



Università
Ca'Foscari
Venezia

Corso di Laurea Magistrale

in Marketing e Comunicazione

Tesi di Laurea

Il sistema E-commerce per ipermercati e supermercati: un'analisi statistica

Relatrice / Relatore

Ch.ma. Prof.ssa Debora Slanzi

Laureando

Valerio Padoan

Matricola

863789

Anno Accademico

2021 / 2022

Ringraziamenti:

Alla mia famiglia, che mi ha sempre sostenuto in questo fantastico viaggio universitario. Il loro incoraggiamento mi ha spinto ad avere fiducia in me stesso e nelle mie capacità anche al di fuori del contesto educativo.

Ad Elena, che passo dopo passo, stiamo costruendo un futuro insieme. Grazie alla sua presenza sono riuscito a superare tutte le paure, affrontandole sempre a testa ad alta. Le devo molto, soprattutto in questo ultimo periodo universitario, dove tra mille ansie, mi è sempre stata accanto, sostenendomi con il suo affetto e amore.

Ai miei amici, oltre ad essere consiglieri fedeli, sono sempre riusciti a strapparmi un sorriso durante questo percorso tra lamentele, gioie e soddisfazioni.

Ai miei colleghi, che hanno reso il mio percorso da studente lavoratore più semplice, facendomi divertire ma soprattutto crescere come persona.

Vorrei infine, esprimere gratitudine alla Professoressa Debora Slanzi. La sua generosa disponibilità ed i suoi consigli preziosi sono stati costruttivi durante lo sviluppo di questa tesi.

INTRODUZIONE	3
---------------------------	---

CAPITOLO I: Evoluzione e prospettive E-commerce

1.1 E-commerce: una panoramica generale.....	5
1.2 Tipologie di e-commerce, differenza con e-business.....	8
1.3 Dati sull'utilizzo E-commerce in Italia	12
1.4 Dati sull'utilizzo E-commerce nel mondo.....	16
1.5 Fattori critici e di successo dell'E-commerce.....	18

CAPITOLO II: La grande distribuzione organizzata

2.1 Storia del commercio in Italia e nel mondo anni 60' ad oggi.....	23
2.2 Le varie forme distributive nel mondo.....	27
2.3 Le nuove tendenze.....	32
2.4 I principali supermercati ed ipermercati che utilizzano l'E-commerce....	37
2.5 Il concetto di Omnicanalità.....	39

CAPITOLO III: Metodologia per la ricerca sociale

3.1 Le ricerche di mercato.....	46
3.2 Tecniche di indagine per la rilevazione dei dati.....	48
3.3 Le rilevazioni campionarie.....	59
3.4 Le scale di misura.....	67
3.5 Introduzione ai metodi statistici.....	75
- 3.5.1 Cluster Analysis.....	76
- 3.5.2 Regressione logistica.....	86

CAPITOLO IV: Il sistema E-commerce per supermercati ed ipermercati: presentazione caso studio

4.1 Il questionario.....	91
4.2 Caratteristiche sociodemografiche.....	97
4.3 Caratteristiche dei consumatori che utilizzano l'E-commerce.....	102

4.4	Caratteristiche dei consumatori che non utilizzano l'E-commerce.....	112
4.5	Applicazione dei metodi statistici ai dati.....	115
-	4.5.1 Cluster analysis sui soggetti che acquistano online.....	115
-	4.5.2 Analisi regressione logistica applicata ai non acquirenti.....	124
4.6	Alcune considerazioni finali.....	127
CONCLUSIONI.....		129
APPENDICE.....		133
ELENCO FIGURE E TABELLE		143
BIBLIOGRAFIA		149
SITOGRAFIA.....		150

INTRODUZIONE

La diffusione dell'E-commerce e delle nuove tecnologie hanno modificato le abitudini ed il comportamento delle persone nell'effettuare i propri acquisti personali.

Il miglioramento è avvenuto anche per le imprese, che grazie all'E-commerce possono usufruire di un nuovo canale per vendere i propri prodotti alle persone, informarle sulle offerte relative alle nuove referenze e servizi ma soprattutto possono raccogliere informazioni importanti sui comportamenti e abitudini per indirizzare poi la propria offerta di prodotto in modo specifico per ogni consumatore.

Ai giorni d'oggi è possibile dunque parlare di modernizzazione 4.0, dove è principe il concetto di omnicanalità. Le aziende non gestiscono più solo diversi canali di vendita ma diversi profili di consumatori. In questo modo si sta raggiungendo il punto dove acquistare online ed acquistare dal negozio per alcuni clienti è la medesima cosa.

È importante sottolineare che il servizio E-commerce ha visto un incremento del suo utilizzo nel 2020, a causa della pandemia Covid-19. Date le limitazioni di spostamento dovute dalle restrizioni politiche e sanitarie, il servizio E-commerce in alcuni casi è diventato essenziale per soddisfare le esigenze e le richieste delle persone. Durante la pandemia molte imprese hanno dunque cominciato ad investire denaro sul servizio online effettuando ad esempio costi di logistica o costi di sviluppo del sito internet.

L'evoluzione principale però è avvenuta da parte dei supermercati ed ipermercati, i quali hanno cercato di aiutare le persone che non erano in grado di spostarsi dalla propria abitazione a causa delle limitazioni, sviluppando un servizio E-commerce dove i consumatori potevano effettuare la propria spesa da casa. Le imprese si stanno sviluppando cercando metodi innovativi per consegnare i propri prodotti in grado di soddisfare la richiesta da parte dei clienti, ad esempio l'utilizzo di Locker per la consegna della merce o il ritiro della spesa già preparata e consegnata al punto vendita più vicino a casa. Le grandi imprese inoltre cercano di innovarsi e sfruttare nuovi servizi per ampliare la propria offerta di vendita introducendo la ristorazione come metodo alternativo per vendere i propri prodotti.

Il presente elaborato vuole contribuire all'analisi cercando di estrarre informazioni riguardo alle abitudini dei consumatori nei confronti dell'E-commerce, evidenziando eventuali fattori che possono incidere sulla decisione di affidarsi o meno al servizio.

Nel primo capitolo viene fornita una panoramica generale sull'utilizzo dell'E-commerce, introducendo la storia e come si è evoluto negli anni. Vengono indicati i dati relativi al

servizio online oggi sia in Italia che nel mondo ed i fattori che lo hanno portato al successo in questi ultimi anni.

Nel secondo capitolo viene delineata la storia e lo sviluppo delle varie forme distributive nel mondo, evidenziando inoltre la differenza tra le varie strutture organizzative come ad esempio tra supermercato, ipermercato e discount. Vengono trattati temi riguardanti le nuove tendenze come il caso Coop, Amazon, Carefur e Alibaba ed infine viene illustrato in dettaglio il tema della omnicanalità e come si lega al concetto di E-commerce e alle forme distributive.

Nel capitolo terzo vengono illustrati i vari modi per raccogliere le informazioni ed effettuare una ricerca di mercato. Vengono dunque introdotte le tecniche di indagine per rilevare dati come il questionario. Inoltre, sono affrontati temi su come i dati raccolti possono essere analizzati per ricavare informazioni importanti su un fenomeno oggetto di studio. Verranno presentati dunque le principali scale di misura ed i metodi statistici per analizzare i dati, approfondendone vantaggi e svantaggi.

Infine, nel quarto capitolo verrà introdotto il questionario utilizzato per la raccolta di dati, presentando le domande effettuate al campione. Quest'ultimo verrà diviso principalmente in due gruppi che rappresentano i soggetti che acquistano online da supermercato o ipermercato e chi non utilizza il mercato online relativo ai vari punti vendita di genere alimentare. I dati verranno poi descritti tramite analisi univariata esplorativa, per avere un quadro generale sulle informazioni rilevate. Infine, saranno analizzati tramite modelli statistici appropriati per contribuire alla descrizione del fenomeno e all'individuazione dei fattori che più identificano i due gruppi di soggetti analizzati. Data la natura del questionario e i metodi di somministrazione dello stesso bisogna sottolineare che i risultati ottenuti sono a scopo puramente descrittivo del campione analizzato e non possono avere nessuna valenza inferenziale rispetto alla popolazione di coloro che acquistano online da supermercato o ipermercato.

CAPITOLO I

1.1 E-COMMERCE: UNA PANORAMICA GENERALE

La nascita dell'E-commerce pone le sue origini già dal 1970 con EDI (Electronic Data Interchange). Si tratta del primo servizio elettronico, con il quale avviene un interscambio tra sistemi informativi sui canali, dove vengono trasferiti documenti tra le varie compagnie per semplificare e computerizzare gli acquisti. La trasmissione dei dati non era così agevole, infatti era richiesta una linea di comunicazione tra i vari utilizzatori e, a causa delle abituali modifiche al database aziendali, questo comportava ingenti spese di denaro nella sua gestione. Paragonato ai tradizionali metodi di generazione e scambio di documenti commerciali tra aziende come lettere, fax, e-mail, l'EDI consentiva di ridurre l'intervento umano e quindi evitare numerosi errori. Inoltre, permetterà di eliminare la documentazione cartacea, abbattendone i costi, e di facilitare e velocizzare gli scambi commerciali. Dall'altro lato però la trasmissione dei dati non era così agevole; infatti, era richiesta una linea di comunicazione tra i vari utilizzatori e, a causa delle abituali modifiche al database aziendali, questo comportava ingenti spese di denaro nella sua gestione. Ne derivava quindi la necessità di revisionare i processi e i flussi di business frequentemente e il bisogno di impiegare personale specializzato per le operazioni o l'esternalizzazione delle stesse.

Nel 1982, in Francia, viene immesso sul mercato da France Telecom, il Minitel. L'hardware della compagnia francese era formato da un sistema di terminali video collegati alla rete telefonica, con la quale si poteva scambiare messaggi, fare transazioni e raccogliere informazioni¹. Il valore aggiunto di questo servizio era dato dalla possibilità, estremamente innovativa per l'epoca, di fare prenotazioni sulla rete. Seppur il Minitel in Francia poteva vantare milioni di utilizzatori, con l'arrivo degli anni Novanta, il servizio offerto dalla France Telecom viene presto sovrastato da Internet, ma rimarrà comunque la prima vera rete di comunicazione civile prima dello sviluppo del Web².

Grazie all'introduzione del Web nel mondo ci si avvicina sempre di più al mondo moderno e nel 1992 si inizia a parlare di E-commerce con la libreria online Book Stacks

¹ TIZIANO VESCOVI, *Il marketing e la rete: la gestione integrata del Web nel business: comunicazione, e-commerce, sales management, business to business*, Milano: Il sole 24 ore, 2007, pag.29

² TIZIANO VESCOVI, *Il marketing e la rete: la gestione integrata del Web nel business: comunicazione, e-commerce, sales management, business to business*, Milano: Il sole 24 ore, 2007, pag.29

Unlimited, che con il passare degli anni si chiamerà Books.com. Si tratta appunto di una delle prime librerie online, ideata in America da Charles M. Stack. Per attirare i clienti, oltre alla possibilità di consultare i vari libri, il sito includeva anche un diario letterario quotidiano, riepiloghi, interviste RealAudio con autori e forum in cui i clienti potevano porre domande e discutere di libri.

Secondo le fonti riportate su Internet³, il primo vero e proprio acquisto online però è avvenuto nel 1994 con la vendita di un CD contenente l'album del cantautore Sting tramite un mercato online chiamato NetMarket fondato da Dan Khon e Roger Lee. Il New York Times riconosce alla società la prima transazione al dettaglio sicura in Internet.

Dal 1995 cominciano a svilupparsi i primi siti commerciali avviati dalle imprese. La vera svolta avviene però grazie al fondatore di Amazon Jeff Bezos che spedisce a Seattle il primo libro acquistato sul suo sito online. L'idea principe dell'azienda era quella di vendere libri online in tutto il mondo, prevedendo per i primi 5 anni di attività una perdita di profitto. Con il passare degli anni, sul sito Amazon si iniziò a vendere non solo libri, ma anche prodotti di genere diverso come CD, film, software, dispositivi elettronici e videogames⁴, consacrando nel 1998 Amazon come l'esempio più concreto di E-commerce al mondo. Il suo successo è stato e continua ad essere dilagante come dimostra il fatto che oggi il valore del brand si aggira sui 684 miliardi di dollari⁵.

A metà degli anni Novanta Pierre Omidyar, iraniano trasferitosi a Parigi apre AuctionWeb. L'idea di base era molto semplice: permettere a chiunque volesse disfarsi di oggetti vecchi di caricare gratuitamente un proprio annuncio con annesse foto e prezzo di partenza che sarebbe poi eventualmente cresciuto in base alle offerte dei possibili acquirenti. Si tratta di uno dei primi esempi di siti di aste online. Il 4 Settembre 1995 avviene la prima transizione con la vendita di un puntatore laser rotto per il modico prezzo di 15 dollari. Il marketplace cambia poi nome in Ebay e migliora sempre di più le sue funzionalità, aggiungendo la possibilità da parte degli utenti di dare un giudizio sui

³ Corriere della sera: Il primo acquisto online: https://www.corriere.it/tecnologia/cards/ecommerce-prime-informazioni-scambiate-digitalmente-nascita-colossi-commercio-elettronico/primo-acquisto-digitale_mobile.shtml

⁴ Il sole 24 ore: I 25 anni di Amazon: dal garage di Bezos al negozio globale di ogni cosa: https://www.ilsole24ore.com/art/dal-garage-bezos-mille-miliardi-storia-amazon-che-ora-vuole-superare-apple-AE4ZzfjF?refresh_ce=1

⁵ Il Corriere: Come è nato l'E-commerce, la storia i primi passi di Amazon, Ebay e Alibaba: <https://www.corriere.it/tecnologia/cards/ecommerce-prime-informazioni-scambiate-digitalmente-nascita-colossi-commercio-elettronico/nascita-ebay.shtml>

prodotti acquistati. Questo innovativo servizio di rating e la geniale mossa di vendere anche biglietti di treni ed aerei fanno esplodere il sito che ha poi continuato a crescere negli anni. Ancora oggi è infatti uno dei siti più utilizzati al mondo e vanta, nel 2020, utili pari a 485 milioni di dollari.

Primi acquisti in Italia

In Italia il primo acquisto online viene effettuato il 3 Giugno 1998, tramite il sito internet IBS.it. Il pagamento online è avvenuto dopo 35 minuti dall'apertura del portale e si trattava di un libro intitolato "La Concessione del Telefono" scritto da Andrea Camilleri e acquistato dal Canada.

Tra il 1997 e il 1998 inizia ad esserci una forte competizione tra web browser come Netscape e internet Explorer per il dominio del mercato. Netscape Navigator nasce come primo browser commerciale di successo grazie ad alcune importanti modifiche apportate anche al linguaggio HTML utilizzato. Rimane l'unico browser in circolazione per poco tempo perché nel 1995 entra in competizione con Internet Explorer della Microsoft. Alla fine, prevale quest'ultimo perché la Microsoft decide di distribuire il proprio browser, per la precisione Internet Explorer 3, già all'interno del sistema operativo Windows 95. A quel punto gli utenti avrebbero dovuto cambiare autonomamente il browser, ma nella maggior parte dei casi ciò non è avvenuto.

A partire dagli anni 2000 il fenomeno dell'e-commerce inizia a crescere grazie al miglioramento della velocità di connessione ad internet e lo sviluppo di nuovi sistemi tecnologici. Le principali compagnie come Visa, Mastercard, Discover e American Express formarono il PCI (Payment Card Industry Security Standards Council) atto a regolarizzare il mercato online con standard di sicurezza predefiniti. Il consorzio aveva l'obiettivo di impedire furti relativi a dati personali durante l'utilizzo di carte di pagamento da parte dei consumatori per le transazioni effettuate online. Viene dunque rilasciato uno standard chiamato PCI DSS (Payment Card Industry Data Security Standard), utilizzato dalle strutture che utilizzano metodi di pagamento con carte.

Pochi anni più tardi grazie all'introduzione degli smartphone, l'e-commerce comincia ad espandersi rapidamente e sempre più clienti dei negozi online cominciano ad acquistare i prodotti tramite telefono. Si conferma così il livello di sicurezza nei pagamenti effettuati online.

I dati di oggi rilevano che più di cinque miliardi di persone sono connesse ad Internet. La crescita dell'e-commerce con il passare degli anni è sempre più esponenziale, si rileva

che nei paesi dell'Asia-Pacifico, il commercio online tocca i 1892 miliardi di dollari. In Europa invece all'incirca il 69% della popolazione ha effettuato almeno un acquisto⁶.

1.2 TIPOLOGIE E-COMMERCE E DIFFERENZE E-BUSINESS

Il termine e-commerce è spesso confuso con e-business. Con l'espressione e-business si intende tutto ciò che è necessario al processo di gestione di un business online; non ci si limita soltanto all'acquisto e alla vendita dei beni come nel caso dell'e-commerce, ma si forniscono anche servizi come, ad esempio, fornire assistenza ai clienti, comunicazione con i dipendenti, clienti o soci commerciali.

Alcuni prototipi di e-business utilizzano Internet nelle proprie attività di mercato. In base al suo impiego si possono distinguere due modelli principali.

- Pure Play o Dot-com: che si riferiscono ad imprese che operano unicamente in Internet.
- Brick and Click o Clicks and mortar: le aziende integrano le attività online con quelle tradizionali quindi con uffici, stabilimenti e punti vendita. Nel primo caso le attività si svolgono puramente online⁷, mentre nel secondo caso sono imprese Hybrid dal momento che uniscono il digitale con il fisico.

I vari modelli di e-business che emergono dalle imprese che integrano le proprie funzioni con le attività di Rete sono:

- Modello mass media: sono imprese che vengono istituite per il commercio di spazi pubblicitari online come accade nei modelli di business dei media TV, radio e stampa. Attraverso una rete di vendita, i rivenditori di spazi web contattano potenziali aziende, agenzie di pubblicità e centri media. Si utilizza quindi un sito internet come strumento per raggiungere i potenziali clienti e audience dagli inserzionisti. Gli attori principali in questo modello sono essenzialmente tre: editore del sito, inserzionista e cliente finale. L'editore ha la funzione di sviluppare due offerte per rendere il suo business un successo. La prima è rivolta al visitatore del sito ossia il cliente finale, al quale devono essere

⁶ Ansa: La crescita dell'e-commerce traina ogni mercato:

https://www.ansa.it/pressrelease/economia/2020/11/20/la-crescita-virtuosa-delle-commerce-traina-ogni-mercato_e41da430-caa2-4811-a77d-d4e932de14ca.html

⁷ TIZIANO VESCOVI: *Il marketing e la rete: la gestione integrata del Web nel business: comunicazione, e-commerce, sales management, business to business*, Milano: Il sole 24 ore, 2007 pag.56

proposti contenuti allettanti che lo spingano alla permanenza e alla frequentazione del sito; la seconda è dedicata all'inserzionista che ha la possibilità di accedere a spazi pubblicitari, link e connessioni che generano maggiore "visibilità". Il compenso dell'editore si misura in base al costo delle impression ovvero al costo che viene sostenuto per aumentare la popolarità del sito oppure tramite le caratteristiche interattive di Internet. Tramite il tasso click through si può infatti misurare ad esempio il comportamento del visitatore quindi quanto rimane all'interno della pagina, dove clicca e che interessi ha all'interno dell'inserzione. Uno dei punti deboli di questo modello è l'investimento richiesto per gestire e mantenere i contenuti di qualità e di interesse adeguati al target di riferimento, tutto questo confrontato con i ricavi dovuti dalla vendita degli spazi online. I primi citati sono relativamente elevati, invece i secondi si riducono a causa dell'elevata offerta di spazi presente online⁸.

- Modello e-commerce: si basa su una vendita online dove il negoziante vende i prodotti ai clienti, pagando il prezzo stabilito. Il modello è utilizzato dalle imprese sia dot-com e clicks and mortar, le quali completano la vendita sia con un canale offline sia con un canale online. Il cliente finale che acquista tramite un sito e-commerce un prodotto digitale è agevolato perché il passaggio avviene tutto attraverso la Rete. Se invece viene commercializzato un bene fisico allora è essenziale la presenza di un addetto logistico nello smistamento dei prodotti venduti. Uno dei punti deboli del modello, infatti, può essere quello del mancato o inadeguato servizio al cliente a causa del sistema logistico inadeguato. Questo può causare scarsa fiducia nei confronti dei sistemi di pagamento elettronici.
- Modello abbonamento e modello a consumo: si tratta di due modelli che discendono da quello e-commerce. Nel modello abbonamento, si parla della fruizione di un servizio come i giornali online ("Gazzetta dello sport", "Il Sole 24 Ore", "Wall Street Journal"), o database di raccolta di informazione. Questi contenuti vengono usufruiti dal cliente tramite un abbonamento a pagamento quindi mediante un pagamento mensile fisso o annuale, indipendentemente dal consumo effettivo. Dunque, il fornitore del servizio in abbonamento avrà il compito di definire il prezzo dell'abbonamento che dovrà essere

⁸ TIZIANO VESCOVI, *Il marketing e la rete: la gestione integrata del Web nel business: comunicazione, e-commerce, sales management, business to business*, Milano: Il sole 24 ore, 2007

remunerativo rispetto all'investimento dei contenuti generati. Per invogliare i clienti ad abbonarsi al servizio è infatti fondamentale che i costi non siano troppo elevati altrimenti si rischia di ottenere l'effetto inverso. Il modello a consumo strutturalmente è simile a quello precedente ad eccezione del fatto che il cliente paga in base ai servizi che intende consumare. Per esempio, l'acquisto di un servizio di intrattenimento come pay-per-view, un aggiornamento di un sistema, informazioni o notizie di interesse da parte del cliente. Un problema che emerge nell'utilizzo di questi due modelli di e-business è la gestione e la definizione dei prezzi che possono essere suddivisi e differenziati in base al servizio. Le soluzioni alla problematica possono essere sostanzialmente due: si applica un prezzo standard ad una singola unità di servizio (ad esempio acquisto di una partita in streaming senza abbonarsi), oppure si utilizza una formula a sistema di pagamento a punteggio dove il cliente acquista un insieme di servizi tramite pagamento in un'unica soluzione e li utilizza a consumo.

