspettative ma che non sono soddisfatte mentre le seconde sono persone che non avevano grandi aspettative ma che sono rimaste comunque soddisfatte del servizio. Sicuramente l’azienda dovrà puntare di più sulle prime, quindi quelle insoddisfatte ma con grandi aspettative perché sono le persone che l’azienda più facilmente riuscirà a convincere del buon servizio TIM. Nel III quadrante invece ci sono le persone che sono poco soddisfatte e che non hanno grandi aspettative sull’azienda e sono, a nostro parere, le persone su cui puntare meno in un’eventuale proposta di marketing.



**Chap 4: SEGMENTAZIONE**

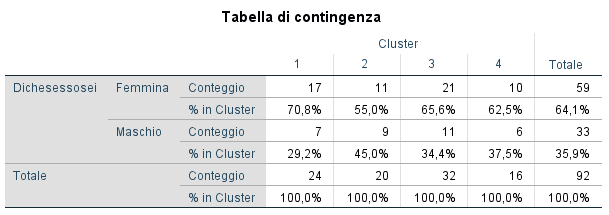
**4.1 Cluster Analysis**

Procediamo adesso con una cluster analysis gerarchica per andare a vedere la segmentazione nel mercato. Per la nostra cluster analysis siamo partiti dalle variabili ruotate che avevamo ottenuto nell’analisi delle CP, capitolo 3, e abbiamo ottenuto il seguente dendrogramma, da qui decideremo in quanti cluster dividere le nostre osservazioni. Per ottenere il seguente dendrogramma abbiamo usato la distanza euclidea, che è quella che ci sembrava modificasse e distorcesse meno i nostri dati. Per quanto riguarda il metodo di agglomerazione abbiamo usato quello di Ward per riuscire a massimizzare la devianza tra i gruppi e avere allo stesso tempo gruppi caratterizzati da una bassa variabilità.

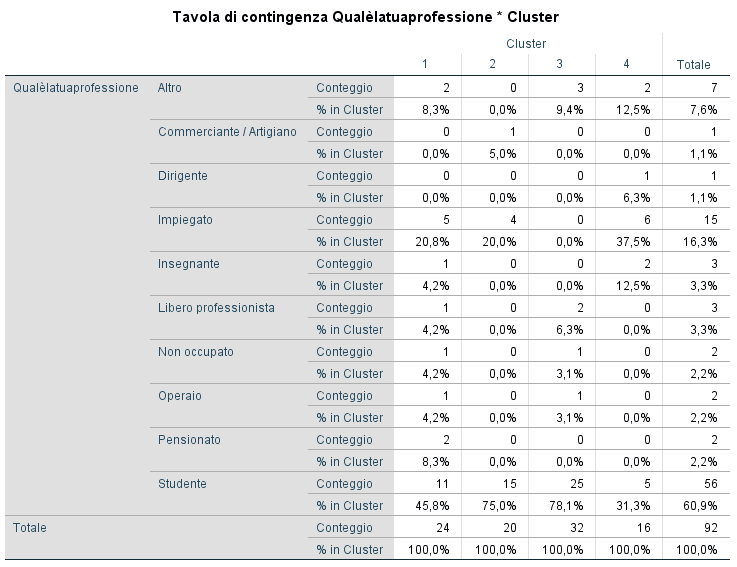


In questo caso siamo stati propensi a dividere il dendrogramma in 4 cluster e quindi dividere le osservazioni in 4 gruppi, che risultavano più o meno omogenei rispetto alla varianza e rispetto alla numerosità all’interno di ogni gruppo. Il taglio del dendrogramma non avviene dove il “salto” tra l’unione del gruppo seguente è maggiore perché ci avrebbe portato a dividere le osservazioni in due gruppi, di cui uno sarebbe stato troppo numeroso rispetto all’altro e quindi non abbiamo seguito questa strada perché non ci sembrava adatta per la nostra analisi.