Oltre ai modelli di e-business descritti precedentemente, le aziende possono adottare composizioni miste di modelli in base alla interpretazione del mercato. Si parla in questo caso di metamediari, ovvero ad una divisione di intermediari che inglobano il commercio con la tecnologia. Gli intermediari costruiscono un mercato virtuale dove si incontrano acquirenti e venditori che agevolano l'incontro da domanda e offerta. In breve, possiamo dire che i modelli di e-business si stanno orientando verso una virtualizzazione dei mercati, ampliando così l'offerta ai clienti attraverso anche l'integrazione di servizi collegati ai prodotti di base forniti tradizionalmente che incoraggiano forme miste di business. Le imprese che svolgono attività tramite Internet utilizzano alleanze e outsourcing per ampliare la propria rete di vendita. Emergono inoltre modelli di business customer driven, dove è presente una forte personalizzazione dell'offerta ossia il cliente acquista attraverso la fruizione di servizi online self-service. L'acquirente diventa fonte di vantaggio competitivo passando da un marketing collettivo ad un marketing one to one.

Ricapitolando, e-commerce ed e-business sono due termini con significato diverso. Comparando le due espressioni si rende noto che l'e-commerce è lo scambio di merci tramite internet ed è un sottoinsieme dell'e-business. Al suo interno vengono svolte transazioni di tipo monetario dove il venditore può interagire con il cliente senza interfacciarsi direttamente face to face. Con il termine e-business invece si esprime tutto ciò che riguarda la gestione degli affari online: è un superset che comprende anche l'e-commerce.

Entrando più nello specifico, bisogna delineare quali sono i tipi di e-commerce principali in base ai soggetti che offrono il servizio ed ai destinatari:

- B2B (business to business): si riferisce alla commercializzazione tra aziende al fine di scambiare in tempo reale informazioni, prodotti e servizi, con transazioni effettuate elettronicamente. Un esempio pratico di B2B riguarda la negoziazione tra un grossista ed un rivenditore. Secondo i dati di Netcom (consorzio italiano e punto di riferimento in materia di e-commerce e retail digitale nazionale) il mercato B2B in Italia ha avuto un valore di 406 miliardi di euro nel 2020. Il 55% delle imprese italiane B2B, attive nelle vendite online con un proprio sito e-Commerce o tramite i marketplace B2B e con fatturato pari o superiore a 2 milioni di euro, hanno registrato una crescita nell'ultimo anno del +6% delle imprese che usufruiscono dei canali digitali di vendita rispetto al +2,5% registrato nei precedenti anni⁹.
- B2C (business to consumer): la contrattazione in questo caso avviene tra azienda e cliente finale quindi l'impresa vende direttamente le informazioni, i prodotti ed i servizi al cliente. Il processo di acquisto tra l'impresa ed il cliente tende ad essere più breve e meno complesso rispetto alla contrattazione che avviene nel B2B in quanto in quest'ultimo si devono realizzare diversi incontri e negoziazioni prima di arrivare ad un vero e proprio accordo. Un'azienda, dunque, può vendere i propri prodotti sia ad altre aziende sia direttamente al cliente finale. Un esempio concreto è quello di Samsung o Apple: le aziende non solo vendono i propri prodotti come B2C quindi ai clienti tramite i negozi o sito internet, ma possono distribuire i propri prodotti nei negozi di elettronica (quindi ad altre imprese) e a loro volta rivenderanno il prodotto ai propri clienti. Anche Amazon è una piattaforma che vende secondo un modello di business B2C. Nasce come una libreria online, ma nel tempo si è sviluppata enormemente diventando uno degli e-commerce di maggior successo a livello globale, che vende non più solo libri, ma milioni di generi diversi di prodotti.

La differenza sostanziale tra B2C e B2B sta nel tipo di comunicazione utilizzata che dipende dal target di riferimento a cui si rivolge. Nel primo caso la comunicazione è diretta ed emotiva, cercando di richiamare l'attenzione del cliente verso l'esperienza di consumo. Nel secondo caso invece si utilizza un approccio più informale perché non c'è

⁹ Engage: Cresce l'adozione delle-commerce B2B tra le aziende italiane. I dati di Netcomm: <https://www.engage.it/dati-e-ricerche/cresce-l-adozione-dell-ecommerce-b2b-tra-le-aziende-italiane-i-dati-di-netcomm.aspx>

la necessità di coinvolgere il pubblico, dal momento che la vendita avviene direttamente ad altre aziende già interessate al prodotto.

- C2C (consumer to consumer): con questa tipologia di commercio elettronico, lo scambio di prodotti viene effettuato tra consumatori. Questo business è stato incrementato ed è diventato popolare negli ultimi anni grazie anche allo sviluppo di piattaforme che hanno lo scopo di mettere in relazione le persone con l'obiettivo di acquistare, scambiare o vendere beni e servizi. Esempi di piattaforme di questo tipo sono ad esempio Ebay, sito internet che offre la possibilità di vendere o acquistare prodotti da altri consumatori, ma anche di effettuare delle aste online; Letgo e Offerup, due piattaforme che consentono agli utenti situati geograficamente vicini di mettersi in contatto per scambiare, vendere, acquistare prodotti, incentivando l'incontro *face to face*.
- C2B (consumer to business): in questo modello di business, a differenza dei precedenti descritti, il consumatore stabilisce le condizioni della transazione: decide il prezzo che è disposto a pagare per uno specifico prodotto o servizio e, successivamente, l'impresa deciderà se accettare o meno l'offerta. Ad esempio, un cliente propone ad una compagnia aerea un prezzo per l'acquisto di biglietto aereo per un determinata tratta e la compagnia decide se accettare o rifiutare l'offerta. Questa particolare forma di transazione non è popolare come quelle descritte precedentemente, ma esistono vari esempi di e-commerce C2B. Tra tutti spicca soprattutto Priceline, compagnia fondata da Joy Walker, con lo scopo di organizzare aste al ribasso per biglietti aerei, pernottamenti in hotel e qualsiasi altro bene o servizio che perde valore se non viene consumato in un giorno specifico.

1.3: DATI SULL'UTILIZZO DELL'E-COMMERCE IN ITALIA

L'e-commerce B2C in Italia è in continuo miglioramento come mostrato in figura 1.1. Nel 2021 i dati relativi agli acquisti di prodotto hanno avuto un tasso più contenuto (+18%) rispetto l'anno precedente (+45%), toccando la soglia di 30,5 miliardi di euro. Per quanto concerne la fruizione degli acquisti dei servizi online rispetto l'anno 2020, i dati rilevano un miglioramento con una ripresa del 36% in valore monetario quindi all'incirca 8,9 milioni di euro.

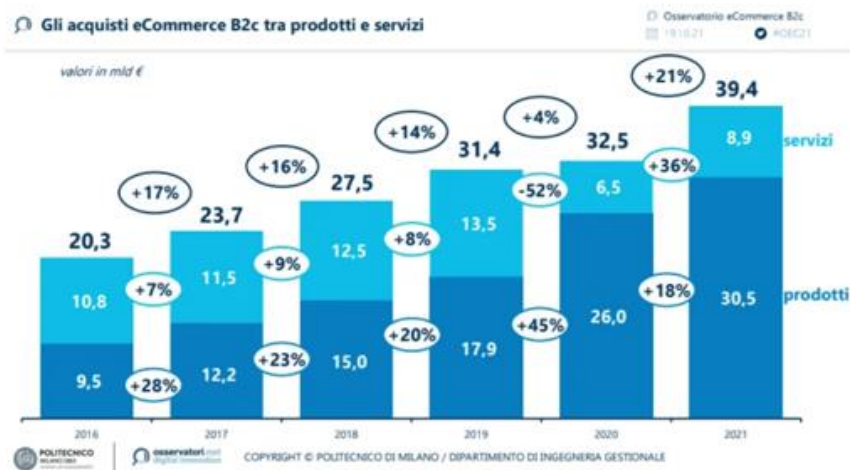


Figura 1.1 Gli acquisti eCommerce B2c tra prodotti e servizi. Fonte: Dipartimento di ingegneria, Politecnico di Milano.

Ciò che ha favorito lo sviluppo dell'e-commerce sono state le restrizioni dovute alla pandemia Covid-19, durante il periodo di lock-down, infatti, il sistema del mercato online è stato fondamentale per reperire i prodotti a molte famiglie italiane, evitando così spostamenti in zone affollate. La pandemia non ha solo cambiato la modalità di acquisto delle persone, ma ha anche modificato i sistemi di vendita di molte aziende che hanno cominciato a sviluppare i propri percorsi di acquisto integrando sia il canale offline che online. Roberto Liscia, presidente di Netcomm (Consorzio del Commercio Digitale italiano), ha sottolineato questo cambiamento. Prima della pandemia, infatti, il 70% dei rivenditori e grossisti non strutturava la propria impresa per il mercato online¹⁰, mentre nel 2020 a livello europeo il mercato dell'e-commerce ha raggiunto all'incirca 756 miliardi di euro, con una crescita pari al 10% rispetto al 2019.

La crescita dell'e-commerce è stata incoraggiata anche dall'utilizzo dei sistemi di pagamento online correlato all'adozione di strumenti tecnologici e servizi internet. In Italia si evince che più di cinquanta milioni di persone accedono a Internet con una media di sei ore al giorno, di cui almeno due di esse vengono spese nei social network¹¹. Per quanto riguarda i dati relativi all'e-commerce si è evidenziato che gli utenti nella fascia

¹⁰ Bitmat: Boom e-commerce: mercato da 39,4 miliardi di euro nel 2021: <https://www.bitmat.it/blog/internet/ecommerce/boom-e-commerce-mercato-da-394-miliardi-di-euro-nel-2021-21-rispetto-al-2020/>

¹¹ Puntoinformatico: Un italiano su due acquista online la svolta della pandemia: <https://www.punto-informatico.it/un-italiano-su-due-acquista-online-la-svolta-della-pandemia/>

d'età compresa tra i 55 e 64 anni, nel 2020, hanno acquistato online all'incirca quanto i consumatori con età tra i 25 e i 34 anni. La figura 1.2 mostra la distribuzione delle classi d'età degli utenti che hanno usufruito dei servizi e-commerce.



Figura 1.2: Distribuzione per classi d'età. Fonte: We are Social. Fonte: Digital Report Italia 2021

Nel 2020 le categorie di prodotto vendute maggiormente sul mercato elettronico sono state: cibo e personal care, con un aumento del +38,7% rispetto all'anno precedente e cibi d'asporto, mobili ed elettrodomestici, con un incremento del +27,5%. Dall'altro lato invece, a causa della situazione di emergenza sanitaria e delle restrizioni alla mobilità, il settore turistico ne ha risentito fortemente e si è infatti abbassato all'incirca del -53,9%¹².

Come viene riportato nella figura 1.3, tra l'anno 2019 e 2020 si è vissuto dei cambiamenti anche sugli acquisti online. Acquisti effettuati online per il tempo libero come hobby o sport ha sempre la quota maggiore il 48%, i settori che hanno risentito di più nel 2020 sono gli spettacoli e turismo. Crescono però settori come l'elettronica del 12% rispetto l'anno precedente, raggiungendo all'incirca il 4% sul fatturato totale; la moda invece resta stabile al 2%, ma con una crescita del fatturato del 14% rispetto il 2% del 2019¹³.

¹² Goodcom: Digital report 2021: I dati italiani: <https://www.goodcom.it/digital-report-2021-i-dati-italiani/>

¹³ Casaleggio: E-commerce in Italia 2021: <https://www.casaleggio.it/focus/rapporto-e-commerce-in-italia-2021/>

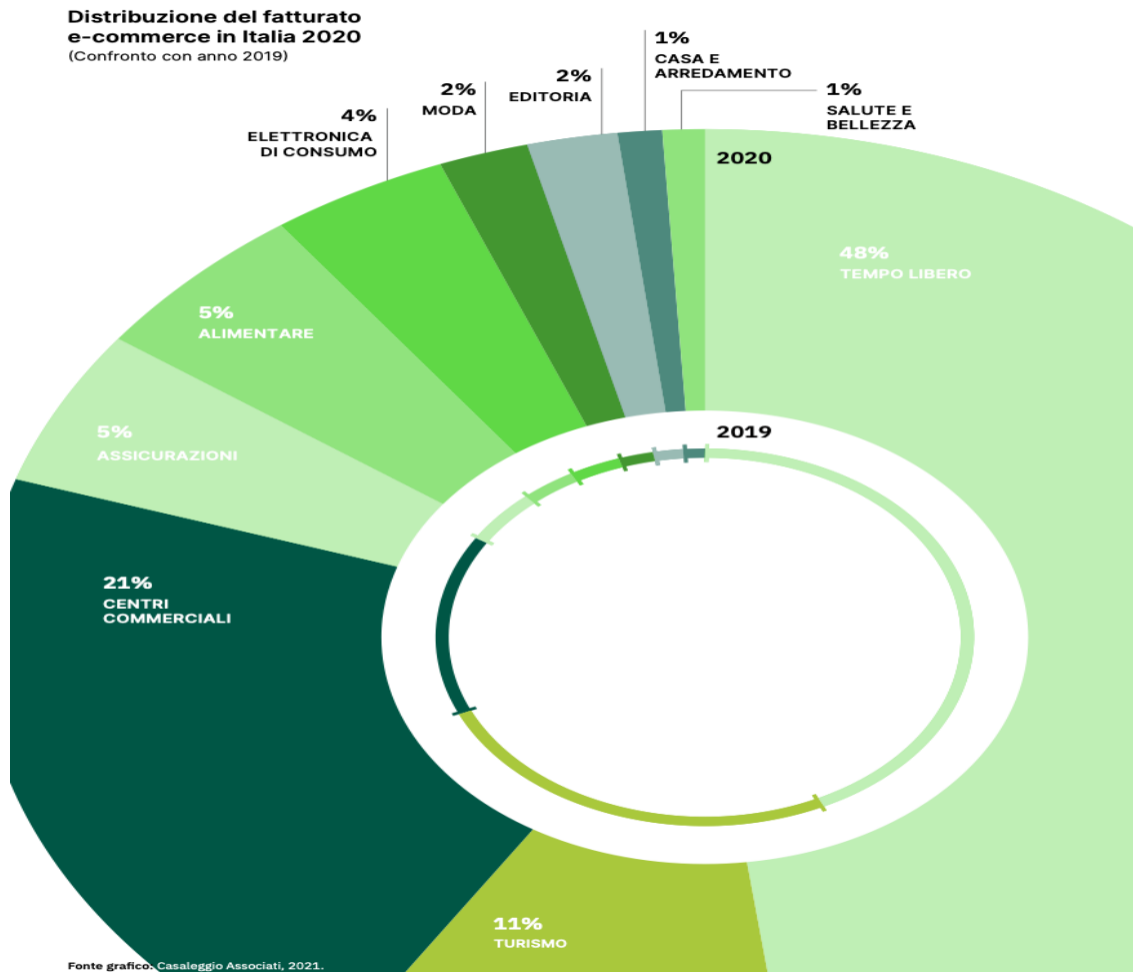


Figura 1.3: E-commerce in Italia 2021. Fonte: Casaleggio

In Italia la soluzione adottata dalle imprese B2C durante la pandemia Covid-19 si sono sostanzialmente concretizzate nell'utilizzo di un sito E-commerce o nella modifica dell'assetto da parte dei negozi di vendita. Per ovviare all'impedimento degli spostamenti della popolazione sono state adottate varie tecniche¹⁴:

- Partnership one to one: si sfruttano sistemi logistici più piccoli e personalizzabili. Un esempio è Tender, il quale gestisce la logistica dei siti e-commerce e botique delle città più rinomate in Italia come Roma, Firenze e Milano. Le consegne vengono effettuate one to one, avvalendosi quindi di ryder. Il servizio è così sia più ecologico che più rapido, con la possibilità di consegnare prodotti con un peso pari a 100 kg grazie alle cargo bike.

¹⁴ Casaleggio: E-commerce in Italia 2021: <https://www.casaleggio.it/focus/rapporto-e-commerce-in-italia-2021/>

- Incarico a piattaforme e-commerce locali: sono nate delle piattaforme online per piccole aziende che non si avvalevano ancora di un sito e-commerce per la consegna dei propri prodotti. Si tratta soprattutto di azioni locali come ad esempio Bergamo Smart Shopping, un servizio per la consegna a domicilio.
- Hub logistici in città: questa soluzione è stata adottata da molti negozi che prima della pandemia erano solamente fisici e aperti al pubblico. Sono stati quindi trasformati in centri last mile: i clienti effettuano l'ordinazione nel sito e-commerce e successivamente la merce viene preparata dai dipendenti dell'azienda in modo da poter essere consegnata già pronta all'arrivo degli acquirenti in negozio. Questo tipo di modalità di consegna dei prodotti viene utilizzata non solo per negozi di moda, ma anche per ipermercati e supermercati.

Le imprese italiane hanno saputo adattarsi ai cambiamenti moderni grazie ad ingenti investimenti che hanno portato cambiamenti strutturali importanti come: lo sviluppo di un sito e-commerce; la digitalizzazione dei pagamenti; l'amministrazione dei magazzini e della logistica ed infine lo sviluppo di nuove strategie relazionali con i clienti. Questo cambiamento ha coinvolto anche i consumatori e si tratta, molto probabilmente, di una situazione permanente che richiederà continui aggiustamenti, miglioramenti e interventi anche in futuro.

1.4: DATI SULL'UTILIZZO E-COMMERCE NEL MONDO

In riferimento all'intero pianeta, gli utenti che hanno accesso ad Internet sono 4,6 miliardi, complessivamente il 59% della popolazione mondiale. Il continente Asiatico vanta il maggior numero di utilizzatori di internet (2,4 miliardi), di cui solo la Cina ne conta la metà¹⁵.

Interessanti sono anche i dati relativi ai dispositivi tecnologici con i quali gli utenti accedono ad Internet. Il primato lo detiene il cellulare con il 55,73% come riportato nella figura 1.4 nel 2020 il mobile ha portato percentuali di fatturato altissime alle aziende, al secondo posto il computer con il 41,46% ed infine con il 2,81% il tablet. A livello globale, dunque, la percentuale di coloro che hanno cercato un prodotto o un servizio sulla rete è

¹⁵ Casaleggio: E-commerce in Italia 2021: https://www.casaleggio.it/wp-content/uploads/2020/12/CA-E-commerce-2021-report-ITA__WEB.pdf

l'81,5%, chi ha visitato un e-store il 91,4%, infine chi ha acquistato un prodotto online il 76,8%¹⁶.

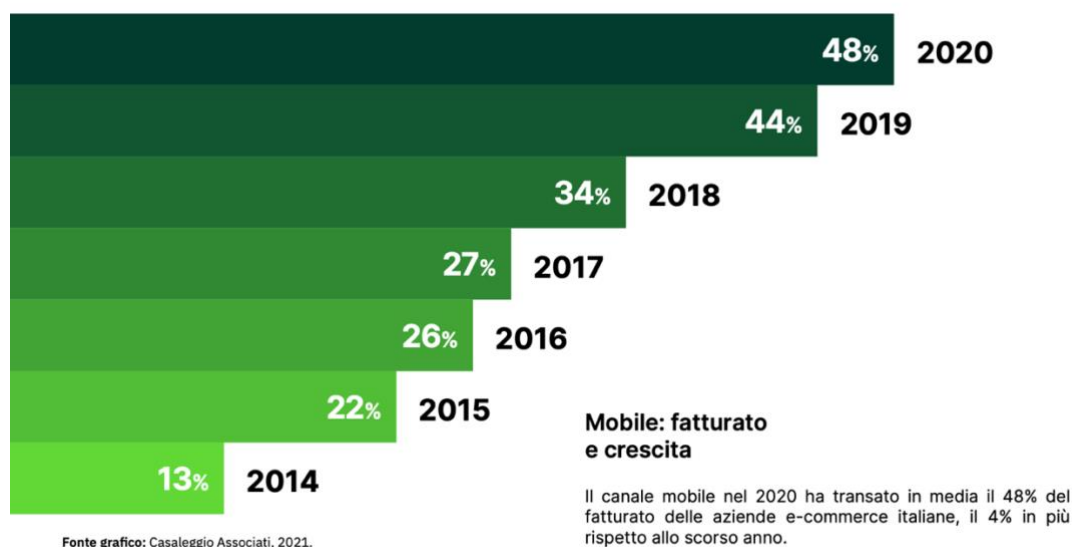


Figura 1.4: Mobile: fatturato e crescita. Fonte: Casaleggio Associati 2021

Secondo le stime del 2021 il fatturato riguardante il servizio e-commerce tra B2C e B2B è pari a 10.780 miliardi. Nel dettaglio il mercato del B2C vale più di 4.000 miliardi di dollari: se si rapporta le vendite e-commerce B2C al totale delle vendite retail, i paesi con più incidenza sono Cina, Corea del Sud, UK, Danimarca, Norvegia e infine USA¹⁷.

In Europa la crescita dell'e-commerce è molto promettente perché circa 480 milioni di persone acquistano già online. I settori principali di vendita sono: alimentare e salute con una quota di 56 miliardi di fatturato, a seguire ci sono casa e arredamento con 60 miliardi, moda con 127 miliardi e infine elettronica di consumo e media con 91 miliardi¹⁸.

Il Regno Unito si è collocato in prima posizione nel 2020 come mercato e-commerce più grande in Europa. Brexit e pandemia hanno determinato importanti risultati di import/export tramite il servizio dell'e-commerce. Il 36% degli utenti ha acquistato all'estero, invece le vendite hanno registrato un incremento su base annua del 57%¹⁹.

¹⁶ Casaleggio: E-commerce in Italia 2021: https://www.casaleggio.it/wp-content/uploads/2020/12/CA-E-commerce-2021-report-ITA__WEB.pdf

¹⁷ Casaleggio: E-commerce in Italia 2021: https://www.casaleggio.it/wp-content/uploads/2020/12/CA-E-commerce-2021-report-ITA__WEB.pdf

¹⁸ Casaleggio: E-commerce in Italia 2021: https://www.casaleggio.it/wp-content/uploads/2020/12/CA-E-commerce-2021-report-ITA__WEB.pdf

¹⁹ Casaleggio: E-commerce in Italia 2021: https://www.casaleggio.it/wp-content/uploads/2020/12/CA-E-commerce-2021-report-ITA__WEB.pdf

In Spagna il dato sugli acquisti online è in crescita: il 15% della popolazione acquista almeno una volta alla settimana online; mentre il 4% afferma di acquistare almeno una volta al mese. Il secondo mercato in cui gli acquisti online sono frequenti è il Belgio, seguono rispettivamente Italia, Paesi Bassi e Polonia. Quest'ultima nell'anno 2019 ha lanciato 11 mila nuovi siti di e-commerce nelle proprie piattaforme di rete. La Germania, nel corso del 2020, ha incrementato il settore dell'e-commerce del 14,6% ossia all'incirca 83,3 miliardi di euro. I dati rilevano infatti che 4 persone su 10 acquistano online e 1 su 3 è un cliente con più di sessanta anni d'età²⁰.

È importante sottolineare che nel 2020, i paesi Europei hanno affrontato problemi riguardanti la logistica, i costi investiti ammontano a 38, 64 miliardi di euro²¹. Questi costi sono dovuti all'introduzione di: nuovi strumenti e tecnologie per cercare di soddisfare le domande da parte dei clienti e le loro richieste; potenziamento delle linee e servizi di consegna per garantire il servizio offerto adottando strumenti che aiutino nella gestione di ordini di dimensione maggiore cercando di consegnare i prodotti nello stesso giorno in cui vengono richiesti. Le imprese per restare in linea con le grosse richieste del mercato virtuale iniziarono ad implementare il settore della logistica con soluzioni di automatismo con la quale era possibile rendere la gestione più flessibile e sostenere gli alti volumi di domande.

1.5: FATTORI CRITICI E DI SUCCESSO DELL'E-COMMERCE

È importante sottolineare che la creazione di un sito E-commerce per una società è una grossa opportunità, perché grazie ad esso si attiva un nuovo canale di vendita. Si possono quindi già specificare due vantaggi rispetto la vendita in un canale tradizionale face to face:

- Temporale: il sito e-commerce è aperto a tutti e in ogni momento della giornata, in modo tale che il cliente possa acquistare direttamente sul portale quando ne necessita;

²⁰ Casaleggio: E-commerce in Italia 2021: https://www.casaleggio.it/wp-content/uploads/2020/12/CA-E-commerce-2021-report-ITA__WEB.pdf

²¹ Casaleggio: E-commerce in Italia 2021: https://www.casaleggio.it/wp-content/uploads/2020/12/CA-E-commerce-2021-report-ITA__WEB.pdf

- Spaziale: grazie ai sistemi di trasporto che negli anni si sono evoluti, gli utenti che intendano acquistare un prodotto dal sito E-commerce, possono riceverlo anche se residenti in qualsiasi parte del mondo.

Il cliente ha la possibilità di visitare il catalogo dei prodotti offerti dall'impresa che utilizza un sito web, conoscerne il relativo prezzo, infine reperire prodotti di vario genere non reperibili localmente.

La vendita tradizionale rispetto la vendita sul canale online è legata al servizio, al potere del brand e alla fiducia che i dipendenti del negozio fisico instaurano con la clientela. Per un sito E-commerce è diverso; è importante avere una buona posizione del sito web sui motori di ricerca, dato che la maggior parte degli utenti, non cerca direttamente un brand specifico ma ricercano direttamente il prodotto di cui ha necessità o è interessata. È importante quindi che un'impresa ottimizzi il proprio servizio E-commerce in ottica SEO.

A sostegno dell'E-commerce c'è anche la velocità con cui gli acquisti vengono effettuati. Grazie alla possibilità di filtrare velocemente la ricerca sui portali online, il cliente può indirizzarsi direttamente sul prodotto desiderato, semplificando il processo di acquisto. In aggiunta i nuovi metodi di pagamento elettronici diventati più sicuri con il passare degli anni facilitano l'accessibilità e la praticità con gli acquisti online. Ad aumentare inoltre la soddisfazione dell'utente che acquista online sono le spedizioni rapide che, se effettuate nel minor tempo possibile, producono una sensazione di fiducia verso l'impresa venditrice.

Per le aziende che usufruiscono di un servizio e-commerce, è importante introdurre il concetto di CRM (Customer Relationship Manager). È uno strumento che favorisce la raccolta e la gestione di dati anagrafici, comportamenti d'acquisto e le preferenze di consumo. L'impresa grazie all'utilizzo del CRM riesce a registrare le informazioni di contatto, l'e-mail, la data di nascita, il numero di telefono e molti altri dati anagrafici, che se organizzati correttamente possono creare un riassunto completo sulle persone con cui l'azienda interagisce e comprendere come quest'ultime cambiano nel tempo i propri comportamenti d'acquisto. L'impresa può dunque mappare una strategia di vendita che rispecchi le esigenze di ogni consumatore, offrendo un servizio target ed una assistenza prima e dopo l'acquisto di un prodotto da parte della clientela. Capire chi sono i "clienti tipo" del proprio negozio, cosa vogliono, quali sono i loro interessi e i loro gusti è fondamentale anche per realizzare delle comunicazioni tagliate su misura per loro, ad

esempio, inviando messaggi mirati nel momento stesso in cui navigano nel sito e-commerce oppure offrendo specifici sconti. Maggiore è la soddisfazione dei bisogni dei clienti, più loro daranno fiducia ed acquisteranno. Alcuni esempi principali di strumenti di CRM utilizzati dalle imprese sono Salesforce, Zendesk e Hubspot.

Per una impresa, utilizzare un sito e-commerce per le vendite può generare ulteriori vantaggi come:

- Risparmio e riduzione dei costi: si possono ottenere modesti risparmi sui costi a livello di supply chain, di gestione degli ordini, di comunicazione e promozione alla clientela. Grazie ad un sito e-commerce, inoltre, si possono gestire nel minor tempo possibili i servizi post-vendita, diminuendo così i relativi costi; infine si possono risparmiare quelli che vengono definiti costi immobiliari, quindi senza dover aprire negozi fisici: l'e-commerce può essere infatti un ottimo rimedio senza dover effettuare costi elevati di investimento.
- Vendita diretta senza intermediari: grazie ad un sito E-commerce si può risparmiare sul personale, atto alla vendita dei prodotti nel punto vendita fisico;
- Gestione dei prodotti in magazzino: per un'impresa l'utilizzo di un e-commerce può essere profittevole perché non sempre c'è bisogno di un magazzino nella gestione degli ordini. In caso ci fosse però è comunque un vantaggio per l'impresa che, grazie ai dati raccolti online, può gestire meglio i propri prodotti, ad esempio capendo quale articolo è meno venduto al fine di proporlo ad un prezzo stracciato. In tutti i casi l'e-commerce facilita l'uscita delle rimanenze del magazzino ed evita inutili accumuli di scorte.

Nonostante i numerosi vantaggi, esistono anche degli svantaggi derivanti dall'e-commerce sia per i venditori che per gli acquirenti:

- Mancanza di un contatto face to face: risulta così difficile instaurare un rapporto con il cliente finale come lo si può stabilire con un negozio fisico e viene a mancare tutto il contesto del rapporto umano e della creazione della fiducia nei confronti del venditore. Inoltre, il cliente online non ha la possibilità di provare il prodotto e quindi molte volte può essere scontento dell'acquisto perché non comprendere adeguatamente le caratteristiche di ciò che sta acquistando;

- Ritardo nella spedizione: uno dei problemi che ancora oggi affligge l'utilizzo di un mercato virtuale sono le difficoltà legate alle spedizioni dei prodotti che a volte sono in ritardo o non avvengono proprio;
- Periodi di down del sito e-commerce: se un sito spesso è fuori servizio, può provocare la perdita di clienti o chi è in visita nel sito non può procedere all'acquisto del prodotto desiderato. Di conseguenza si crea anche una reputazione negativa nei confronti del brand che può causare seri danni alle vendite.
- Un ultimo svantaggio è quello derivato dalla possibilità del cliente di paragonare i prodotti desiderati online. In un negozio fisico, infatti, difficilmente un cliente ricerca prezzi di altre marche perché viene indirizzato ed aiutato dal commesso che lavora per quel determinato negozio. Se da una parte questo risulta essere un vantaggio per l'acquirente che ha la possibilità di comprare in modo più consapevole, dall'altra parte risulta un problema per i venditori perché subiscono maggiormente l'influenza dei competitors.

In generale quindi, l'adozione di un sito e-commerce per un'impresa può essere una opportunità di vendita in più da sfruttare. L'azienda deve riuscire a gestire le nuove tecnologie e a sfruttare i dati raccolti dall'utilizzo di strumenti come il CRM, per avvicinarsi alle esigenze dei clienti e creare soluzioni personalizzate. Inoltre, uno dei fattori di successo per un sito e-commerce risulta essere la sua corretta indicizzazione online grazie alle tecniche di SEO (Search Engine Optimization) che permettono un aumento della brand image e della web reputation.

L'università degli studi di Brescia ha identificato le variabili che impattano di più sulle decisioni d'acquisto e sono²²:

- divertimento e soddisfazione personale nell'acquisto;
- sicurezza e fiducia ispirata;
- utilità percepita dal cliente;
- influenza dei social network;
- piena disponibilità di internet e della tecnologia;
- competenza dei navigatori;

²² Gli elementi maggiormente impattanti sulla propensione all'acquisto online, Ricerca realizzata per conto di Aicel, Università degli studi di Brescia anno 2019

- facilità d'uso del sito;
- esperienza nell'acquisto online.

È importante, dunque, che il sistema e-commerce interagisca il più possibile con i clienti per far sì che questi una volta entrati nel portale online o dopo un acquisto ritornino ad effettuare acquisti continui. È fondamentale sfruttare l'omnicanalità ossia la centralizzazione del consumatore che deve vivere la stessa esperienza su tutti i punti di contatto che l'azienda offre quindi sia fisico tramite negozio che online tramite e-commerce. Questo argomento verrà però approfondito maggiormente nei capitoli successivi.

CAPITOLO II: LA GRANDE DISTRIBUZIONE ORGANIZZATA

2.1: STORIA DEL COMMERCIO IN ITALIA E NEL MONDO DAGLI ANNI 60' AD OGGI

In Italia, prima della nascita dei supermercati la vendita dei generi alimentari era caratterizzata da una struttura commerciale polverizzata, formata da produttori che vendevano al dettaglio i propri prodotti in piccoli negozi; quest'ultimi erano concentrati per la maggior parte nelle vie centrali delle città o dei paesi più importanti. Ad ogni modo la modernizzazione del settore distributivo italiano cresceva a rilento rispetto agli altri Paesi in Europa, a causa di problemi del sistema socioeconomico e, dagli anni 70' in poi, per le barriere all'ingresso nei settori amministrativo e legislativo.

Le principali modernizzazioni che hanno caratterizzato l'evoluzione del commercio in Italia sono le seguenti:

Prima modernizzazione

Nel lontano 1957 nasce il primo supermercato in Italia precisamente a Milano, chiamato "Supermarket italiani" che, con il passare degli anni, si trasformò nel colosso Esselunga²³. Negli anni Sessanta, dunque, cominciano i primi commerci "self service" dove il cliente può recarsi all'interno del supermercato ed acquistare ciò che desidera, pagando alla cassa. Grazie alla grossa offerta di prodotti e alle nuove tecnologie per la conservazione, i consumatori iniziano ad acquistare generi alimentari o prodotti di uso comune facendo stock nelle proprie case. I supermercati, dunque, iniziano ad acquisire vantaggi di costo che non erano conseguibili dai piccoli dettaglianti. Alla fine degli anni Sessanta, si possono contare 369 supermercati e nessun ipermercato aperto, chiaro segno della lenta crescita di questo settore produttivo in Italia. La situazione comincia a cambiare attorno agli anni Ottanta, raggiungendo un numero di circa 1321 supermercati su tutto il suolo nazionale e una quota di mercato pari a 16,1% come stava succedendo anche in Spagna. Lo sviluppo delle forme distributive procede a rilento a causa delle barriere all'entrata imposte dal settore pubblico provocando:

- Un rallentamento di forme distributive come gli ipermercati;

²³ Attualmente quest'ultimo gestisce quasi l'8,7% delle vendite in supermercati e ipermercati con all'incirca 170 negozi aperti in tutta la penisola.

- La sottodimensione dei supermercati, rispetto alla loro dimensione ideale;
- Mancata omogeneità della struttura economica e sociale da parte delle regioni italiane, causando uno sviluppo diverso dei supermercati nelle varie aree della penisola;
- Scarso sviluppo di canali ed aperture che non erano abbastanza ottimali dal punto di vista della logistica.

Seconda fase di modernizzazione

Dopo gli anni Sessanta, con il boom economico e con l'inizio del consumismo, incominciano una serie di sviluppi tecnologici-strutturali che incentivano la crescita dei supermercati. Innanzitutto, vengono costruite sempre più strade e sempre più agevoli, permettendo lo spostamento di un gran numero di persone dalle aree di campagna a quelle urbane. In secondo luogo, si sviluppano nuovi mezzi di trasporto come le auto e, a mano a mano che il mercato automobilistico cresce, anche i prezzi di acquisto diventano più abbordabili per le famiglie italiane. Infine, il fattore più decisivo è il cambiamento di mentalità della popolazione che, avendo una disponibilità economica maggiore rispetto al passato, inizia a fare la spesa periodicamente, spostandosi uno o due giorni alla settimana.

Durante gli anni Ottanta, i supermercati in Italia triplicarono arrivando all'incirca a 4490 unità, dando il via però anche a diversi nuovi problemi da affrontare. I bacini di domanda formati da più punti vendita ormai sempre più moderni tendevano a sovrapporsi, iniziando una competizione tra le grandi aziende di distribuzioni che offrivano sostanzialmente gli stessi prodotti. Ne deriva la necessità di differenziazione tra concorrenti, sia dalle formule di vendita, sia dalle caratteristiche delle insegne con medesima forma. Le caratteristiche dei supermercati tradizionali italiani, alla fine degli anni '90 erano infatti le seguenti:

- Dimensioni ridotte;
- Non si attuava una competizione sui prezzi ma solo sulla qualità dei prodotti venduti;
- Non si erano ancora diffuse le marche commerciali, quindi si vendevano per la maggior parte prodotti con marca industriale;

Iniziano quindi a spuntare nuove forme distributive, che vengono importate dai Paesi esteri come: gli Ipermercati negli anni '80 e i Discount negli anni '90. Si tratta di due strutture totalmente diverse caratterizzati da aspetti molto vantaggiosi:

- Una elevata superficie di vendita, dai 2500 m² (dimensione necessaria a quel tempo per alcuni paesi Europei tranne che per Italia e Francia) ai 20000 m²;
- Un assortimento dei prodotti in scaffale ampio che potevano essere sia prodotti alimentari che no;
- La dislocazione in aree non urbane ma periferiche, ad esempio appena fuori dalle autostrade, in modo tale che i clienti lo potessero raggiungerlo facilmente;
- Un prezzo di vendita inferiore a seguito di un basso costo d'acquisto grazie all'ampia disponibilità di prodotti, e alla elevata rotazione delle scorte, a seguito di un basso costo d'acquisto da parte dell'ipermercato;

La formula distributiva però non era un elemento unico per differenziarsi dai diversi competitor e si cerca quindi di spingere verso un'identità di insegna. Le insegne dei supermercati erano molto simili, dunque le grandi distribuzioni di prodotti alimentari e no, si potevano differenziare riducendo il peso dei marchi industriali nell'assortimento e sviluppando una capacità di progettazione autonoma di prodotti. Grazie allo sviluppo tecnologico dei sistemi informativi, supermercati ed ipermercati creano la propria marca commerciale per i prodotti di consumo, divenendo uno strumento competitivo fondamentale. Da semplice intermediario di prodotti, le grandi distribuzioni diventano autonome e cercano di soddisfare direttamente i bisogni della domanda da parte della clientela, sottraendo alcune funzioni di marketing attuate dalle industrie di beni commerciali. Cominciano dunque ad avere un'attenzione più mirata verso il consumatore, sviluppando nuove insegne, ma facendo anche pubblicità e promozioni di acquisto per incentivare le persone ad acquistare i prodotti alimentari e no.

Terza modernizzazione

L'attenzione verso una relazione con il cliente negli anni ha permesso la diffusione di un sistema E-commerce, dove sostanzialmente cambia la modalità di acquisto perchè gli acquisti non tendono più ad essere effettuati fisicamente nel punto vendita, ma tramite l'utilizzo della tecnologia come smartphone, tablet e computer. Inoltre, la spesa viene

direttamente portata a casa o prelevata ad un punto di ritiro spesa specifico, eliminando totalmente l'atto fisico di recarsi al negozio per fare acquisti.

I primi tentativi di spesa online risalgono al 1984 con Tesco, catena di negozi britannica, e al 1989 con Peapod, negozio statunitense. Con la diffusione nel 1995 dei colossi Amazon ed Ebay, le persone cominciano ad effettuare con continuità acquisti online, grazie anche all'introduzione di nuove tecnologie ed al miglioramento delle reti Internet. Con l'aumento dei canali di vendita, aumenta anche la possibilità dei clienti di soddisfare i propri bisogni di acquisto, di conseguenza migliorano le relazioni tra aziende e clienti grazie all'introduzione di più touch point. Le aziende cominciano a incrementare le promozioni e pubblicità non solo in modalità fisica ma anche online, sviluppando nuove attività di marketing.

L'aumento del numero di canali a disposizione, da una parte quello "fisico" e dall'altra la possibilità di accedere in qualsiasi momento della giornata al canale online, diminuiscono le barriere all'uscita da parte delle aziende, su cui quest'ultime puntavano per mantenere la "customer retention" cioè il mantenimento o conservazione di una relazione tra consumatore ed azienda.

Ciò che emerge dalla terza modernizzazione è che l'utilizzo di un sito E-commerce da parte delle aziende, porterà una serie di cambiamenti in futuro soprattutto nei settori grocery:

- Avranno origine fenomeni di acquisizione o fusione tra aziende, principalmente per chi vorrà utilizzare nuovi canali di vendita; perciò, serviranno nuove conoscenze o competenze che possono essere ricavate all'esterno, da una realtà consolidata e specializzata;
- Grazie al nuovo canale online, il retailer potrà non solo raggiungere nuovi clienti, ma potrà raccogliere ulteriori informazioni utili che non vengono reperite facilmente sul mercato fisico. Offrendo dunque un'immagine aziendale di affidabilità e qualità per il servizio offerto;
- I retailer cercheranno nuove fonti di efficienza, grazie integrazione di nuovi modelli di business "snelli" e alla volontà di ridurre i costi operativi. Ciò comporterà dunque ad un minor investimento sulle aperture fisiche e di conseguenza ad un ridimensionamento delle superfici di vendita.
- Con l'e-commerce aumenterà la tendenza del ricercare inizialmente un prodotto online, a cui seguirà successivamente la fase d'acquisto

direttamente nel punto vendita storico. Questo potrà avvenire solo quando un retailer riuscirà, tramite il suo sito e-commerce, ad instaurare una sensazione di affidabilità e sicurezza nel cliente.

Ai giorni d'oggi si può parlare di una quarta modernizzazione, caratterizzata dalla omnicanalità. Grazie alla diffusione della tecnologia, dei social network e dall'esperienza accumulata durante gli anni sull'uso dei vari strumenti, si è arrivato ad un punto dove acquistare online ed acquistare dal negozio per alcuni clienti è la medesima cosa. Le aziende cercano di soddisfare sempre di più le esigenze dei consumatori che si ritrovano in mezzo tra una vita analogica e digitale, influenzati sempre di più dai mass e social media. Ciò che emerge dall'omnicanalità è che l'azienda non gestisce più diversi canali ma diversi profili di consumatori, interagendo tramite vari mezzi di comunicazione.