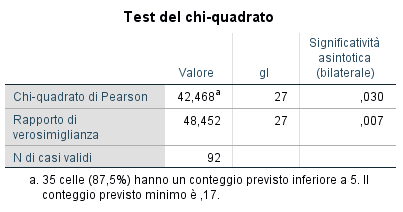
Ora andiamo ad analizzare le vere e proprie tabelle di contingenza per vedere quanto sono caratterizzati i nostri gruppi rispetto ad alcune variabili indagate all’interno del nostro questionario.



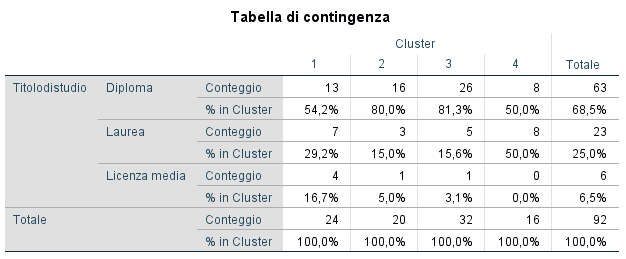
La prima tabella di contingenza che siamo andati a prendere in considerazione è quella che vede la distinzione di Sesso all’interno del campione. Dopo il commento iniziale già riportato nel capitolo 1.2 che vedeva in nostro campione squilibrato dalla parte delle femmine rispetto ai maschi, in questa tabella di contingenza non possiamo dire che i cluster siano caratterizzati dal sesso visto che le distribuzioni del genere sono omogenee e nessuna prevale sull’altra all’interno del cluster. Infatti si nota che all’interno dei cluster le percentuali di maschi e femmine rimangono fedeli alla popolazione del nostro campione. Il test del chi quadro risulta superiore alle usuali soglie e quindi le caratterizzazioni del sesso non sono significative in ogni caso.



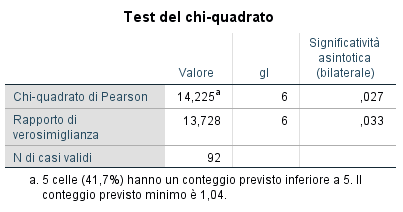
Un’altra variabile che andiamo ad analizzare per vedere se ci sono eventuali caratterizzazioni all’interno della nostra analisi cluster è la professione degli intervistati. Analizzando questa tabella risultano invece alcune caratterizzazioni significative ai fini della nostra indagine. Nel primo cluster risulta un grande numero di pensionati, molto superiore alla media del campione. Quindi i pensionati vengono spiegati in maniera soddisfacente dal nostro primo cluster. Nel secondo e terzo gruppo la caratterizzazione che rimane evidente è quella degli studenti, che nel nostro campione è molto alta infatti arriva a un 60% e sfiora quasi l’80% nel terzo cluster. Quindi in questi due gruppi la prevalenza di studenti è molto superiore alla media del nostro campione. Nel 4 cluster invece sembra che ci sia una palese caratterizzazione di tutte quelli che si possono considerare i lavori standard, come insegnante, dirigente, altro e in particolare impiegato che risulta molto più alto come livello rispetto a quello del campione. Il nostro commento in questo caso è quello che c’è una caratterizzazione in base all’età implicita dove nel primo cluster ci sono le persone più adulte, nel secondo e terzo cluster ci sono tutti i giovani e quelli che stanno ancora studiando, mentre nel quarto cluster risultano esserci tutte le persone che stanno lavorando.

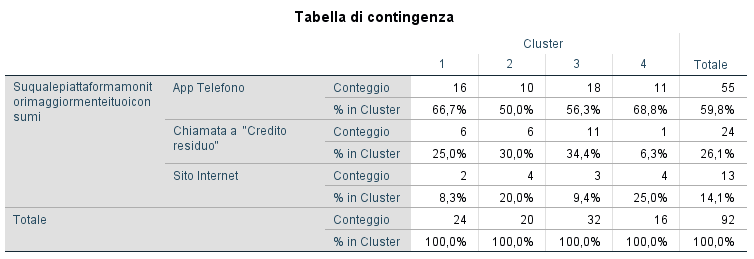


Il test conferma che la professione caratterizza veramente la partizione.