2.2: LE VARIE FORME DISTRIBUTIVE NEL MONDO

La Grande Distribuzione Organizzata è un concetto che fa riferimento al sistema di vendite al dettaglio, nello specifico, si parla di reti di supermercati o altri canali di vendita di forma differente, che vanno a sostituire quello che era definito come negozio tradizionale. La Treccani, infatti, conferisce una definizione al termine GDO come: “realizzata tramite la concentrazione dei punti vendita in grandi superfici (non minori di 200 m² ma che arrivano anche a superare 4000 m²) e la gestione a carico di catene commerciali che fanno capo a un unico marchio”²⁴.

All'interno della macro insieme della GDO, si possono distinguere due acronimi:

- Grande distribuzione (GD): struttura centralizzata, dove esiste un'unica proprietà, la quale gestisce direttamente le strutture secondarie ed i punti vendita, determinando le strategie operative, gestionali e di marketing;
- Distribuzione organizzata (DO): definita anche come distribuzione associata, dove più aziende si uniscono tramite consorzio, formando un gruppo di acquisto come ad esempio per Sigma, Crai e Conad.

Nella GDO si possono incontrare varie formule organizzative, che si differenziano principalmente per: dimensione, assortimento, prezzi, caratteristiche delle esposizioni al proprio interno ed infine servizi offerti ai propri clienti, come riportato nella figura 2.1.

²⁴ Inside Marketing: GDO: <https://www.insidemarketing.it/glossario/definizione/gdo/>

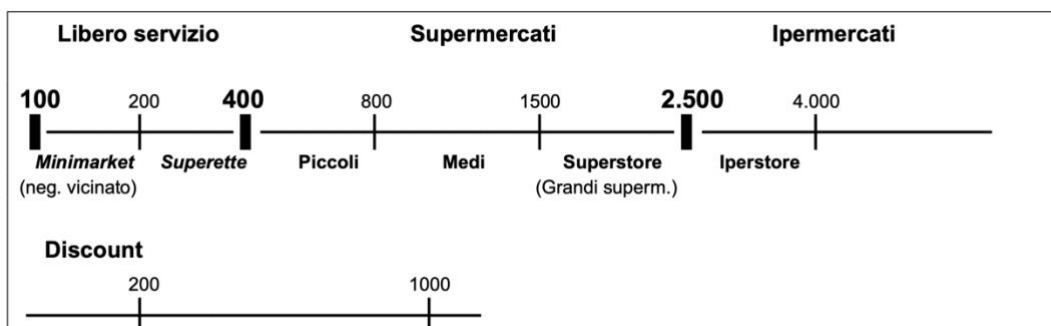


Figura 2.1 Tipologie di vendita al dettaglio in base alla dimensione. Fonte: InsideMarketing

Le principali formule organizzative che si possono incontrare sono:

SUPERETTE

Le superette si sviluppano maggiormente a fine anni 90', con all'incirca 8000 punti vendita²⁵. La dimensione del punto vendita aumentare e LS (libero servizio) varia tra i 200 e 400 m². Le superette forniscono un'offerta accessibile a clienti che intendono effettuare acquisti di integrazione, senza perdere molto tempo. La dimensione ridotta rispetto ai grandi dettaglianti permette alla superette di avere una interazione maggiore con i clienti, aumentando così il loro grado di fiducia e limitando dunque i pericoli che possono incombere con supermercati o ipermercati nelle zone vicine. Si possono trovare due tipi di superette:

- Superette tradizionale: svolgono una funzione, come citato prima, di integrazione della spesa dei clienti. Sono collocate in piccoli paesi o nelle zone limitrofe delle città e utilizzate da coloro che non intendono spostarsi verso le grandi distribuzioni;
- Superette evolute: simile alla superette tradizionale e con la stessa mansione, ma con la differenza che sono collocate in punti di passaggio strategici verso cui confluiscano molti potenziali clienti. Si propongono come soluzione alla spesa giornaliera integrativa, non facendo caso alla presenza di ipermercati e supermercati nelle vicinanze;

Esempi di superette possono essere Sigma che utilizza altri due marchi chiamati Ok e Super Day. Il marchio Pam, allo stesso modo, utilizza per esempio l'insegna

²⁵ ENRICO COLLA, *Gli ipermercati: sviluppo e maturità delle grandi superfici di vendita in Italia e in Europa*, Milano, ETAS libri, 1992, pag. 145

Metà Supernegozi per le superette tradizionali e Pam Express per le superette evolute.

SUPERMERCATI

I supermercati sono punti vendita al dettaglio con dimensioni comprese tra i 400 m² e i 2500 m². Si collocano maggiormente nelle aree urbane e sono dotate di superfici superiori rispetto le superette citate prima, inoltre possono essere inserite all'interno di un contesto come quello dei centri commerciali. All'interno dei supermercati si offre all'incirca il 90% di prodotti referenze di largo consumo (grocery), il cui prezzo si colloca ad un livello intermedio rispetto alla media dei negozi al dettaglio. Negli anni si sono sviluppate varie strutture di supermercato come ad esempio:

- Supermercati discount: la cui dimensione e prodotti collocati in assortimento sono simili ad un supermercato tradizionale. La differenza però sta nelle strategie di leadership dei costi attuate dalle imprese. Il livello del servizio ed il prezzo medio delle referenze vendute sono leggermente inferiori rispetto ad un punto di vendita al dettaglio tradizionale.
- Supermercati qualitativi: essi adottano una strategia di differenziazione rispetto agli altri punti vendita al dettaglio, si distinguono infatti grazie alla presenza di prodotti in assortimento più strutturati. Di conseguenza hanno un prezzo di vendita al di sopra della media generale.

IPERMERCATI:

Come citato nei capitoli precedenti, generalmente la dimensione degli ipermercati può variare dai 2500 ai 4000 m². Il minimo di 2500 m² è una condizione fondamentale, però non sufficiente in Italia ed in Francia²⁶. In Germania invece l'appellativo di Ipermercato è riservato solo ai negozi al dettaglio che hanno una superficie maggiore di 4000 m². Gli aspetti che differenziano l'ipermercato dalle altre forme distributive sono:

- Profondità di assortimento: l'obiettivo principale è quello di vendere grosse quantità rispetto, ad esempio, ai supermercati e per questo si possono trovare diversi generi di prodotti alimentari, anche provenienti dall'estero;

²⁶ ENRICO COLLA, *La grande distribuzione in Europa: evoluzione delle formule distributive, strategie e strutture aziendali, rapporti con l'industria*, Milano, ETAS libri, 1995, pagina 96

- Localizzazione: se supermercati e superette sono distanziati in aree urbane, gli ipermercati sono collocati in zone extra-urbane perché necessitano di ampi spazi o in aree dove la circolazione dei veicoli automobilistici è maggiore, per esempio fuori dai caselli delle autostrade.
- Parcheggio di grosse dimensioni: rispetto a un supermercato, i clienti hanno bisogno di parcheggi per le autovetture molto più agevoli, dato che la spesa media è maggiore rispetto ai punti vendita al dettaglio tradizionali.
- Orari di apertura: gli ipermercati tendenzialmente sono aperti tutti i giorni della settimana, con orari continuati. Nei piccoli supermercati invece ciò non avviene perché sono aperti dal lunedì al sabato con orari prestabiliti.
- Inserimento dell'ipermercato all'interno dei centri commerciali: la soluzione, per far sì che l'ipermercato non sia isolato, è quella di inserirlo all'interno di un centro commerciale o galleria commerciale, rendendolo attraente grazie al flusso di gente maggiore.

DISCOUNT

Come viene evidenziato dagli esempi citati prima le forme distributive si differenziano le une dalle altre per la dimensione di vendita, l'assortimento, la presenza di prodotti con marchio commerciale ed infine per il prezzo di vendita. I discount principalmente si dividono in due tipologie:

- Hard discount: in Gran Bretagna vengono regolarmente chiamati "limited assortment store", la loro superficie di vendita varia tra 350 e 800 m²²⁷, il numero di referenze si aggira attorno ai 500 e 1000 e i prezzi di vendita sono inferiori del 15-30% rispetto ai prezzi medi che si possono trovare nei supermercati. Nei discount vengono offerti prevalentemente prodotti di marchio commerciale; i marchi industriali invece non superano il 20% nell'assortimento, inoltre quest'ultimi non sono di prima fascia, ma di seconda o terza. Le insegne che utilizzano questa forma distributiva di Hard discount sono le tedesche come Lidl ed Aldi o la danese Netto.
- Soft discount: in Gran Bretagna viene chiamata "discount supermarket" con un assortimento più grande rispetto al precedente tra le 1500 e 2000 referenze. C'è

²⁷ ENRICO COLLA, *La grande distribuzione in Europa: evoluzione delle formule distributive, strategie e strutture aziendali, rapporti con l'industria*, Milano, ETAS libri, 1995, pagina 85

una prevalenza di prodotti surgelati, congelati e freschi, e di marca industriale che avendo più spazio nell'assortimento a volte supera anche il numero dei marchi commerciali. La differenza di prezzo rispetto i supermercati è di circa 5-10%. A differenza degli hard discount la localizzazione varia in base alla dimensione; quindi, si possono trovare sia nei centri urbani di piccola estensione oppure, se al contrario hanno una misura più elevata, si collocano nei centri commerciali o in quartieri secondari.

- Supermercati discount e Warehouse store: sono due formule che mantengono comunque un margine di vendita tra il 5-15%, la differenza sta sulle maggiori dimensioni rispetto l'hard e il soft discount. I supermercati discount hanno una superficie che varia fra i 1000 e 2000 m², mentre il warehouse store fra 2000 e 4000 m²²⁸. Il numero delle referenze è nettamente maggiore rispetto i discount citati precedentemente, infatti il numero varia tra le 5000 e 10000 referenze di prodotto. Le tipologie di prodotto che vengono vendute sono: gastronomia, pasticceria, ortofrutta, surgelati, articoli per la casa. È importante inoltre sottolineare che le referenze vendute sono per la maggior parte prodotti di marca industriale e non commerciale. Quindi in linea generale, soft discount, supermercato discount e warehouse sono simili; invece, l'hard discount è più diversificato.

Tra le due tipologie di discount citate, la forma hard è utilizzata spesso in Italia e Spagna, invece la soft in Gran Bretagna, Belgio e Germania. Il vantaggio dei discount è quello di avere prezzi inferiori di vendita rispetto le altre forme distributive. Durante le negoziazioni si cerca di eliminare ogni costo superfluo, ma non solo perché, grazie all'utilizzo di impianti logistici all'avanguardia e alla grossa quantità acquistata, si possono ottenere più sconti sul prezzo dei prodotti acquistati. Alcuni discount, come in Gran Bretagna e USA, attuano una politica di "everyday low price", tradotto, "prezzo basso tutti i giorni": ciò avvantaggia le aziende a gestire in modo ottimale le scorte e la rotazione della merce sugli scaffali di vendita, evitando perciò di gestire le numerose promozioni di acquisto.

²⁸ ENRICO COLLA, *La grande distribuzione in Europa: evoluzione delle formule distributive, strategie e strutture aziendali, rapporti con l'industria*, Milano, ETAS libri, 1995, pagina 86

2.3: LE NUOVE TENDENZE

L'avvento dell'E-commerce nei supermercati ed ipermercati e con la recente pandemia che ha colpito tutto il mondo, sta cambiando la natura di acquisto della spesa da parte delle persone. Le forme distributive, dunque, cominciano a rispondere a nuove richieste da parte della clientela, infatti come citato da Giuseppe Stigliano, docente di Retail Marketing Innovation alla IULM e CEO di Wunderman Thomson Italy: “Stiamo passando da *'make people want things'* a *'make things that people want'*”²⁹. Ciò a cui si riferisce è il passaggio da un mercato dove le aziende cercavano di incentivare le persone a comprare ad uno dove le aziende realizzano i propri prodotti o servizi in base alle richieste di mercato.

Con l'inizio della pandemia, la digitalizzazione è divenuta un elemento essenziale per poter acquistare prodotti o servizi senza dover spostarsi da casa. Per soddisfare la forte richiesta di beni alimentari da parte dei clienti, i quali erano costretti a rimanere nelle proprie abitazioni a causa delle restrizioni imposte dagli Stati per la pandemia Covid-19, i supermercati ed ipermercati hanno scelto come soluzione ottimale l'utilizzo di siti e-commerce. Viene evidenziato nella figura 2.2 l'evoluzione dei consumi post-Covid 19, e viene reso noto come le vendite siano aumentate esponenzialmente dal 2019 ad oggi, con un aumento di percentuali come il numero di famiglie acquirenti o come la spesa media di famiglia.



Figura 2.2, Totale Largo Consumo Confezionato. Previsioni IRI condotte a livello di macrocategoria ECR prima dell'evento pandemico da Covid-19. Elaborazioni RemLab su dati GFK panel famiglie. Fonte: IRI Online.

²⁹ Retail e GDO: trend, innovazione, cambiamenti e best practice: <https://www.tuttofood.it/news/news/retail-e-gdo--trend--innovazione--cambiamenti-e-best-practice.html>

Dal 2020 le aziende di distribuzione cominciano a migliorare i sistemi di E-commerce, per rendere il servizio sempre più efficiente. Un esempio è il gruppo COOP che nel 2021 ha costruito un impianto di packaging completamente automatizzato in Svizzera, sfruttando imballaggi di dimensione effettiva del prodotto in modo tale da limitare al minimo lo spreco di materie prime come la carta.

Ma non solo, già dalla presentazione fatta all'Expo del 2015, il gruppo COOP aveva già le idee chiare su come potesse essere un supermercato del futuro. Con la partecipazione di Accenture e Avanade, la COOP ha sviluppato un modo per migliorare le esperienze di acquisto, impegnandosi a soddisfare una delle esigenze principali dei clienti: l'informazione, in modo semplice ed intuitivo. Lo slogan quindi che racchiude tutto ciò che si è descritto su questo capitolo è: "l'esigenza chiave dell'era digitale: ottenere informazioni chiare, dettagliate, finalizzate a coinvolgere in modo semplice ed intuitivo"³⁰.

Un esempio può certamente essere il modello di supermercato predisposto da COOP a Milano, il quale potrà essere preso come prototipo da altre aziende. Quest'ultimo infatti dispone di novità strutturali e innovative come:

- **Banchi interattivi:** il cliente può tramite un semplice monitor inquadrando il codice a barre della referenza in questione, visualizzare tutte le caratteristiche di composizione del prodotto che si vuole acquistare, il prezzo, le promozioni, prodotti che possono essere sostitutivi ad esso;
- **Scaffali verticali:** il layout degli scaffali sono progettati in modo tale da favorire l'attenzione verso le referenze e reperire facilmente i prodotti interessati. Questi ultimi inoltre vengono contrassegnati da cartelli che evidenziano le promozioni in corso.
- **Visualizzazione dei dati in tempo reale:** sulla parete del supermercato COOP vengono installati all'incirca 54 monitor, nei quali vengono mostrati i dati e contenuti in tempo reale come i prodotti che maggiormente vengono venduti, le offerte e promozioni in corso e le ultime notizie che vengono divulgate anche nei social come Facebook ed Instagram.
- **La struttura del supermercato:** con l'utilizzo di materiali come il legno e metallo e scaffali "a vista" l'azienda cerca di accogliere i propri clienti con un'atmosfera

³⁰ GDO e digitalizzazione come sarà il supermercato del futuro?: <https://www.meiko.it/it/mondo-meiko/blog/gdo-e-digitalizzazione-come-sara-il-supermercato-del-futuro>;

calda e rilassante. Come espresso dall'architetto progettista di Inres, Francesco Scanu: “Abbiamo pensato a uno store da vivere oltre la spesa, come esperienza a tutto tondo legata al mondo del cibo”³¹.

Recentemente l'azienda COOP ha introdotto anche due nuovi servizi: Fiorfood Cibo & Incontri, dove viene adibito uno spazio per la ristorazione e Coop Drive dove è possibile ordinare la propria spesa online e riceverla entro due ore dalla commissione in negozio senza scendere dal proprio autoveicolo.

È importante dunque sottolineare come COOP cerchi di soddisfare le richieste dei propri clienti in moltissime modalità, rendendo facile anche effettuare la spesa. Come viene anche descritto da Alberto Pozzi, Amministratore delegato della Retail Practice Accenture: “Con il Supermercato del Futuro di Coop realizziamo quella convergenza tra fisico e digitale che permette di creare un'esperienza di acquisto realmente coinvolgente e immersiva. Coop dà così forma al supermercato del domani, combinando in un'unica esperienza: personalizzazione dei servizi, accesso ad informazioni approfondite sui prodotti e utilizzo di dispositivi connessi, che accompagnano il consumatore nell'acquisto...”. Viene riportata nella figura 2.3, un concept del servizio offerto da Coop nei propri store, in particolare quello di Fiorfood, dove i dettagli a volte possono fare la differenza.



Figura 2.3: Coop Fiorfood.

³¹ ALFACOD: Coop punta sul “supermercato del futuro”: coinvolgimento ed experience: <https://www.alfacod.it/blog-supermercato-futuro-coop-customer-experience>;

Le nuove tendenze spingono le grosse aziende ad indirizzare le strutture dei propri supermercati per implementare il rapporto con i clienti come dimostra La Esse, aperto di recente a Milano³². La costruzione del supermercato La Esse è circa 400 m², dove al suo interno si possono trovare servizi come bar, ristoranti e spazi di co-working, oltre a poter acquistare prodotti freschi e freschissimi anche da asporto. Si deve aggiungere inoltre il servizio offerto Click & Vai, con il quale il cliente può ordinare la spesa e ritirarla in varie modalità. Ad esempio, il cliente può ritirare la propria spesa nello slot indicato tramite il codice che viene fornito dal supermercato senza ulteriori costi aggiuntivi; oppure tramite assistenza diretta nel punto vendita dove il cliente può effettuare il ritiro della merce direttamente in supermercato; o ancora, similmente a come già evidenziato, il cliente può effettuare il ritiro direttamente nei parcheggi di ritiro e la spesa viene caricata direttamente sull'autovettura. La scelta del servizio Click & collect è una delle migliori soluzioni per il ritiro della propria spesa, infatti come riportato nella figura 2.4, il servizio ha una quota di utilizzo che con il passare degli anni è sempre di più in aumento.

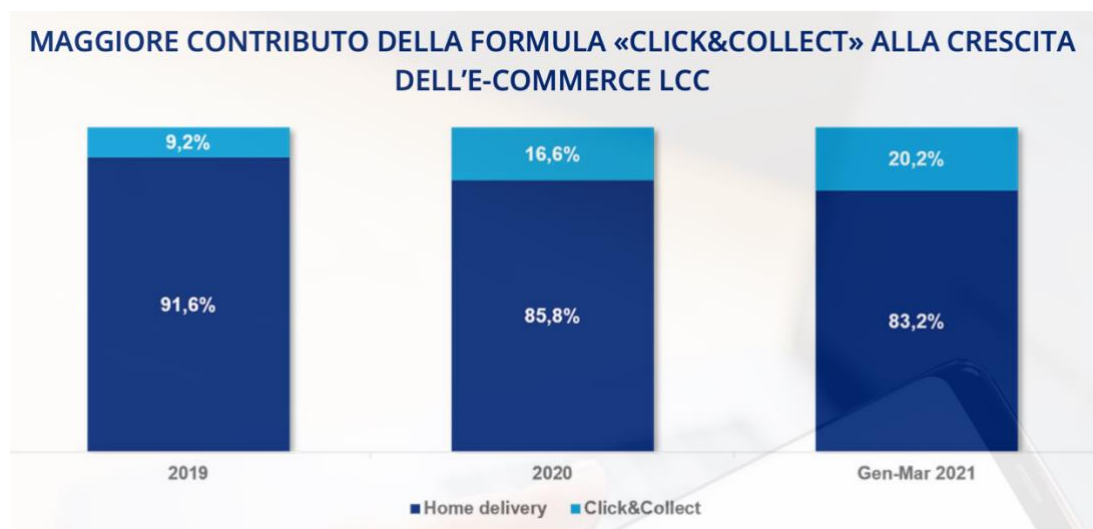


Figura 2.4: Quote delle vendite in base ai servizi Home delivery e Click&Collect. Fonte: IRI Liquid Data

Un'altra novità riguardante il mondo dei supermercati è quella portata da Carrefour, il risto-retail. Infatti, a Parigi dalla compagnia francese nasce Bon-Appétit, un ristorante dedicato alla cucina di prodotti bio e solidali che possono essere consumati sia all'interno della struttura che portati a casa.

³² Il supermercato del futuro è tutti un esperimento: [https://ledhors.it/news/48-il-supermercato-del-futuro-%26egrave%3B-tutto-un-esperimento](https://ledhors.it/news/48-il-supermercato-del-futuro-%26egrave%3B-tutto-un-esperimento;);

Tralasciando il modello risto-retail, La Cloche d'Or, aperto da Auchan in Lussemburgo, ha voluto portare delle innovazioni che hanno fatto aumentare le proprie vendite tra queste vi sono la coltivazione di erbe aromatiche all'interno del complesso o la possibilità di trovare oggetti scenografici come bagagli che girano attorno un nastro trasportatore in cui all'interno vengono esposte le referenze, rappresentando dunque un contesto simile ad un aeroporto. Dunque, Auchan cerca di puntare sulle emozioni che vengono provocate ai clienti nel vedere queste sceneggiature all'interno di un ipermercato grande all'incirca 12000 m².

Infine, le novità arrivano anche per le aziende che nascono online ma che vogliono utilizzare un servizio fisico per aumentare le proprie vendite. Gli esempi più concreti sono Honestbee e Amazon.