Un’altra tabella di contingenza che andiamo ad analizzare è quella relativa alla variabile Titolo di Studio, che nel nostro caso non prende in esame tutti i livelli che erano stati chiesti ma solamente quelle 3 a cui almeno una persona ha dichiarato di appartenere. In questo caso andando ad analizzare la tabella vediamo che nel primo cluster non c’è una grossa caratterizzazione, cosa che ci aspettavamo dato che il primo cluster vedeva una prevalenza di pensionati che non è detto che abbiano un determinato titolo di studio. Nel secondo e nel terzo cluster si conferma l’ipotesi fatta in precedenza, quindi quella che sono formati principalmente da studenti infatti risultano caratterizzati in particolare da diplomi, dato che l’anno che ha rilevato il maggior numero di intervistati nel nostro questionario è il 1995, che quindi contiene ragazzi che hanno conseguito il diploma ma non la laurea. Nel 4 cluster c’è una sovrabbondanza di laureati, quindi è il cluster più graduato e questo supporta il fatto che ci siano in questo cluster principalmente impiegati e insegnanti che probabilmente hanno conseguito la laurea prima di fare i suddetti mestieri. Anche in questo caso i cluster sono ben delineati e appaiono sempre più chiare e limpide le caratterizzazioni di ogni gruppo. Anche in questo caso il test di significatività risulta positivo e quindi la caratterizzazione del titolo di studio risulta sensata.





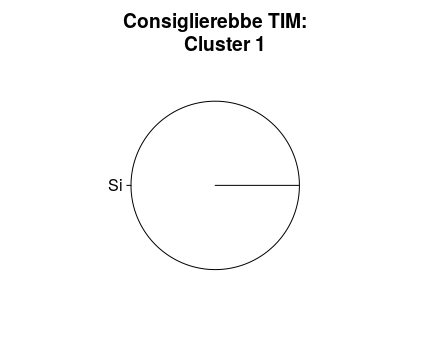
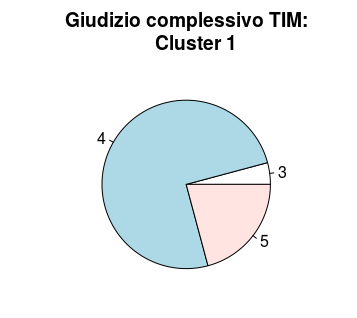
In questa caratterizzazione dei cluster dove viene presa in considerazione la variabile chiesta nel sondaggio riguardo a dove si monitorano i propri consumi del telefono le caratterizzazioni sembrano poche, e non sono neanche significative dato che il valore del test del chi-quadrato risulta pari a 0.313 e quindi superiore alle usuali soglie. L’unica cosa che ci teniamo a sottolineare è riguardo al quarto cluster dove c’è un buon numero di persone che monitorano i proprio consumi sul sito internet, quindi risulta una caratterizzazione sicuramente non di persone giovani che conferma il dato dei punti precedenti.

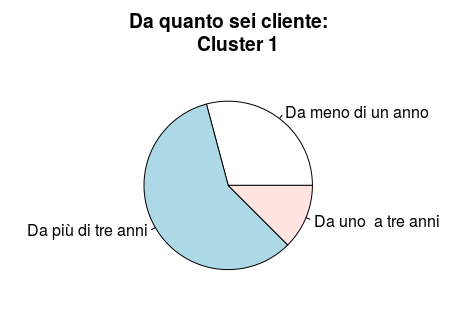
Quindi risulta chiaro da queste tabelle come ci sia un' evidenza, nonostante l’impossibilità di fare delle fasce d’età, di una caratterizzazione dei cluster piuttosto netta discussa già in questo capitolo. Sulla base di quanto studiato fino ad ora, analizzeremo il giudizio complessivo del nostro questionario per vedere quale cluster ha una visione migliore dell’azienda TIM.