La prima, fondata nel 2015, è una azienda che collabora con supermercati e rivenditori per recapitare merci ai clienti utilizzando i suoi store e l'ordine viene fatto interamente con le app del telefono o tablet. L'azienda, che ha sede a Singapore, opera in otto mercati in tutta l'Asia: Hong Kong, Singapore, Taiwan, Thailandia, Indonesia, Malesia, Filippine e Giappone. In alcuni mercati ha esteso il proprio servizio alla consegna di cibo e, a Singapore, gestisce un negozio online/offline in stile Alibaba chiamato Habitat che rispecchia il modello "brick & mortar". Questa modalità online/offline ha attirato un target di consumatori con una media d'età tra i 26 e i 35 anni.

Il colosso del mercato online, Amazon, nel corso 2020 ha acquisito Whole Foods, ed in California ha aperto il suo primo negozio fisico. L'azienda americana vuole puntare a crescere un'attività alimentare utilizzando servizi come Amazon Fresh, un servizio di consegna di prodotti alimentari freschi come: carne, ortaggi di stagione, pesce e prodotti gastronomici preparati giornalmente. Questo servizio inoltre è collegato con l'e-commerce, infatti i clienti possono sia acquistare comodamente da casa la propria spesa o ritirarla direttamente al punto vendita. Un'altra novità è Amazon Dash, i clienti che optano di recarsi fisicamente a fare la spesa, quando inseriscono all'interno del carrello le referenze, sanno quanto stanno spendendo grazie ad una combinazione di algoritmi e sensori che permettono di identificare il prodotto all'interno di esso ed infine hanno la possibilità di pagare senza passare per la cassa tramite la corsia Dash Cart.

È importante evidenziare come Amazon si stia espandendo non solo con continuità nel mercato online, ma cerca anche di ampliare la propria area di vendita anche nelle vendite offline.

È un chiaro esempio di omnicanalità e di come l'unione tra online e offline possa essere importante sia per le aziende sia per i clienti che possono vendere o reperire i prodotti in entrambi i modi³³.

È importante infine evidenziare come riportato nella figura 2.5, come supermercati ed ipermercati abbiano preso quote di mercato, grazie all'utilizzo dell'E-commerce, il discount è la forma organizzativa che ha accumulato una buona quota di crescita come i supermercati, che detengono il primato.

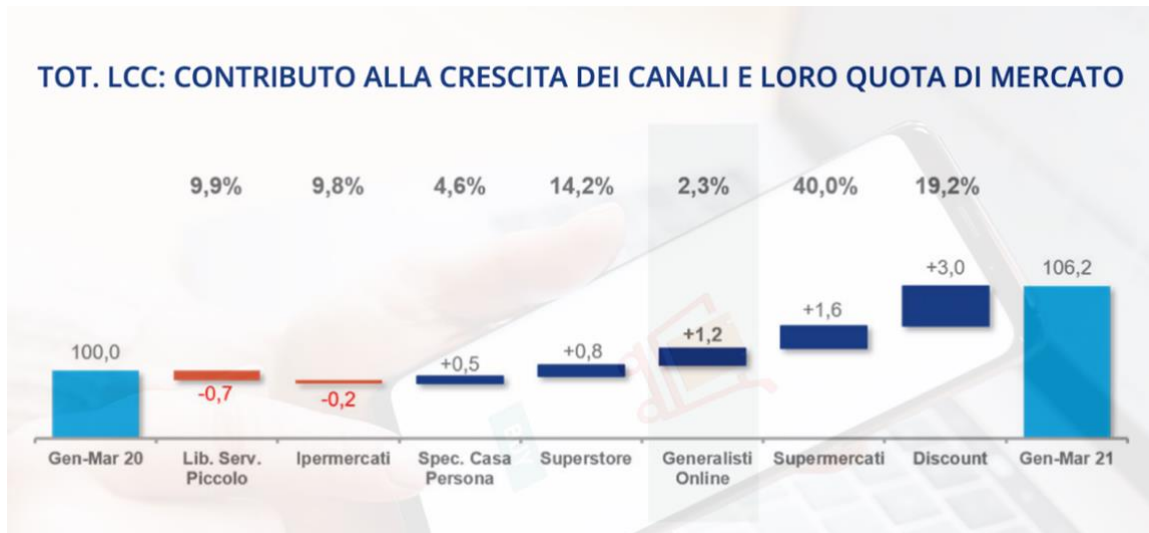


Figura 2.5, Contributo alla crescita dei canali e la loro quota di mercato. Fonte: IRI Liquid Data

2.4: I PRINCIPALI SUPERMERCATI ED IPERMERCATI CHE UTILIZZANO L'E-COMMERCE

Gli oltre 3,5 milioni di Italiani che hanno iniziato ad accedere alle piattaforme online hanno fatto registrare una crescita sul fatturato di oltre 64 miliardi nel 2021, +39% come riportato nella figura 2.6, se da una parte molti piccoli negozi fisici italiani hanno avuto difficoltà nel restare aperti, dall'altra il canale digitale rappresenta una risorsa fondamentale nella strategia di molti colossi alimentari.

³³ Apre in California il primo store di freschi targato Amazon: <https://www.foodweb.it/2020/08/amazon-fresh-il-primo-store-di-freschi-targato-amazon-fresh-apre-in-california/>

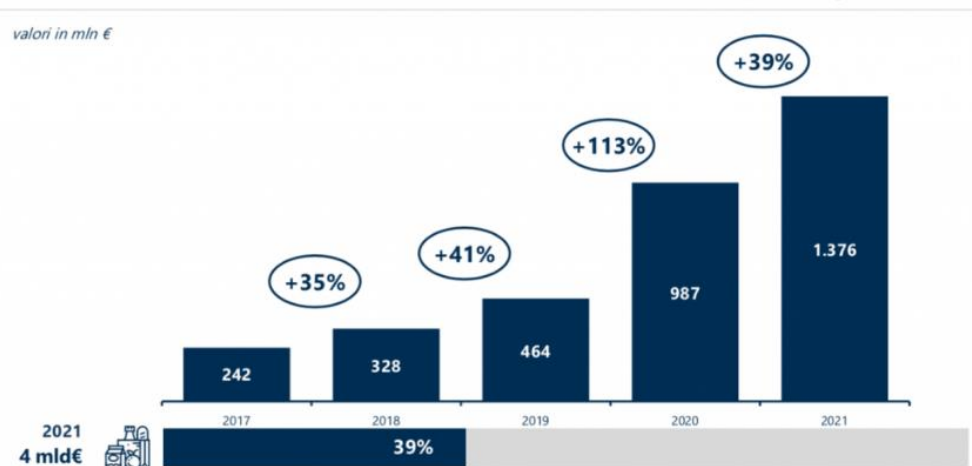


Figura 2.6, Crescita fatturato E-commerce alimentari. Fonte: Dipartimento di ingegneria gestionale, Politecnico di Milano

I 3 principali player che si contendono il mercato online sono:

- Carrefour la catena europea dei supermercati alza il tiro sulla digital transformation, prevedendo da qui a 5 anni un fatturato per il canale online che si aggira attorno al 10 miliardi, l'azienda francese considera questo canale come un asset strategico, sfruttando player del settore logistico come Deliveroo. L'obiettivo è far giungere al cliente la spesa a casa il 15 minuti. L'azienda, a livello consolidato di gruppo, chiude il 2021 con una crescita del +7%, con un fatturato di 79 miliardi di euro, con un EBITDA di 4,4 miliardi (6%), generando nuova liquidità per circa 1 miliardo di euro;
- Coop è un sistema di 109 cooperative operanti nella commercializzazione di prodotti di largo consumo ed alimentari, dagli anni 70' domina il mercato della grande distribuzione e, avendo un ruolo di leader italiano nel retail, ha avuto la necessità di potenziare anche il canale online. L'azienda punta forte sull'e-commerce, passando da 1500 ordini al giorno a 15000 (+1000%), vuole offrire un servizio a 360° al cliente, offrendo non solo prodotti di carattere alimentare, ma anche prodotti che vanno a coprire i settori come tempo libero, elettrodomestici, lavastoviglie, lavatrici, frigoriferi. L'azienda nel 2021 ha registrato un fatturato pari a 14,3 miliardi di euro, con un EBITDA pari a 7 milioni (5% rispetto al fatturato, un indicatore migliore rispetto al colosso nettamente più grande Carrefour).
- Esselunga è la prima società italiana nella GDO; fondata nel 1957 da un imprenditore statunitense, diventa in pochi anni leader nel settore alimentare; accanto alla

distribuzione di beni a largo consumo, è la prima ad offrire prodotti biologici. “Esselunga a casa” è il servizio di spesa online offerta nel nord e centro Italia, servizio che inizia ad offrire nei primi anni 2000, il servizio e-commerce è un percorso di ampio raggio articolato in soluzioni a vari target ed a differenti modalità di spesa. La strategia del colosso è quella di “colpire” zone prettamente turistiche e ad alta densità di utenza, andando ad espandere il proprio servizio principalmente nelle zone costiere italiane (riviera ligure, riviera romagnola, riviera toscana...). “Esselunga a casa” offre oltre 15 mila items, come frutta, verdura, carne, pesce, salumi, prodotti enogastronomici, la spesa consente inoltre di accumulare punti da cumulare ed ottenere ampio catalogo di premi. Il gruppo ha realizzato nel 2021 vendite per 8 miliardi di euro (+2% rispetto al 2020), con un EBITDA pari a 690 milioni (8% del fatturato), superando sia Coop che Carrefour³⁴.

2.5: IL CONCETTO DI OMNICALITÀ

Un metodo adottato da molte aziende, come descritto anche nei precedenti capitoli, è quello dell’omnicanalità.

Nasce con l’obiettivo di integrare tutti i touch point sia fisici che virtuali in maniera trasparente, offrendo al cliente la possibilità di trovare coerenza sulle informazioni sia online che offline, creando fluidità e eliminando tutti i possibili dubbi e difficoltà d’acquisto.

Le aziende che adottano l’omnicanalità non intendono più gestire diversi canali, ma diversi profili di consumatori con i quali interagiscono tramite vari mezzi. Ad esempio, se un cliente deve effettuare un acquisto complesso può ricercare le informazioni sia in Internet che nel punto vendita fisico e lo stesso vale per un acquisto che può essere fatto nella medesima circostanza.

È importante però specificare che questo metodo di interazione tra azienda e cliente non è sempre raggiungibile, quindi esistono vari tipi di canali:

- Monocanalità: viene utilizzato un solo canale di vendita e di comunicazione come, ad esempio, nel punto vendita fisico;

³⁴ Foodaffairs: La spesa online in Italia, <https://www.foodaffairs.it/2021/07/07/trend-nel-2021-la-spesa-online-in-italia-nel-settore-foodgrocery-superera-i-4-miliardi-di-euro-38-sul-2020-dati-osservatorio-netcomm-politecnico-di-milano-su-ecommerce-food/>. Tutti i dati riportati nel capitolo fanno riferimento a questo sito.

- **Multicanalità:** vengono utilizzati più canali di vendita e di comunicazione che non sono collegati tra loro perché le attività possono essere sviluppate in modi differenti o divise da barriere. Ne sono un esempio le offerte assicurative perché si possono trovare diverse polizze a seconda se l'agenzia lavora online o fisica. I vantaggi sono numerosi: si evita la cannibalizzazione ossia quella condizione in cui le vendite di un canale non danneggiano quello di un altro; è più agevole sostituire i canali con altri più moderni o più facili da utilizzare; si ottiene una complementarità tra canali, quindi si assegna i canali in base al mercato da soddisfare e che risulta dunque diviso in regioni differenti, profili consumatori e prodotti diversi;
- **Crosscanalità:** connessioni tra più canali di vendita e di comunicazione che favoriscono il processo d'acquisto offrendo un'unica traiettoria. Ciò significa che si utilizzano diversi canali fisici e virtuali, ma infine convergono in un unico punto come, ad esempio, quando un acquisto viene effettuato online, ma il prodotto viene consegnato nel negozio fisico. Le caratteristiche che emergono da questo tipo di canale possono essere: dissociare le funzioni (un esempio può essere il pagamento online o al ritiro del prodotto); agevolare transizione nei canali dal momento che nel passaggio tra un canale all'altro c'è il rischio di perdere il cliente; quindi, si cerca di sfruttare metodi lock-in che agevolano il passaggio tra i canali. Alcuni esempi possono essere: un buono promozionale da utilizzare sia online che offline, assortimento più esteso o differenziato per il negozio fisico o online. Infine, un canale non deve arrecare danno ad un altro, quindi ad esempio se coesistono due canali come l'E-commerce ed il negozio tradizionale questi devono creare una sorta di coerenza e cercare di informare ed aiutare il più possibile il cliente.

Si sono descritti vari modi di interpretazione dei canali, ma l'omnicanalità è in realtà molto più complessa. Il suo fine ultimo è infatti la trasparenza nei canali e per questo punta ad allargare il ventaglio di acquisto per l'acquirente che ha un mondo di contenuti disponibili da utilizzare e non esiste più una sequenza tipo (online o offline).

È fondamentale, inoltre, utilizzare la matrice funzioni-canali per chiarire quali funzioni sono svolte nei vari canali. Si tratta di una matrice che suddivide da un lato i canali e sotto canali, e dall'altro le funzioni. È utile per comprendere quanto l'omnicanalità sia difficile da ottenere dato che ciascun canale ha diverse caratteristiche che rendono difficile la

standardizzazione a causa della loro natura. Si cerca di specializzare il canale per renderlo perfetto e utilizzare meno risorse nel suo svolgimento. Non sempre però specializzare è la scelta migliore perché possono avvenire fenomeni come il sottoutilizzo dei canali o una aspettativa di polivalenza del cliente che crede così che tutto sia possibile in ciascun canale quando in realtà non è vero.

È importante quindi che un'azienda adotti diversi metodi di sviluppo per cercare di ottenere l'omnicanalità:

- Modalità di sviluppo interno: il canale viene sviluppato in base alle risorse e competenze interne all'impresa. Questa modalità favorisce: la capacità di assestamento ossia la facilità nell'utilizzo di uno specifico canale quando non è così diverso quelli già utilizzati; un maggior controllo nel gestire un'attività che è interna all'azienda; l'apprendimento perché grazie al canale è possibile apprendere direttamente dai clienti, consentendo di controllare meglio i distributori tradizionali. Esistono però anche dei contro come la mancanza di competenze-chiave per gestire tutti i canali; prolungamento dei tempi di sviluppo dato che bisogna assumere nuovo personale e formarlo ed infine il costo nel coordinare i diversi canali che nel breve termine possono provocare un basso profitto;
- Modalità di sviluppo esterno: si costruisce un nuovo canale acquisendo all'esterno un'azienda già avviata e dunque specializzata in quell'ambito. Il principale beneficio di questa modalità di sviluppo riguarda la rapidità di acquisire nuovi strumenti fin da subito e la conseguente minimizzazione dei rischi. Se l'azienda dovesse essere complementare, si potrebbero creare potenziali economie di scala ed acquisire risorse-chiave, ad esempio, sottraendo ai concorrenti risorse rare nel mercato. Tra i contro rientrano sicuramente un trasferimento di conoscenze ridotto, perché il know-how acquisito non è sempre facilmente replicabile, ma anche problemi di incompatibilità di obiettivi e meccanismi manageriali tra le due aziende;
- In comune le modalità di sviluppo interno ed esterno hanno la detenzione di risorse e competenze. Quindi può emergere una terza modalità: il contratto di collaborazione. A renderlo efficace sono la minor onerosità dell'investimento e il fatto che i contratti di collaborazione permettono di relazionarsi con aziende che sono altamente specializzate in quel dato canale. A sfavore c'è la mancata

possibilità di valutare le performance del partner e, come con la modalità di sviluppo esterno, c'è uno scarso trasferimento di competenze. Alcuni esempi di contratti di collaborazione sono l'esternalizzazione ossia un semplice subappalto come la costruzione di un sito Internet e le alleanze e partenariati. I partenariati avvengono tra imprese che sono concorrenti e dunque possono esserci programmi fedeltà, zone geografiche complementari o la costruzione di sito web online in comune, mentre le alleanze avvengono tra imprese non concorrenti.

Dopo aver delineato il concetto dell'omnicanalità, è necessario descrivere anche la sua modalità organizzativa, che può essere presa come modello per molte aziende. Una strategia omnicanales richiede il supporto di una struttura organizzativa adeguata a tutte le esigenze non solo aziendali, ma anche a quelle che riguardano il cliente. Più le azioni di marketing sono inserite a livello aziendale, più l'impresa ha canali da gestire e più quest'ultima sentirà il bisogno di attuare una riorganizzazione interna globalizzata. Le imprese, dunque, tendono a coordinare le attività attraverso la costruzione di un back-office e di un front-office. Le sfide principali che emergono dall'utilizzo dell'omnicanalità sono: logistica, sistema informativo, gestione del personale e conflitti orizzontali. Nel seguito verranno presentati brevemente le caratteristiche di ciascun aspetto.

Relativamente alla logistica i principali costi che emergono sono quelli di approvvigionamento, di trasporto e di movimentazioni merci. Emerge l'esigenza di avere un equilibrio tra ciascun servizio ed il suo relativo costo, in modo tale da ottenere economie di scala e approvvigionamento. Esempi possono essere: una logistica specifica per il canale internet dove i consumatori vengono gestiti tramite una segmentazione specifica come i profili o le aree geografiche; delegare ai vari punti vendita locali fisici il picking che significa sostanzialmente che l'ordine viene ritirato nel punto vendita o dal punto vendita al cliente. Nasce anche la possibilità di ordinare direttamente dal magazzino del punto vendita in questione, per accelerare i tempi della consegna. Infine, è importante ricordare come esempio anche una centralizzazione locale dove risiede una logistica unica sia per internet che per il punto vendita, così l'azienda ha il controllo totale sulle varie vendite e gestione del magazzino.

La gestione dei flussi di merce è associata ai flussi di informazioni e bisogna dunque saper spostare le informazioni all'intera azienda. Attraverso un sistema informativo efficiente esistono due tipi di dati principali. I primi sono i dati sui prodotti dove la referenza vendita deve essere la stessa, qualsiasi sia il canale. Ad esempio, in Internet al

cliente vengono fornite informazioni ricche e dettagliate sul prodotto, e foto di alta qualità. Uno dei vantaggi dell'omnicanalità è quello di abbassare le scorte di sicurezza e quindi se un canale non ha disponibilità di una specifica referenza, è possibile reperirla immediatamente dalle scorte del magazzino del punto vendita fisico. Il secondo dato è quello riguardante i consumatori: unendo i dati sui profili clienti ricavati sia sul canale online che sul canale fisico si possono comprendere e personalizzare al meglio le interazioni tra l'azienda ed il cliente. Un esempio di questo metodo è la carta fedeltà, grazie alla quale si può tener traccia della customer journey.

Riguardo alla gestione del personale: i dipendenti dell'azienda devono apprendere nuovi strumenti e nuovi modi di lavoro. Il front-office deve acquisire nuove competenze per far fronte a clienti più informati, per restare in linea con le nuove tecnologie e per rendere il servizio più allineato con le esigenze dei clienti. Durante la customer journey esistono dei micro-momenti in cui lo shopper avverte dei piccoli bisogni informativi come il prezzo del prodotto o dove reperirlo. L'azienda deve dunque cercare di soddisfare il consumatore, inserendo in rete o tramite pubblicità fisiche più informazioni possibili. La gestione del personale deve essere preparata ad utilizzare e gestire queste tecnologie per risolvere i micro-momenti. Per affrontarle è necessario instaurare regole che siano comuni nella valutazione delle performance e nella retribuzione degli addetti e stabilire indicatori di performance a 2 livelli: performance globale e performance di ciascun canale.

Si possono prevedere formazioni interne per far evolvere le competenze o cercare di far relazionare gli addetti di più canali per avere, in caso di problemi, una soluzione rapida. Un metodo per incentivare questa metodologia di risoluzione dei problemi è quello di avere dispositivi interni che mettano in stretta relazione gli addetti di ciascun canale.

Nello sviluppo di canali, l'impresa crea una rete dove, ad esempio, può sorgere una collaborazione con un'azienda specializzata esterna. Se queste ultime non perseguono gli stessi obiettivi, possono comparire durante lo svolgimento delle attività dei conflitti con le aziende partner e ciascuna impresa tenderà di ottenere più potere sull'altra. È da segnalare dunque che l'omnicanalità genera due tipi di conflitti: orizzontali e verticali.

I conflitti orizzontali sono definiti come sovrapposizioni tra le funzioni svolte come ad esempio opposizioni culturali tra fisico e digitali come può avvenire tra la vendita di un prodotto online o nel punto vendita fisico. Un esempio concreto che raffigura il conflitto orizzontale è quello tra franchisor e franchisee, nel quale il primo crea l'e-commerce ed il secondo vede l'e-commerce come un concorrente. Possono dunque emergere conflitti

per incompatibilità nella fissazione dei prezzi, margini e volumi e questa situazione generalmente viene definita come una vendita “rubata” al punto vendita.

Esistono però dei metodi per ridurre i conflitti orizzontali si possono utilizzare questo tipo di organizzazioni:

- Delega dell’ordine quando si delega al punto vendita più vicino del trattamento dell’ordine e della consegna, a partire dagli stock del punto vendita;
- Priorità dell’ordine al punto vendita per un paio d’ore;
- La redistribuzione dei vantaggi ossia quando centralizzare il magazzino delle vendite elettroniche è più pratico e conveniente, una parte del margine può essere ridistribuito ai punti vendita per remunerare la loro partecipazione alla transazione o percentuale sulle vendite online.