**4.2 Definizione dei Cluster**

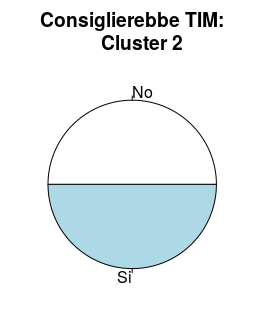
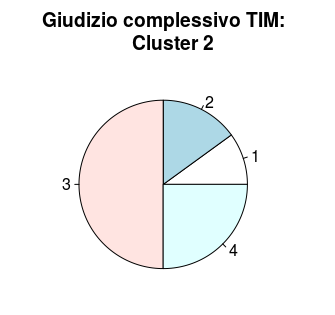
Riportiamo quindi qui sotto alcune analisi che abbiamo effettuato per cercare di caratterizzare al meglio ogni cluster e cercare di individuare quale cluster risulta maggiormente fedele all’azienda. Abbiamo indagato tra tutte le variabili possibili il giudizio finale dato all’azienda per vedere in quale cluster era più alto, se consigliavano ad amici e parenti di entrare in TIM e da quanti anni sono clienti TIM per vedere una sorta di analisi rispetto ad ogni cluster.

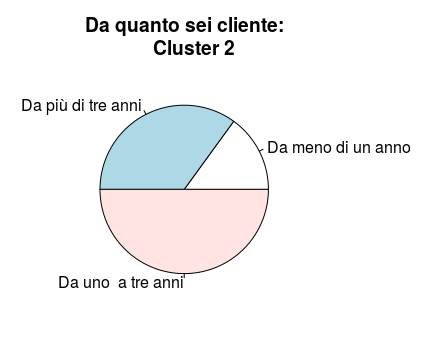
Nel primo cluster i risultati sono i seguenti: dalle nostre analisi sulle tabelle di contingenza abbiamo sottolineato che questo gruppo è abbastanza variegato, infatti la media dell’età in questo gruppo risulta 34 anni, e per quanto riguarda la soddisfazione dell’azienda è un gruppo che è molto soddisfatto di TIM.

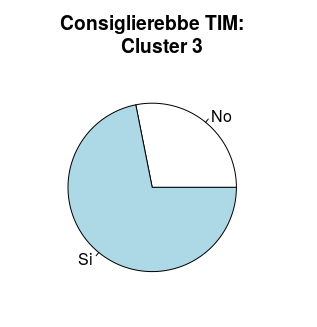
 Cluster 1: **Gli entusiasti**

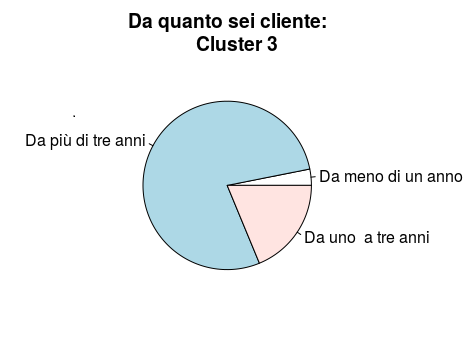
Per quanto riguarda il secondo cluster i risultati delle analisi sono significative: è un cluster formato da persone giovani che sono abbastanza scontente del servizio TIM ma sono anche le persone che da meno tempo usufruiscono di questo servizio, la media dell’età in questo gruppo è pari a 26 anni, quindi sono giovani e per un eventuale campagna di marketing sono i clienti da cui TIM dovrebbe cercare di ottenere una maggiore fidelizzazione.

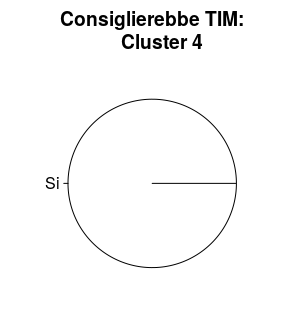
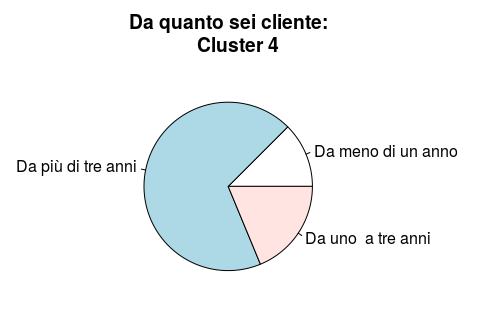
Cluster 2: **Gli Scontenti**