I conflitti verticali: riguardano l’accesso al punto vendita, la ripartizione dei margini, le remunerazioni dei servizi accessori, il merchandising. I conflitti verticali sono relativi a tutte quelle divergenze che possono nascere ad esempio tra un punto vendita fisico ed il suo e-commerce. L’impresa sviluppa l’e-commerce per trovare sbocchi supplementari, per comprendere i consumatori perché interagisce direttamente con essi; conoscere obiettivi del distributore per aumentare il suo potere su di esso, riducendo i suoi comportamenti opportunistici.

Ai giorni d’oggi, dunque, soprattutto per gli Ipermercati e Supermercati si sta vivendo, una “quarta modernizzazione” definita anche “Retail 4.0”. Oggi ipermercati e supermercati cercano sempre di allineare nel miglior modo possibile la gestione dell’offline e dell’online. Questa modernizzazione è dovuta non solo allo sviluppo continuo di nuove tecnologie, ma anche alla generazione Z, definita anche iGen³⁵, dove i ragazzi nati negli anni compresi tra il 1997 e 2012 utilizzano strumenti, Internet, tecnologie e Social Network con molta naturalezza perchè percepiscono la tecnologia come la via più semplice per comunicare e per cercare intrattenimento. Nel 2015 i ragazzi con una età compresa tra i 12 e i 17 anni utilizzavano già uno smartphone³⁶, influenzando dunque i loro comportamenti come modalità di acquisto, e utilizzo di servizi da remoto. Imparare ad utilizzare in modo contemporaneo e senza distinzioni il canale online ed

³⁵ Il Retail dal primo supermercato all’omnicanalità: La storia, l’evoluzione e le opportunità di crescita: <https://iquii.com/2018/04/18/retail-omnicanalita-sfide-opportunita/>

³⁶ Il Retail dal primo supermercato all’omnicanalità: La storia, l’evoluzione e le opportunità di crescita: <https://iquii.com/2018/04/18/retail-omnicanalita-sfide-opportunita/>

offline può non solo fortificare il posizionamento di mercato delle aziende, ma si possono instaurare relazioni con clienti e raccogliere informazioni utili per gestire le vendite in base alle esigenze e alle domande da parte dei consumatori. La raccolta, dunque, di Big Data Analytics può essere utile per sfruttare e per capire i comportamenti di acquisto dei consumatori e cercare di offrire servizi mirati e sempre più personalizzabili.

CAPITOLO III: METODOLOGIA PER LA RICERCA SOCIALE

3.1: LE RICERCHE DI MERCATO

Una delle principali analisi effettuate nel marketing sono le ricerche di mercato. Quest'ultime vengono svolte dalle aziende o dai ricercatori per cercare di comprendere ed analizzare le esigenze, le abitudini, i comportamenti e le preferenze da parte delle persone su un determinato settore o tematiche specifiche come marche, prodotti o servizi. Philip Kotler, uno dei più grandi pionieri del marketing management qualifica il marketing come: "Funzione aziendale che identifica i bisogni e i desideri attuali insoddisfatti, ne definisce e misura l'entità, determina quali mercati target l'organizzazione può servire al meglio e decide prodotti, servizi e programmi appropriati per servire questi mercati. Pertanto, il marketing funge da collegamento tra i bisogni di una società e il suo modello di risposta industriale"³⁷. Ciò che Kotler intende precisare è come il Marketing possa aiutare a cogliere le esigenze delle persone: effettuare dunque un'analisi di mercato è vantaggioso perché indagando sul comportamento delle persone, è possibile raccogliere dati ed informazioni.

L'AMA (American Marketing Association) fornisce una spiegazione ancora più dettagliata sulla relazione tra le ricerche di mercato ed il marketing: "La ricerca di mercato è la funzione che collega il consumatore, il cliente e il pubblico al marketer attraverso le informazioni, le quali vengono utilizzate per identificare e definire opportunità e problemi di marketing; generare, perfezionare e valutare azioni di marketing, monitorare le prestazioni di marketing e migliorare la comprensione del marketing come processo. La ricerca di mercato specifica le informazioni necessarie per affrontare questi problemi; progetta il metodo di raccolta delle informazioni; gestisce e attua il processo di raccolta dei dati; analizza i risultati; e comunica i risultati e le loro implicazioni"³⁸.

Innanzitutto, una ricerca di mercato, riguarda tutte quelle operazioni volte a gestire e raccogliere dati che verranno poi utilizzate dal management e dal marketing per identificare le varie opportunità che il mercato offre.

Dopo aver definito il problema della ricerca, inizia l'attività di raccolta dei dati, il ricercatore dunque deve comprendere quale tipo di dato può utilizzare per il suo studio e

³⁷ M.J. Kotler. *Marketing management*, Prentice Hall, Englewood Cliffs, N.J., 1984

³⁸ Portland Research Group: <https://www.portlandresearch.com/about/what-is-market-research>

quale metodo di raccolta sia il più adatto per ottenere il maggior numero di informazioni possibili.

Vengono dunque distinti due tipologie di fonti informative che possono essere dirette ed indirette:

- **Dati primari (dirette):** sono dati raccolti ai fini della ricerca, sotto il controllo diretto del ricercatore. Sono dati che non esistono in nessuna forma e sono raccolti per la prima volta, hanno natura campionaria con scopo esplorativo e descrittivo. La condizione necessaria per ricavare dati primari è quella di raccogliere direttamente i dati da un campione, popolazione o ente. I principali mezzi utilizzati per ottenere dati primari sono: interviste, sondaggi, questionari ed esperimenti. I vantaggi nell'utilizzo dei dati primari sono la specificità delle informazioni in base alle esigenze del ricercatore e chi ricerca i dati può decidere di renderli noti pubblicamente, brevettarli o addirittura venderli. Solitamente sono tutti dati aggiornati perché vengono raccolti in tempo reale e non da vecchie fonti. Il ricercatore ha il pieno controllo dei dati raccolti attraverso la ricerca primaria, pianificandone la raccolta e le tecniche di analisi da utilizzare.
- Gli svantaggi dei dati primari sono l'onerosità ed il tempo impiegato per la ricerca e la difficoltà di raccogliere i dati a causa della complessità della ricerca. Il dato primario, si forma su circostanze controllate. Ad esempio, si suppone di avere una variabile input come la somministrazione farmaco e una variabile output come la salute dei pazienti, con l'obiettivo di analizzare l'effetto del farmaco sui pazienti. In statistica i dati vengono generati in maniera controllata, ad esempio prendendo in esame due gruppi simili tra loro in base all'età o sesso. Uno dei due gruppi viene sottoposto al trattamento, quindi, viene assegnato il farmaco, all'altro no. Durante l'esperimento vengono rilevate le variabili output dei due gruppi, ovvero la salute dei pazienti. Le differenze che emergono dal trattamento dei due gruppi consentono di capire l'effetto di una variabile rispetto un'altra. In generale tecniche di progettazione degli esperimenti vengono usate in medicina come nel marketing, dove i dati vengono raccolti prima e dopo il trattamento per riuscire ad intuire quale sia la variabilità tra i due campioni in oggetto.

- **Dati secondari (indirette):** sono dati raccolti da studi precedenti. Vengono effettuati per uno studio a scopo conoscitivo e possono essere utilizzati anche da altri ricercatori. Un esempio su come vengono raccolti i dati secondari è l'indagine panel, nella quale i dati sono raccolti a intervalli regolari sulla stessa unità statistica, come individui, imprese e regioni. I vantaggi che emergono dalla disponibilità di dati osservati in diversi momenti sono la valutazione dei comportamenti dinamici e la previsione più accurata di trend futuri. Gli svantaggi sono spesso legati al costo dei sondaggi e il ritardo tra la raccolta e la pubblicazione dei dati (almeno uno o due anni).
- Principalmente i dati secondari sono di due tipi: interni come, ad esempio, i dati raccolti da un'impresa (documenti contabili, analisi o sondaggi da parte delle risorse umane) esterni come i dati raccolti da organizzazioni esterne. Tra queste organizzazioni rientrano il Sistema Statistico Nazionale (SISTAN), che ruota attorno all'ISTAT, che a sua volta comprende altri enti pubblici, come: i ministeri, le regioni e le Camere di commercio. Altri vantaggi dall'utilizzo dei dati secondari sono la facilità di accesso dei dati che possono essere forniti da diverse piattaforme a cui può accedere il ricercatore e forniti a volte in modo gratuito e il tempo dedicato alla raccolta dei dati che solitamente è molto ridotto rispetto a quello dei dati primari. Gli svantaggi che emergono invece sono la scarsa affidabilità delle risorse che comporta ad una ulteriore verifica dei dati raccolti; i ricercatori potrebbero dover gestire dati che irrilevanti per la propria ricerca prima di trovare finalmente i dati richiesti; infine, i dati sono tipicamente generali e non personalizzabili in modo specifico da soddisfare le esigenze del ricercatore come avviene per i dati primari.

3.2 TECNICHE DI INDAGINE PER LA RILEVAZIONE DEI DATI

Con il termine tecnica di indagine si mira a definire l'insieme delle modalità per reperire le informazioni oggetto di studio sulle unità statistiche interessate alla rilevazione. L'utilizzo di una tecnica di indagine più idonea all'oggetto di studio è fondamentale per riuscire a raccogliere più informazioni possibili e ricavare gli aspetti più significativi che andranno poi analizzati dal ricercatore. Nella figura 3.1 vengono elencati alcune delle tecniche di indagine che possono essere utili per rilevare i dati.

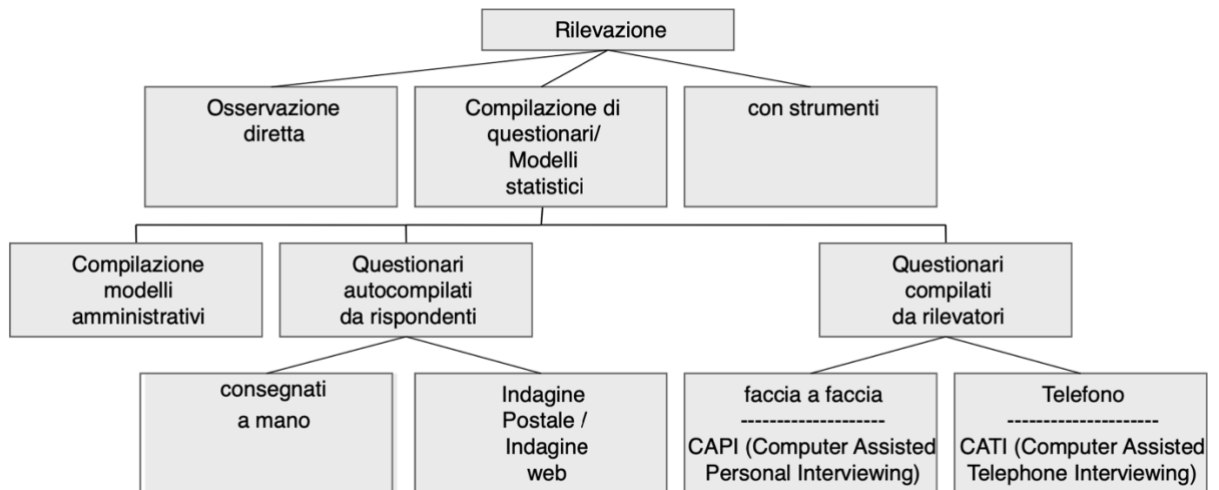


Figura 3.1, Tecniche di rilevazione di dati. Fonte: Dipartimento di Scienze Statistiche, Università di Padova

Le rilevazioni dei dati

Le rilevazioni, come riportato dalla figura 2.2, si dividono in osservazione diretta, compilazione di questionari o modelli statistici ed infine tramite strumenti.

Entrando nel dettaglio, la compilazione tramite questionari si suddivide in altre sottocategorie e sono: compilazione di modelli amministrativi, questionari auto compilati dai rispondenti che possono essere compilati tramite una consegna a mano o indagine postale o web; infine, i questionari compilati dai rilevatori che vengono stilati faccia a faccia (CAPI) o telefonicamente (CATI).

Vengono riportati di seguito alcuni esempi di tecniche di indagine per la raccolta di dati.

Interviste faccia a faccia

Le interviste faccia a faccia vengono condotte da un intervistatore faccia a faccia con l'intervistato. Tramite la lettura di domande l'investigatore registra le risposte così come vengono date dall'intervistato

I vantaggi sono:

- Identificare direttamente chi viene intervistato
- Definizione di quesiti più complessi
- Interviste con durata più lunga
- Risposte più dettagliate e precise

Gli svantaggi sono:

- Tempi più lunghi per la raccolta di dati

- Costi elevati di impianto complessivo per indagini su vasta scala come, ad esempio, la preparazione e la formazione degli investigatori
- Maggiori rischi di condizionamento

Intervista telefonica

L'intervista è conseguita telefonicamente. Le domande vengono poste con un ordine ben preciso e le risposte vengono rilevate esattamente come vengono fornite dall'intervistato.

I vantaggi sono:

- Costo inferiore rispetto l'intervista faccia a faccia
- Reperimento delle risposte più veloce
- Semplicità organizzativa più semplice, non c'è bisogno un'organizzazione più complessa come nell'indagine faccia a faccia

Gli svantaggi sono:

- Mancata copertura telefonica (non tutti hanno un numero fisso, quindi, è difficile reperire il numero di telefono)
- Difficoltà di risposta (alcune categorie di popolazione fanno più fatica a rispondere al telefono; un esempio possono essere gli anziani)
- Meno coinvolgimento nelle interviste
- Il rispondente non è identificato con certezza

Per le modalità di intervista faccia a faccia e telefonica è possibile valutare diversi aspetti favorevoli di questi due sistemi di intervista come la capacità di gestire questionari lunghi in cui l'intervistatore può assistere il rispondente durante l'intervista favorendo l'accuratezza delle misurazioni e la tempestività e rapidità di elaborazione dei dati che in questi tipi di rilevazioni sono maggiori rispetto alle altre che verranno citate prossimamente.

Questionario

Il questionario è uno degli strumenti principali per raccogliere informazioni nel campo delle ricerche di mercato. Le principali tipologie che possono emergere sono: questionario cartaceo e questionario informatizzato (CATI Computer Assisted Telephone

Interview e CAPI Computer Assisted Telephone Interview). A sua volta si possono distinguere diverse modalità per la sua compilazione come ad esempio: autocompilazione o autosommistrazione (questionario postale), compilazione diretta o mediante intervista telefonica o faccia a faccia. Vengono riportate alcuni esempi di questionario:

Questionario postale autocompilato

Il questionario viene indirizzato direttamente al rispondente per posta o corriere. Dopo la sua compilazione del questionario, quest'ultimo viene spedito indietro o consegnato ad un addetto che lo ritira a domicilio.

I vantaggi:

- Bassi costi di realizzazione
- Organizzazione semplice
- Limitati rischi di condizionamento
- Adatto a questionari lunghi perché il rispondente può prendersi il suo tempo per le risposte

Gli svantaggi:

- Tempi più lunghi per la raccolta dei dati
- I rispondenti sono limitati
- Se il tema non è coinvolgente minore sarà la capacità di ottenere le risposte
- Non si possono aiutare le persone a comprendere il tipo di domanda come nelle interviste faccia a faccia e telefono

Diario

Il diario è un tipo di questionario che viene utilizzato per analizzare eventi con una certa frequenza e con una limitata importanza come le spese di basso livello o attività quotidiane effettuate dalle famiglie. Richiede dunque una compilazione accurata ma diluita nel tempo, infatti questo strumento consente di raccogliere informazioni quando l'evento specifico avviene durante la giornata senza dover incorrere ad uno sforzo di memoria come avviene per le interviste faccia a faccia o per via telefonica.

I vantaggi sono:

- Capacità di registrare in modo dettagliato i dati, perché si parla di eventi poco rilevanti ed effettuati con molta frequenza (esempio: visione della televisione durante la giornata, uso del telefono)

Gli svantaggi sono:

- Struttura questionario complicata
- Il soggetto incaricato a rispondere alle domande può scordarsi di compilare il questionario; quindi, non si ha una frequenza omogenea delle risposte
- Necessità di un soggetto che prelevi le risposte

Dopo aver descritto due esempi di modalità di questionario, è importante anche evidenziare quali altre tipologie di somministrazione del questionario possono essere utilizzate. Le più frequenti sono: CAPI, CATI e CASI.

CAPI (Computer Assisted Personal Interview)

Questa tecnica prevede la presenza dell'intervistatore e l'utilizzo di un computer, con un questionario elettronico per lo svolgimento dell'intervista. L'utilizzo del questionario elettronico permette la gestione di interviste più complesse, aiuta a prevenire errori di acquisizione dei dati e rende i dati disponibili in tempi brevi.

I vantaggi sono:

- Sono possibili lunghi colloqui, che possono includere anche ulteriori sondaggi di chiarimento;
- Sono possibili interviste complesse perché le regole di salto sono gestite direttamente dal software;
- Possono essere poste domande più complicate;
- Sono possibili controlli e riconciliazioni con dati già disponibili
- Le interviste possono essere molto fluide e piacevoli perché l'intero questionario può essere formulato da domande personalizzate
- La codifica assistita delle variabili testuali può essere eseguita durante il colloquio
- Le istruzioni e le definizioni lunghe possono essere fornite agli intervistatori attraverso aiuti in linea
- È possibile utilizzare materiale aggiuntivo (mappe, fogli di risposta, ecc.)

Gli svantaggi sono:

- Particolare attenzione deve essere posta nella progettazione e nello sviluppo del questionario elettronico al fine di ridurre al minimo “l'effetto segmentazione”, per fornire all'intervistatore il massimo aiuto e assistenza e per garantire l'efficienza del software procedura
- Le prime fasi di programmazione e test possono essere complesse, lunghe e costose
- Sono possibili meno domande in stile tabella
- Genera pregiudizi di desiderabilità sociale associati a domande delicate³⁹

CATI (Computer Assisted Telephone Interview)

Con questa tecnica l'intervistatore svolge l'intervista telefonicamente, utilizzando un computer con questionario elettronico, lavorando generalmente in un call center dove sono presenti altri intervistatori e almeno un supervisore è a disposizione per dare aiuto, se necessario. Il ritmo dell'intervista è più incalzante perché la conversazione si svolge al telefono. Inoltre, le interviste non sono così costose e il campione può essere diffuso su tutto il territorio.

I vantaggi sono:

- Sono possibili interviste complesse perché le regole di salto sono gestite direttamente dal software
- Sono possibili controlli e riconciliazioni con dati già disponibili (con dati di precedenti rilevazioni o archivi amministrativi)
- Le interviste possono essere molto fluide e piacevoli perché l'intero questionario e la formulazione delle domande possono essere personalizzati
- La codifica assistita delle variabili testuali può essere eseguita durante il colloquio

³⁹ A.A.V.V., *Handbook of Recommended Practices for Questionnaire Development and Testing in the European Statistical System, Questionnaire Design*, 2006, pag.25

- L'attività di intervista può essere monitorata direttamente in quanto l'intervistatore è presente nel call center o attraverso l'analisi quotidiana di una serie di indicatori, che permette di intervenire immediatamente per ottimizzare il questionario o per formare ulteriormente gli intervistatori (riducendo il bias dell'intervistatore)
- Le questioni delicate possono essere gestite più facilmente rispetto al colloquio faccia a faccia

Gli svantaggi sono:

- Non sono consigliabili interviste molto lunghe
- Le domande e gli elenchi delle risposte devono essere brevi
- Le domande non dovrebbero richiedere lunghe riflessioni, poiché interrompe il flusso dell'intervista
- Le prime fasi di programmazione e test possono essere complesse, lunghe e costose
- Sono possibili meno domande in stile tabella
- Non sono consigliabili troppe risposte testuali
- Genera pregiudizi di desiderabilità sociale associati a domande delicate
- Non sono possibili aiuti visivi per gli intervistati⁴⁰

CASI (Computer Assisted Self-Interviewing): WBS (o CAWI), DBM, EMS e TDE

Con queste tecniche il colloquio è autosomministrato con l'utilizzo di un questionario elettronico.

- **Web Based Surveys (WBS):** la raccolta dei dati è gestita online. Il rispondente accede al questionario tramite una connessione web e lo compila
- **Disk by Mail (DBM):** l'intervistato riceve via mail un floppy disk o un CD con il questionario elettronico da compilare. È tenuto a installarlo sul proprio PC, a compilarlo e ad inviarlo indietro

⁴⁰ A.A.V.V., *Handbook of Recommended Practices for Questionnaire Development and Testing in the European Statistical System, Questionnaire Design*, 2006, pag.26

- **E-Mail Survey (EMS):** ha le stesse caratteristiche del DBM, con l'unica differenza che il questionario viene inviato via e-mail
- **Inserimento dati a toni (TDE):** l'intervistato compone un numero collegato a un computer. Risponde alle domande poste dal computer, utilizzando la tastiera del telefono. Questa tecnica può essere utilizzata quando i dati richiesti sono numerici o possono essere direttamente associati a codici numerici.

I vantaggi sono:

- Le regole di salto possono essere automatizzate nel questionario elettronico
- È possibile includere un semplice insieme di regole per evitare errori di intervallo e coerenza del questionario elettronico
- L'intero questionario e la formulazione delle domande possono essere personalizzati
- La codifica assistita di variabili testuali molto semplici può essere eseguita durante il colloquio

Gli svantaggi sono:

- Possono essere gestiti solo questionari semplici e brevi
- Le domande e gli elenchi delle risposte devono essere brevi
- Particolare attenzione deve essere prestata durante la progettazione e lo sviluppo del questionario elettronico per effettuare la navigazione in Internet⁴¹

È importante sottolineare l'importanza dei questionari come strumento di raccolta dei dati, perché non solo è il mezzo che sarà adottato per la ricerca di questa tesi ma costituiscono la base di ogni misurazione statistica basata sull'indagine. Sono di gran lunga gli strumenti di misura più importanti che gli statistici utilizzano per raccogliere più informazioni sui fenomeni oggetto di studio. Le informazioni che vengono fornite dal questionario dipendono dalla presentazione di come vengono posti i vari quesiti, dalla

⁴¹ A.A.V.V., *Handbook of Recommended Practices for Questionnaire Development and Testing in the European Statistical System, Questionare Design*, 2006, Pag.25

preparazione dell'intera struttura delle domande e dagli obiettivi che il ricercatore intende conseguire.

Il questionario innanzitutto è uno strumento di misura. Lo scopo principale è quello di rendere operativa la domanda e di fornire informazioni da parte dell'intervistato in un formato che consentirà poi la raccolta di dati per l'analisi statistica. I concetti per prima cosa devono essere resi chiari in modo tale da consentire al progettista del questionario di effettuare le analisi e gli intervistati devono riuscire a comprendere e rispondere correttamente alle domande. Al fine di fornire una misurazione valida e affidabile, la formulazione, la struttura e il layout devono tenere conto della natura e delle caratteristiche della popolazione rispondente. Quindi devono essere poste agli intervistati domande pertinenti da consentire la raccolta dei dati in modo efficiente e con il minimo errore, facilitando la loro codifica e acquisizione, riducendo al minimo la quantità di modifica e imputazione richiesta.

I passi principali per realizzare un questionario efficiente sono:

- **Revisione della letteratura**

Ogni progetto di questionario dovrebbe iniziare con una revisione della letteratura esistente sull'argomento oggetto della ricerca o investigare su sondaggi con argomenti di ricerca analoghi.

- **Specificazione degli obiettivi dell'indagine**

Prima di iniziare a redigere il questionario, dovrebbero essere specificati gli obiettivi dell'analisi. Si devono prendere decisioni su concetti di base come la popolazione target, sul disegno del campionamento, sulle risorse disponibili ed eventualmente sulla modalità di raccolta dei dati da prediligere. Nel caso di sondaggi su argomenti nuovi, dovrebbe essere condotta un'intensa consultazione incentrata su chi risponderà alle domande al fine di identificare i concetti richiesti.

- **Concettualizzazione**

Una volta stabiliti i concetti di indagine, è necessario tradurli in variabili osservabili. La complessità dei concetti teorici richiede dunque una rigorosa selezione di tratti empirici che possono essere osservati in un'indagine. Alcuni metodi come l'analisi dimensione/indicatore, l'analisi semantica, il campionamento del contenuto, l'interazionismo simbolico e la mappatura concettuale possono facilitare questo passaggio cruciale.

- **Esplorazione dei concetti: focus group e interviste di approfondimento**

Sono disponibili numerosi metodi qualitativi per avere un'idea di come gli intervistati pensano a ciò che i progettisti del sondaggio hanno concepito. Tali metodi includono ad esempio focus group e interviste approfondite.

I focus group sono composti da un piccolo numero di membri della popolazione target guidati da un moderatore. L'obiettivo è imparare come i potenziali intervistati utilizzano i termini relativi all'argomento, come comprendono concetti generali o terminologia specifica e come percepiscono le domande in termini di sensibilità o difficoltà.

Le interviste approfondite o qualitative si concentrano sulla comprensione dei concetti da parte degli intervistati, su come gli intervistati interpretano determinate domande e su come arrivano alle loro risposte. A differenza dei focus group, le interviste in profondità non si basano su una discussione di gruppo.

Definizione di un elenco di variabili e di un piano di tabulazione

Una volta definiti gli obiettivi e concetti, è opportuno definire un piano di tabulazione concreto e un elenco di variabili che specifichino l'output atteso dall'indagine. L'elenco delle variabili e dei valori deve essere visto come una lista contenente ogni definizione di ogni elemento che verrà analizzato. Per quanto riguarda l'elenco delle variabili, si raccomanda di operare una distinzione tra variabili di background come, per esempio, le variabili demografiche o variabili tecniche basate sulle risposte ottenute.

- Decisione sulla modalità di raccolta dei dati

La scelta di una modalità di raccolta dei dati adeguata deve tenere conto del numero dei contenuti e dell'ambito delle variabili di indagine. Inoltre, dovrebbero essere presi in considerazione aspetti come la sensibilità delle domande o possibili problemi di reperimento delle informazioni e le possibili preferenze della popolazione target. È importante notare che il questionario non dovrebbe essere progettato senza una precedente decisione sulla modalità di raccolta dei dati.

- Design di un questionario

Oltre alla formulazione e alla struttura del questionario, gli elementi di progettazione visiva richiedono alcune considerazioni. Gli stimoli non solo verbali, ma anche non verbali guidano gli intervistati e gli intervistatori durante la compilazione del questionario. L'uso dei giusti stimoli non verbali (come simboli

grafici e layout) è fondamentale per ottenere stime valide e affidabili e per migliorare l'usabilità sia per gli intervistati che per gli intervistatori.

Le domande possono essere poste seguendo un ordine ad imbuto o imbuto rovesciato. Quando si parla di disposizione ad imbuto, vengono poste a inizio questionario domande di ambito generale e successivamente si passa a domande più specifiche; questo metodo viene utilizzato per focalizzare l'interesse sul tema principale del questionario e serve per raccogliere e valutare sensazioni e atteggiamenti dell'intervistato. Nella disposizione ad imbuto rovesciato invece, al contrario, vengono poste prima le domande specifiche ed in seguito le domande più generiche; questa disposizione è utile per raccogliere informazioni direttamente su uno specifico tema, e viene utilizzata, a differenza della struttura citata prima per valutare situazioni più oggettive. Esistono diverse tipologie di domanda che il ricercatore deve scegliere per comporre il questionario:

- Domande aperte: le risposte sono realizzate direttamente dall'intervistato, grazie ad esse si possono raccogliere informazioni su un tema a cui non si comprende il fenomeno indagato. Questo genere di domande hanno tempi lunghi di risposta e sono le più costose in termini di raccolta, in base al livello culturale di coloro che rispondono al questionario, si possono ottenere risposte di alta o bassa qualità, infine possono essere difficili da codificare in caso di analisi di dati.
- Domande chiuse: sono le più semplici da utilizzare e generalmente vengono usate su argomenti dove il fenomeno indagato è conosciuto dall'intervistato e si vuole intervistare su larga scala. Hanno un tasso di risposta maggiore e sono più facili da codificare rispetto le domande aperte. Le risposte che si ricevono sono mirate e previste dal ricercatore.
- Domanda semichiusa: in questo caso si possono avere domande prestabilite dal ricercatore con l'aggiunta di una domanda aperta a cui l'intervistato può dare la propria opinione.
- Domanda gerarchizzate: è possibile ordinare le risposte elencate dall'intervistatore in base alle preferenze dell'intervistato.

È possibile, inoltre, gestire il questionario tramite delle domande filtro, in cui si possono saltare quesiti in base alla risposta data dall'intervistato. Vengono utilizzate quando il ricercatore vuole indirizzare i rispondenti su domande di interesse specifico, evitando quindi condizionamenti sulla risposta ed evitare domande dettagliate inutili per l'analisi.

3.3: LE RILEVAZIONI CAMPIONARIE

Come introdotto nei capitoli precedenti, il questionario è un ottimo strumento per la raccolta delle informazioni da parte del ricercatore. La difficoltà di ricercare fonti secondarie che aiutano a soddisfare un obiettivo conoscitivo, incentiva la raccolta diretta dei dati per ricavare le informazioni ed analizzarle al fine di comprendere e prevedere aspetti specifici sul comportamento della popolazione di interesse. Nella maggior parte dei casi però durante una ricerca di mercato, essa non può essere svolta nella sua interezza (intera popolazione), ma il ricercatore di solito deve accontentarsi di un sottoinsieme di questa realtà. Quindi durante le ricerche di mercato è diffuso il ricorso delle rilevazioni campionarie per una serie di obiettivi come identificare il mercato in base ad una determinata tematica come un prodotto o misurare la soddisfazione di una unità.

Solitamente si osserva un campione di un'unità concernente ad una popolazione ristretta piuttosto di effettuare una rilevazione totale, quindi, completa di tutte le unità. Le ragioni che favoriscono questa scelta sono la tempestività, l'economicità e la precisione dei dati. A riguardo si possono distinguere due errori che emergono nella misurazione dei dati: errore campionario, legato all'osservazione della popolazione, ed errore non campionario, legato alle componenti ad esempio per mancata risposta o inadeguatezza delle unità presenti nella popolazione.

Sono coinvolti diversi aspetti dalle rilevazioni campionarie come:

- **Definizione della popolazione:** la popolazione rappresenta l'insieme degli elementi o unità statistiche di interesse per l'indagine o per lo studio;
- **Specificazione dell'unità di campionamento:** un campione è un sottoinsieme delle unità statistiche in una popolazione che dovrebbe rappresentare l'intera popolazione. Dunque, bisogna sottolineare anche che il campionamento è un processo di selezione di unità statistiche di una popolazione che ne rappresentano le caratteristiche

- **Specificazione dell'elenco delle unità campionarie:** l'elenco si determina da una lista di unità che rappresenta gli elementi di una popolazione.
- **Metodo di campionamento (criterio di selezione delle unità campionarie):** il metodo di campionamento utilizzato ha un impatto sull'informazione e sulla qualità ottenuta. La natura del campionamento può essere principalmente di due tipi: campionamento probabilistico e campionamento non probabilistico.

Il campionamento probabilistico

In questo metodo di campionamento, ad ogni unità statistica della popolazione è associata una data probabilità nota e non nulla di essere inclusa nel campione. Essa si può dividere in:

- **Campionamento casuale semplice:** è la tecnica di campionamento di base in cui ogni elemento della popolazione ha una probabilità di selezione nota ed uguale di essere estratto. Ogni elemento è selezionato indipendentemente dagli altri elementi. Dunque, il campionamento casuale semplice, è un'estrazione senza reinserimento o in blocco, che ha due caratteristiche principali: probabilità nota di estrazione e indipendenza tra le estrazioni. Per attuare questo disegno di campionamento, il ricercatore avrà, in primo luogo, bisogno di un elenco completo dei membri della popolazione; a ciascuna delle N unità dell'elenco verrà quindi assegnato un numero e verranno estratti casualmente gli n numeri corrispondenti ai soggetti da includere nel campione. Il campionamento casuale semplice è la base del campionamento della composizione casuale o “random-digit-dialing” (RDD) utilizzato in molti sondaggi telefonici, in particolare negli Stati Uniti⁴². I metodi RDD sono stati sviluppati in modo che i campioni telefonici potessero includere famiglie i cui numeri di telefono non

⁴² PIERGIORGIO CORBETTA, *Social Research, Theory Methods and Techniques*, Sage Publication, London, 2003, pag. 218

sono inclusi negli elenchi telefonici. Fatta eccezione per le interviste telefoniche, il semplice campionamento casuale è usato raramente nella ricerca sociale, sia perché la selezione si basa interamente per caso e non incorpora informazioni già note sulla popolazione, sia perché negli studi su larga scala comporta uno schema di raccolta dati costoso e difficile da organizzare.

- Campionamento sistematico: è una procedura statisticamente equivalente al campionamento casuale semplice dal punto di vista del risultato. L'unica differenza sta nella tecnica di individuazione dei soggetti. Le unità di campionamento non vengono più selezionate casualmente, ma scorrendo l'elenco dei soggetti e selezionando sistematicamente un'unità per ogni dato intervallo. Se è nota la dimensione N della popolazione di riferimento ed è stata stabilita la dimensione n del campione, si seleziona un'unità ogni $k = N/n$ unità della popolazione, partendo da un numero scelto casualmente compreso tra 1 e k (k è chiamato intervallo di campionamento)⁴³. Il campionamento sistematico è equivalente al campionamento casuale tranne in quelle situazioni (abbastanza rare) in cui l'intervallo di campionamento coincide con qualche caratteristica periodica della popolazione di riferimento. L'utilità del campionamento sistematico non risiede tanto nella sua semplificazione della procedura di selezione quanto nella sua capacità di produrre campioni casuali anche quando non è disponibile un elenco della popolazione e la dimensione della popolazione è sconosciuta. Nella ricerca sociale, il campionamento sistematico è spesso utilizzato proprio perché in molti casi non è disponibile un elenco della popolazione di riferimento.
- Campionamento stratificato: questo disegno di campionamento è organizzato in tre fasi. Innanzitutto, la popolazione di riferimento è suddivisa in sottopopolazioni (dette strati) il più possibile

⁴³ PIERGIORGIO CORBETTA, *Social Research, Theory Methods and Technique*, Sage Publication, London, 2003, pag. 218

omogenee rispetto al fenomeno da studiare. Ciò avviene utilizzando come criterio di stratificazione una variabile correlata al fenomeno. In secondo luogo, viene selezionato un campione da ogni strato mediante una procedura casuale. Infine, i campioni prelevati da ogni strato vengono raggruppati in modo da produrre un campione complessivo. L'obiettivo è di rendere il campione rappresentativo rispetto a questi gruppi, che altrimenti, utilizzando un campionamento casuale semplice, sarebbero sottorappresentati. Questo tipo di campionamento assicura che i sottogruppi (strati) di una data popolazione siano adeguatamente rappresentati all'interno dell'intera popolazione campione di uno studio di ricerca. È, inoltre, una tecnica efficace per studiare una tendenza o un problema che potrebbe differire tra i sottogruppi.

- Campionamento a cluster: Questa procedura viene utilizzata quando la popolazione può essere suddivisa in gruppi. È il caso delle famiglie, delle classi scolastiche, dei reparti di lavoro, dei reparti ospedalieri e così via. Tali raggruppamenti sono chiamati cluster. In questa procedura non vengono selezionate le unità (gli individui), ma i cluster: tutte le unità appartenenti ai cluster selezionati sono quindi incluse nel campione. In questo modo si sfrutta la vicinanza spaziale tra individui appartenenti allo stesso cluster. La procedura è estremamente utile quando, come in questo esempio, non è disponibile un elenco delle unità, ma è possibile selezionare i cluster mediante una procedura casuale.

Il campionamento non probabilistico

Quando un campione probabilistico non è fattibile, o quando è noto in anticipo che non può essere implementato nella fase di raccolta dei dati, verrà adottato fin dall'inizio un disegno del campione non probabilistico.

I tipi più comuni sono:

- Campionamento per quote: questo è probabilmente il modello di campionamento più utilizzato, soprattutto nelle ricerche di mercato e nei sondaggi di opinione. Per attuare la procedura, la popolazione

deve essere prima suddivisa in un certo numero di strati definiti da alcune variabili di cui si conosce la distribuzione (es. genere, fascia di età e livello di istruzione). Successivamente si calcola il 'peso' proporzionale di ogni strato, cioè la porzione di popolazione complessiva che appartiene a ciascun gruppo. Infine, si moltiplicano questi pesi per la dimensione campionaria n . Fino a questo punto, il campionamento per quote non è diverso dal campionamento stratificato. La differenza, tuttavia, sta nel fatto che, entro i limiti posti alla dimensione complessiva di ciascuna quota, l'intervistatore è libero di scegliere i soggetti dell'intervista a sua discrezione. Pertanto, il campionamento per quote è un campionamento stratificato con la scelta delle unità lasciata all'intervistatore, per cui la dimensione della quota limita questa libertà di scelta e garantisce che il campione complessivo riproduca la distribuzione della popolazione rispetto alle variabili su cui le quote sono basate. I limiti del campionamento delle quote sono evidenti. La libertà di scelta dell'intervistatore fa sì che, entro i vincoli prefissati, possa adottare criteri utilitaristici di selezione privilegiando i casi più facilmente accessibili (conoscenti, soggetti conviventi nella stessa zona o appartenenti allo stesso gruppo sociale), evitando situazioni che possono comportare qualsiasi grado di difficoltà e non persistendo quando i soggetti sono riluttanti a essere intervistati.

- Campionamento a scelta ragionata: in questa procedura, le unità di campionamento non vengono scelte in modo casuale, ma sulla base di alcune loro caratteristiche. Questo tipo di campionamento viene utilizzato, ad esempio, su campioni molto piccoli per evitare fluttuazioni casuali che potrebbero allontanare eccessivamente il campione dalle caratteristiche della popolazione. Il campionamento a scelta ragionata non offre ovviamente le garanzie fornite dal campionamento probabilistico, né la semplicità di attuazione vista nel campionamento per quote. Tuttavia, come si è detto, trova una comoda applicazione nel caso

di campioni molto piccoli (soprattutto quando le unità di analisi non sono individui, ma istituzioni, comuni, ecc.).

- Campionamento a valanga: questo modello di esempio è particolarmente utile nello studio di quei gruppi sociali in cui i membri tendono a nascondere la propria identità per ragioni morali, legali, ideologiche o politiche. La procedura viene utilizzata anche per studiare gli "elementi rari": piccoli gruppi sparsi su una vasta area ma che in qualche modo si tengono in contatto tra loro (membri di religioni minoritarie, gruppi e associazioni particolari, ecc.). Il campionamento a valanga implica l'identificazione di soggetti da includere nel campione tramite rinvio da altri soggetti. Il percorso inizia con un ristretto numero di soggetti che possiedono i requisiti desiderati, attraverso i quali vengono individuati ulteriori soggetti con le stesse caratteristiche. Man mano che il processo va avanti, il numero di soggetti dovrebbe aumentare in modo esponenziale, da qui l'analogia con il termine "valanga". Questo campionamento ha lo svantaggio di selezionare gli individui socialmente più attivi e più visibili (sebbene all'interno del quadro ristretto del gruppo). Inoltre, vi è il rischio che questa catena di selezione possa essere incanalata lungo percorsi troppo specifici. Per evitare tali rischi è generalmente necessario come nel caso del campionamento per quote imporre vincoli basati su quanto già noto del fenomeno in esame.

Errore Campionario

Durante l'ottenimento delle informazioni da un questionario, possono emergere degli errori sulle risposte. I principali tipi di errori sono:

- Errori motivazionali: chi è tenuto a rispondere al questionario non è abbastanza motivato nell'immedesimarsi nel ruolo di "informatore"; dunque, tende a rilasciare risposte non precise;
- Errori sociali: l'intervistato cerca di fornire risposte non esatte fornendo un'immagine di sé non veritiera;

- Errore di memoria: le informazioni richieste sono difficili da reperire perché l'intervistato non ha ricordo su quel tipo di questione;
- Errori di comunicazione: è possibile che durante la lettura delle domande da parte dell'intervistato non vengano comprese pienamente la tematica o l'oggetto dell'indagine
- Errori strutturali: l'intervista è troppo lunga o breve, quindi l'intervistato non rispetta l'ordine delle domande o risponde in modo non adeguato.



Figura 3.2, Errori nell'indagine Campionaria. Fonte: Teoria e tecnica dell'indagine statistica e del campionamento, Dipartimento di Scienze Statistiche, Università di Padova

Come riportato nella Figura 3.2, l'errore globale si divide in errore campionario ed errore non campionario. L'errore campionario è inversamente proporzionale alla dimensione del campione, controllabile a priori dal ricercatore attraverso la teoria della probabilità. L'errore non campionario. Diversamente ed è direttamente proporzionale alla dimensione campionaria, l'errore è positivamente influenzato dalla complessità dell'indagine. Un esempio di errore non campionario in cui generalmente ci si può imbattere è la distorsione da selezione definito anche selection bias. Si può indurre a selection bias se: vengono selezionate unità che per il ricercatore sono "rappresentative", viene specificata in modo non corretto la popolazione obiettivo, vengono sostituite le unità del campione per

raggiungere la numerosità obiettivo desiderata ed infine si possono incontrare mancate risposte del questionario.

Verifica del questionario

Infine, per riuscire a comporre un questionario è importante effettuare una serie di controlli per comprendere se il questionario è stato strutturato in maniera efficiente per l'oggetto di ricerca da parte dell'investigatore.

I controlli sono:

- **Revisione estesa:** viene effettuato un controllo da parte di un gruppo di esperti specializzati sul tema oggetto della ricerca. L'obiettivo dei gruppi di esperti è valutare il questionario per potenziali problemi sia per gli intervistatori che per gli intervistati. Di solito queste revisioni sono condotte all'inizio del processo di sviluppo del questionario. Esistono due modi per condurre gruppi di esperti: il primo è mediante una discussione strutturata su ciascuna domanda, con uno schema di codifica standardizzato; il secondo è sotto forma di una discussione non strutturata formalmente su ciascuna domanda, ma senza una lista di codifica standardizzata. Lo scopo quindi del gruppo di esperti è quello di migliorare la stesura del questionario tramite suggerimenti.
- **Pre-test:** il questionario viene sottoposto ad una verifica sul campo di solito composta tra le 200 e 300 unità, somministrandolo ad individui che non sono omogenei tra loro riguardo le caratteristiche strutturali. È importante che all'interno del campione analizzato ci siano degli intervistati che rilasciano commenti o suggerimenti sul questionario, per migliorarlo sulla completezza e correttezza sulle informazioni che devono essere rilevati, il grado che il questionario è accolto da parte degli intervistati e la semplicità di gestione del questionario.
- **Test di alternative:** si testano altre alternative del questionario differenziando il tipo di domanda o la sequenza delle domande. I risultati infine possono essere analizzati per comprendere quale delle due sia la scelta migliore.