Nel terzo cluster ci sono sempre persone giovani, la media delle età nel cluster è di 25 anni e sono però un gruppo più soddisfatto di quello precedente perché da un giudizio complessivo maggiore all’azienda e sono, nonostante siano clienti giovani sono da tanto tempo clienti TIM e questo porta a rendere il loro giudizio più importante e significativo del precedente. Importante è il fatto che nonostante siano da tanto clienti siano comunque abbastanza soddisfatti del servizio.

Cluster 3: **Gli Storici**



Nel quarto cluster, invece, viene confermata la nostra ipotesi di persone adulte e lavoratrici, dato che la media nel cluster risulta la più alta, pari a 35 anni. Questo gruppo dà un giudizio mediamente positivo, più alto sia del terzo che del secondo cluster, ma sono persone che per lo più sono clienti TIM da tanto e accettano le condizioni di TIM, sono contenti del servizio ma non in maniera assoluta come le persone del primo cluster. Sono le persone più sicure del servizio TIM e che non vogliono cambiare, non hanno una possibile espansione in quanto a proposte di marketing.



Cluster 4: **Gli Indifferenti**

**Chap 5: I NON CLIENTI**

**5.1 Identificazione dei soggetti**

**Descrizione del campione**

Come anticipato nel capitolo 1.2 il 66.4% del campione analizzato non è cliente TIM, quindi 239 dei 360 intervistati non posseggono una scheda telefonica TIM.

*Non ho mai trovato una promozione adatta a me=**3.029*

*E' troppo costosa=**2.925*

*Non mi fido di TIM=**1.82*

*Non ha una rete effciente=**1.921*

*Mi è stata sconsigliata da amici=**1.782*

*Sono vincolato ad un altro operatore=**2.314*

*Altro=**2.113*

Queste sono le medie dei voti assegnati dai non clienti, alla domanda “Non sono cliente TIM perché..” dove 1 esprimeva massimo disaccordo e 5 massimo accordo con l’affermazione.

Come si può notare la situazione peggiore è quella delle offerte, che riporta una media di 3.029.

**Possibile promozione**

Notando che il motivo principale per cui gli intervistati non erano clienti era la mancanza di una promozione adatta, abbiamo deciso di provare a muoverci in questa direzione.

Una delle domande del nostro questionario era “Come è composta in percentuale la tua promozione “tutto incluso” ideale?”; proprio da qui cercheremo di trarre un’offerta che si adatti a quel segmento di mercato che vorrebbe entrare in TIM ma che non ha mai trovato l’offerta giusta.

Considereremo quindi solo i non clienti che hanno risposto con valore maggiore di 3 alla domanda “Non sono cliente TIM perché non ho mai trovato una promozione adatta a me” e che hanno completato la domanda rispetto la promozione in maniera corretta. Riterremo idonei quindi 78 soggetti e faremo la media delle loro “Offerte ideali” divise per tipo di traffico (dati, voce, sms).

*Come è composta in percentuale la tua promozione "tutto incluso" ideale? [Traffico Dati]*

*62.30769*

*Come è composta in percentuale la tua promozione "tutto incluso" ideale? [Chiamate]*

*25.76923*

*Come è composta in percentuale la tua promozione "tutto incluso" ideale? [SMS]*

*11.92308*

Riportiamo qui sopra i risultati dell’analisi, da cui possiamo dedurre il desiderio dei non clienti di avere una tariffa “tutto incluso” che sia composta da un’offerta strutturata come riportato sopra, che sia basata al 62.3% su traffico dati e che a chiamate e messaggi venga dato il resto dello spazio, suddiviso in 25.77% chiamate e 11.92% SMS.

**5.2 Confronto clienti-non clienti**

Dopo aver analizzato il campione dei non clienti e le varie possibilità di renderli clienti andremo a confrontare il gruppo dei non possessori di scheda TIM con quelli che invece sono in possesso di quest’ultima.

In particolare confronteremo come viene recepita la qualità della rete TIM dai fruitori di questa e da chi di questa si è fatto un’idea tramite pubblicità o amici e parenti.

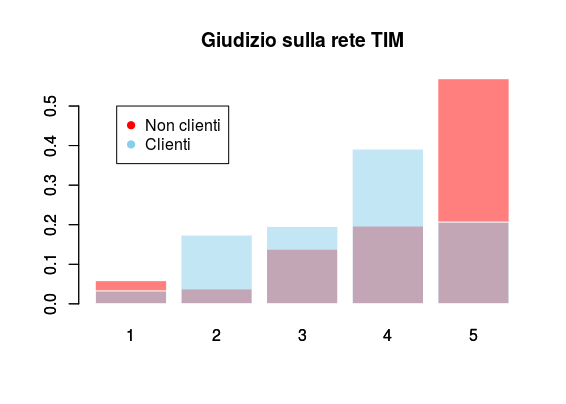
Possiamo fare questo confronto perché i clienti TIM si sono trovati a rispondere ad una domanda così formulata

“Come valuti l’uso di dispositivi con sim TIM? “ [stabilità della rete 1 2 3 4 5 ]

mentre i non clienti hanno risposto alla domanda

“Quanto sei d’accordo con le seguenti affermazioni? ‘Non sono cliente TIM perché...’” [Non ha una rete efficiente 1 2 3 4 5 ]

Produrremo un’analisi basandoci sulle risposte a queste 2 domande, in particolare useremo una trasformata della seconda domanda che riporti la scala di giudizio nello stesso verso della prima domanda.



Dal grafico qui riportato possiamo immediatamente notare quello che poi il confronto tra le medie dei 2 gruppi ci confermerà: i non clienti hanno un’opinione della rete TIM maggiore rispetto a chi l’ha effettivamente provata, le medie di discostano di 0.61 punti, che su una scala da 1 a 5 rappresentano comunque il 10% di scostamento; troviamo infatti una media di 4.18 nei non clienti, ed una media di 3.57 nei clienti.

Questo è indice di una grande campagna pubblicitaria di TIM ed è un indice di qualità percepita dai non consumatori che l’azienda potrebbe leggere come incentivante e segno di una “facile conquista di nuove fette di mercato”.

**Chap 6: CONCLUSIONI**

Cercando di trarre delle conclusioni dal lavoro svolto, diremmo che sin dal capitolo 2.2 era chiaro fosse necessario un intervanto per sopperire alle mancanze lì analizzate. In particolare per ovviare al problema maggiore, quello della capacità di risposta con gap -*1.4076,*

proporremmo di lavorare al miglioramento del servizio di call center.

Con riferimento al capitolo 3, visto il plot degli score delle CP, definendo come popolazione target tutti i soggetti nel II quadrante, proporremmo una campagna marketing per cercare di migliorare le percezioni del servizio offerto da TIM. Questi infatti erano i soggetti con maggiore potenzialità, quindi con alte aspettative ma basse percezioni.

Dalla cluster analysis, capitolo 4, ricaviamo i profili dei clienti TIM. Come visto in fase di definizione dei cluster abbiamo 4 tipi di soggetti: gli entusiasti, gli scontenti, gli storici e gli indifferenti. Viste le considerazioni lì affrontate, opteremmo per improntare una campagna marketing sul profilo “dello scontento”.

Messi in luce tutti i limiti di questa analisi, possiamo ritenerci soddisfatti del lavoro svolto e degli obbiettivi raggiunti, coscienti del fatto che con una più ampia diffusione del questionario e con un’analisi conseguentemente più precisa, i risultati della stessa sarebbero potuti essere addirittura generalizzati alla popolazione italiana.

Lavoro svolto a fini didattici da: Lorenzo Gregori e Giovanni Misseri.