- Indagine pilota: a differenza del controllo pre-test l'indagine pilota ha lo scopo di utilizzare gli elementi costitutivi dell'indagine dunque anche il questionario. Testando in maniera molto ristretta: l'insieme dei veicoli di somministrazione del questionario, le operazioni di post-codifica del questionario ed il piano di tabulazione, viene fornito un test in "miniatura" di tutta l'indagine adottata però su un numero ristretto del campione oggetto di analisi.

3.4: LE SCALE DI MISURA

La ricerca di mercato richiede delle misurazioni dell'oggetto di studio. Per misurazione si intende l'assegnazione di valori numerici e proprietà ad attributi, caratteristiche delle unità ed oggetti, secondo principi ben definiti. Il ricercatore deve scegliere dunque uno strumento per tradurre la dimensione latente in unità, quindi, deve definire una scala di misura. È importante perché la scelta della scala di misura influisce sul tipo di modello statistico che il ricercatore può andare ad utilizzare per analizzare i dati. Una scala di misurazione deve soddisfare: una corrispondenza biunivoca tra i numeri e le caratteristiche da misurare (uno a uno, per esempio, classi di reddito che non devono essere sovrapposte); le regole per l'assegnazione dei numeri dovrebbero essere standardizzate e applicate in modo uniforme; infine le regole non devono cambiare tra gli oggetti o nel tempo.

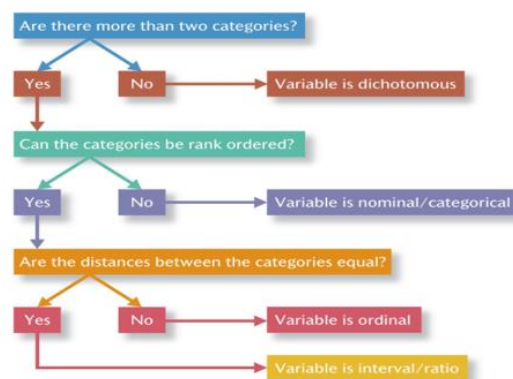


Figura 3.3, I diversi tipi di scale di misurazione. Fonte: Metodi quantitativi per la segmentazione e posizionamento, Università Ca' Foscari di Venezia

Come viene evidenziato dalla Figura 3.3, esiste una variabile di due categorie come, ad esempio, occupato e non occupato definita variabile dicotomica. Quando sussistono più di due categorie, ma che non hanno un ordinamento naturale o un grado di intensità si parla di: scala nominale o categorica. Quando invece le modalità che hanno un ordine naturale ma non è possibile misurare la distanza (esempio basso/medio/alto), si parla di variabile ordinale (scuola media, scuola superiore). Se le modalità hanno ordine naturale e possiamo esprimere la distanza tra le modalità (esempio la variabile quando la variabile è quantitativa) allora la variabile è definita a scala o rapporti.

Le scale di misura sono strumenti per individuare le modalità con le quali una variabile determinata da una domanda può manifestarsi. Una differente classificazione può essere tra: misurazione qualitativa che comprende scala nominale e ordinale; misurazione quantitativa, che comprende scala intervallo e scala rapporti. Nello specifico:

- Scala nominale: descrive una variabile con categorie che non hanno un ordine naturale o una classificazione. Si consente solo il confronto tra due elementi uguali (appartenenti alla stessa classe) o diversi (appartenenti a classi diverse). È possibile codificare le variabili nominali con i numeri, ma l'ordine è arbitrario e qualsiasi calcolo, come il calcolo di una media, una mediana o una deviazione standard, non avrebbe senso. Esempi di variabili nominali sono: codice postale, preferenze politiche (repubblicano, democratico, indipendente), stato occupazionale: (occupato, disoccupato), genere letterario: (commedia, drammatico, satira, epico, tragedia). Le unità, dunque, sono classificate in base ad una classe o modalità.
- Scala ordinale: vale il principio di ordinamento ma non la differenza tra i valori. (non si può capire la distanza, ma si possono calcolare moda e mediana). Le classi possono essere classificate in base a un criterio che consente di determinare quale classe è maggiore e quale è minore anche se non è possibile misurare la distanza o non è possibile valutare l'entità della differenza degli oggetti. Esempi di variabili ordinali sono stato socioeconomico (reddito basso, reddito medio, reddito alto), livello di istruzione (scuola superiore, laurea, laurea magistrale, dottorato di ricerca)

livello di reddito (meno di 50K, 50K-100K, oltre 100K) grado di soddisfazione (totalmente insoddisfatto, abbastanza insoddisfatto, molto soddisfatto, totalmente soddisfatto).

- Scala intervallo: viene quantificato il valore delle caratteristiche di una unità. Si possono avere proprietà nell'ordine della distanza ma la posizione dello 0 è arbitraria o non esiste quindi non c'è nessun punto di riferimento univoco. Ad esempio, se uno studente ricevesse un punteggio QI di 84, allora sarebbe 16 punti al di sotto della media che indicherebbe un potenziale sotto la media per le attività educative, non esiste punto zero per il QI. Si possono calcolare indici come moda e mediana, ma anche un minimo massimo e variazione standard.
- Scala rapporti: ha tutte le proprietà di una variabile intervallo ed inoltre ha anche una chiara definizione di 0. Ad esempio: soldi sul conto corrente, numero di veicoli posseduti negli ultimi 10 anni, numero di persone in un nucleo familiare, tempo di reazione in un'attività del computer, anni di esperienza lavorative, velocità in miglia orarie.

Tecniche di scaling

Le tecniche di scaling sono una serie di procedure che vengono utilizzate per misurare attributi come comportamenti, preferenze ed opinioni. Esistono due modalità di misura per valutare gli attributi secondo una regola specifica:

- Scala di valutazione comparativa: all'intervistato viene chiesto di confrontare una marca o un prodotto o opzioni con qualcos'altro. La misurazione si basa sul confronto tra oggetti esempio confrontare due o più prodotti, con uno standard in comune;
- Scala di valutazione non comparativa: gli intervistati devono solo valutare un singolo prodotto o marchio o opzioni, la misurazione si basa sulla valutazione individuale di ciascuno degli oggetti, valutare un prodotto in maniera oggettiva senza fare riferimento ad un oggetto concorrente.

Tecnica di valutazione comparativa

Una scala comparativa viene impiegata quando si desidera un comune riferimento per tutti i rispondenti, per esempio, una o più marche.

Le tecniche principali di valutazione comparativa sono:

- **Ridimensionamento accoppiato/appaiati:** definita anche con il nome di “Paired comparison”, in essa agli intervistati viene chiesto di confrontare due marchi o prodotti o opzioni. Questa tecnica è generalmente limitata ad un attributo e un limitato numero di marche, dal momento che per n marche è richiesto un $n(n-1)/2$ confronti accoppiati. Un esempio: presentare all'intervistato 4 paia di marche di shampoo. Per ogni paio, chiedi di indicare quale delle due marche di shampoo preferirebbe per uso personale, si può effettuare una misurazione su scala nominale quindi indicare quale si preferisce indicando per ogni coppia di marca di shampoo quale sia la migliore o peggiore; o effettuare una misurazione su scala ordinale quindi si conta il numero di volte che la marca di shampoo viene preferita;
- **Ridimensionamento della somma costante:** agli intervistati viene richiesto di dividere un numero specifico di punti o percentuali come parte di una somma totale (dividere un ammontare specifico di punti tipo 0 a 100, in modo di riflettere la considerazione che si ha su uno specifico oggetto, esempio 8 attributi su uno shampoo, 0 se ha un particolare aspetto o 100 se ha doppia importanza. Ad esempio: vengono presentati otto attributi dei saponi da bagno. L'intervistato deve assegnare 100 punti tra gli attributi in modo che la tua assegnazione rifletta l'importanza relativa che attribuisce a ciascun attributo. Più punti riceve un attributo, più importante è l'attributo. Se un attributo non è affatto importante, assegnagli zero punti. Se un attributo è due volte più importante di un altro attributo, dovrebbe ricevere il doppio dei punti, un esempio di scala di attributi che si possono utilizzare per il confronto tra le marche di shampoo sono: mitezza; schiuma; ritiro; prezzo; fragranza; imballaggio; idratante; pulizia generale;

- **Misure ordinali di valutazione (Rank order rating scales):**
 questa tecnica di misurazione fornisce all'intervistato una serie di elementi da mettere in ordine in base ad un criterio prestabilito. L' "ordine" potrebbe essere qualcosa come preferenza, simpatia, importanza o efficacia. Questa può essere una semplice struttura ordinale come A è maggiore di B o essere eseguita per posizione relativa, ad esempio, si può dare a ciascuna lettera un valore numerico poiché in A è 10 e B è 7. Lo svantaggio di questa tecnica di valutazione comparativa è quelli di produrre soltanto dati ordinali e consente un limitato numero di analisi;

Tecniche di valutazione non comparative

Le tecniche di ridimensionamento non comparativo sono un insieme di metodi di ridimensionamento in cui ogni elemento viene ridimensionato indipendentemente da ciascuno degli altri. Al rispondente, dunque, viene richiesto di considerare un prodotto e di valutarlo ma a differenza delle tecniche di valutazione comparative non c'è uno standard prefissato.

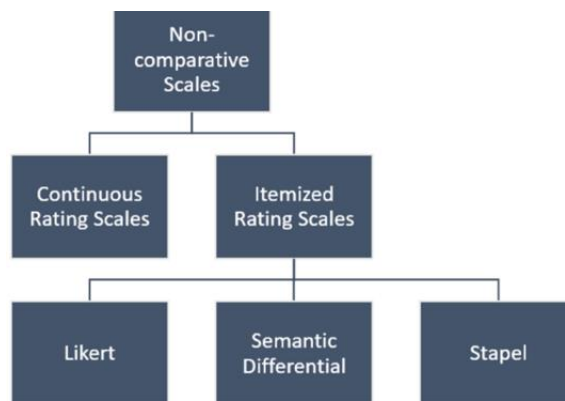


Figura 3.4 Tecnica di valutazione non comparative. Fonte: Survey-measurement and scaling⁴⁴

Come riportato nella figura 3.4, la tecnica di valutazione non comparativa si sviluppa principalmente in due parti:

⁴⁴ Question Star, Survey: Measurement and Scaling: <https://www.questionstar.com/textbook-principles-of-survey-research/survey-measurement-and-scaling/non-comparative-scales/>

- **Scala di valutazione continua (Continuous rating scales):** dove gli intervistati valutano gli oggetti mettendo un segno nella posizione appropriata su una linea che va da un estremo all'altro della variabile di criterio. Un esempio è riportato nella figura 3.5.

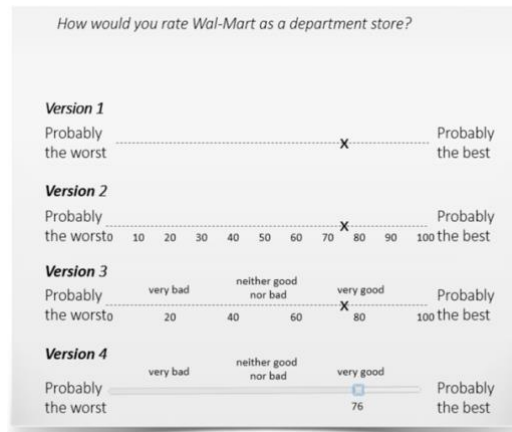


Figura 3.5 Esempio di scala di valutazione continua non comparativa. Fonte: Survey-measurement and scaling⁴⁵

- **Scala di valutazione dettagliata (Itemized rating scale):** è una scala ordinale che ha una breve descrizione o numeri associati a ciascuna categoria, ordinati in termini di posizioni della scala. Agli intervistati viene chiesto di selezionare la categoria che meglio descrive l'oggetto a cui si vuole dare una valutazione. Questa scala di valutazione si suddivide in: scala Likert, scala del differenziale semantico o semantic differential e stapel scape.

Scala Likert

Gli intervistati specificano il loro livello di accordo o disaccordo su una scala simmetrica accordo-disaccordo per una dichiarazione; dunque, il rispondente deve dare una risposta su una particolare affermazione. Questa tecnica prende il nome dallo psicometrico Rensis Likert, che per primo la propose all'inizio degli anni '30. Il titolo «Scala Likert» copre un'ampia varietà di scale, note anche come scale di valutazione sommate. Questa tecnica è ancora la procedura più utilizzata nella valutazione dell'atteggiamento, ed è stata

⁴⁵ Question Star, Survey: Measurement and Scaling: <https://www.questionstar.com/textbook-principles-of-survey-research/survey-measurement-and-scaling/non-comparative-scales/>

svilupata e arricchita notevolmente sin dalla sua introduzione. La procedura su cui si basano le scale Likert è semplice ed intuitiva. Il formato delle singole domande nelle scale Likert è tradizionalmente quello di una serie di affermazioni, dove il rispondente deve dire se, e in che misura, è d'accordo o disaccordo con ciascuna. Nella versione iniziale, lo stesso Likert proponeva sette alternative (che vanno da "concordo fortemente" a "fortemente in disaccordo"). Il numero è stato successivamente ridotto a cinque che vanno da "fortemente in disaccordo" a "fortemente d'accordo", in cui al rispondente viene chiesto di indicare il grado di accordo o disaccordo con ciascuna delle affermazioni relative all'oggetto stimolo in analisi, come riportato nella figura 3.6.

Semantic differential scale:

Strong	1	2	3	4	5	Weak
Decisive	1	2	3	4	5	Indecisive

Figura 3.6 Esempio scala Linkert; Fonte: Metodi quantitativi per la segmentazione e posizionamento, Università Ca' Foscari di Venezia

Semantic differential

All'intervistato viene chiesto di scegliere dove si trova la sua posizione su una scala tra due aggettivi polari. La tecnica differenziale semantica è stata sviluppata negli anni '50 dallo psicologo americano Charles Osgood e collaboratori. Lo scopo di questi ricercatori è stato quello di creare uno strumento altamente standardizzato per misurare il significato che i concetti hanno per le persone. Il problema di cosa "significa" per una persona un termine, una situazione sociale o un oggetto in generale è estremamente complesso e carico di significato sia filosofico che psicologico, rilevante se si vuole indagare sull'importanza del comportamento. Inoltre, Osgood sosteneva che le coppie di attributi bipolari non dovrebbero avere alcuna relazione con l'oggetto in esame e che le stesse coppie di aggettivi dovrebbero essere utilizzate in ogni caso⁴⁶. Anche aggettivi che apparentemente non hanno nulla a che fare con l'oggetto in esame, ad esempio, le coppie maschile-femminile o giovane-vecchio, possono rivelarsi utili per rivelare in profondità,

⁴⁶ PIERGIORGIO CORBETTA, *Social Research, Theory Methods and Techniques*, Sage Publication, London, 2003, pag. 180

e persino subconscio, significati che l'oggetto ha per l'intervistato. Un esempio di domanda di scala del differenziale semantico è riportato nella figura 3.7.

Likert Scale:

I feel my manager is strong: <input type="checkbox"/> Strongly disagree <input type="checkbox"/> Disagree <input type="checkbox"/> Neither agree nor disagree <input type="checkbox"/> Agree <input type="checkbox"/> Strongly agree
I feel my manager is decisive: <input type="checkbox"/> Strongly disagree <input type="checkbox"/> Disagree <input type="checkbox"/> Neither agree nor disagree <input type="checkbox"/> Agree <input type="checkbox"/> Strongly agree

Figura 3.7, Esempio di Semantic Differential; Fonte: Metodi quantitativi per la segmentazione e posizionamento, Università Ca' Foscari di Venezia

Stapel scale

Una scala di valutazione unipolare con 10 categorie numerate da -5 a +5 senza punto neutro (zero). Questa scala è solitamente una scala verticale in cui il singolo aggettivo è posto al centro dell'intervallo pari (da -5 a +5). Al rispondente viene chiesto di identificare in che modo ogni termine descrive correttamente o scorrettamente l'oggetto stimolo scegliendo una categoria di risposta appropriata. Poiché questa è una scala non comparativa e categoriale, la rende molto simile a una scala differenziale semantica con l'unica differenza che è la presenza di un solo aggettivo tra una categoria positiva e una negativa piuttosto che avere due aggettivi opposti. Gli intervistati devono indicare la rilevanza di ciascuna caratteristica per un oggetto nella domanda. Un esempio è riportato nella figura 3.8.



Figura 3.8, Esempio Stapel scale. Fonte: Survey-measurement and scaling⁴⁷

⁴⁷ Question Star, Survey: Measurement and Scaling: <https://www.questionstar.com/textbook-principles-of-survey-research/survey-measurement-and-scaling/non-comparative-scales/>

La tabella 3.9 riassume le principali tecniche di valutazione non comparative e ne descrive brevemente vantaggi e svantaggi:

Scala	Caratteristiche	Esempi	Vantaggi	Svantaggi
Continuous Rating scale	Il rispondente mette un segno su una linea continua	Reazione agli spot televisivi	Facile da costruire	Il punteggio può essere ingombrante, a meno che non sia computerizzato
Linkert scale	Gradi di accordi su una scala da 1 (fortemente in disaccordo) a 5 (fortemente d'accordo).	Misurazione degli atteggiamenti	Facile da costruire, amministrare e capire	Più dispendioso in termini di tempo
Semantic Differential	Scala a sette punti con etichette bipolari	Immagini del marchio, del prodotto e dell'azienda	Versatile	Polemica sul fatto che i dati siano a intervalli
Stapel Scale	Scala unipolare a dieci punti, da -5 a +5, senza punto neutro (zero)	Misurazione di atteggiamenti e immagini	Facile da costruire, amministrare per telefono	Confuso e difficile da applicare

Tabella 3.9, Vantaggi e svantaggi tecniche di valutazione comparative. Fonte: Elaborazione personale

3.5: INTRODUZIONE AI METODI STATISTICI

Dopo aver introdotto i concetti sullo sviluppo di un questionario, è importante approfondire quali siano i principali metodi per gestire i dati e ricavare informazioni importanti da essi. Verranno dunque introdotti i metodi che saranno utilizzati successivamente nel capitolo 4. Per approfondire il caso studio di questo elaborato i principali metodi scelti per l'analisi dei dati raccolti sono la Cluster Analysis e la Regressione. Ciascuno di essi presenta delle peculiarità utili per trarre informazioni rispetto agli obiettivi di ricerca selezionati per l'elaborato.

3.5.1: CLUSTER ANALYSIS

L'analisi dei cluster è un'analisi esplorativa che cerca di identificare dei cluster o gruppi all'interno dei dati, con specifiche caratteristiche di seguito presentate.

Definita anche come analisi di segmentazione o analisi della tassonomia, la cluster analysis cerca di individuare gruppi sulla base della loro omogeneità e somiglianza. In questo tipo di analisi non vengono fatte alcune distinzioni tra variabili dipendenti e indipendenti ed inoltre tramite essa è possibile gestire dati binari, nominali, ordinali e di scala (intervallo o rapporto).

Il ricercatore deve essere in grado di interpretare la cluster analysis e comprendere se i dati raccolti siano significativi per il suo studio. Nell'applicazione aziendale e nel contesto decisionale, l'analisi dei cluster può essere un processo chiave per conoscere gli attributi distinguibili di una vasta popolazione. Ad esempio, i metodi di analisi dei cluster aiutano a segregare la popolazione in diversi bucket o gruppi di marketing in base all'obiettivo della campagna, che possono essere altamente efficaci per iniziative di marketing mirate. Ciò può far risparmiare tempo, fatica e denaro concentrandosi su iniziative separate per ciascun gruppo di pubblico o al contrario decidere di concentrarsi su uno solo.

Esistono diversi approcci che possono essere utilizzati per sviluppare un'analisi dei cluster. Questi metodi possono essere classificati principalmente come segue:

Metodi gerarchici che si dividono in:

- 1) Metodi agglomerativi, in cui le unità statistiche sono raggruppate ciascuna in un cluster separato diverso e separato dagli altri. Inizialmente, ciascuna delle n osservazioni costituisce un cluster separato, quindi saranno formati n cluster. I due cluster più simili secondo la regola della distanza che verrà successivamente spiegata nel capitolo (metodo legame singolo, metodo di collegamento completo...) verranno aggregati. Di seguito verrà ricalcolata la distanza tra i cluster e si ripeterà il passaggio citato pocanzi; quindi, i cluster più simili verranno combinati fino a raggiungere un singolo cluster.
- 2) Metodi divisivi, in cui tutte le unità iniziano nello stesso cluster e la strategia viene applicata al contrario del metodo agglomerativo. Le unità vengono dunque divise fino a raggiungere il punto dove ognuna di esse si trova in un cluster separato.

Metodi non gerarchici, noti spesso come metodi di clustering k-means, dove si cerca di dividere le osservazioni in un pre-specificato numero di cluster

I metodi citati pocanzi, si basano sul concetto di similarità e dissimilarità tra le osservazioni e dunque, richiedono la definizione di una misura distanza.

Le misure di distanza tra le osservazioni

Le variabili utilizzate nell'analisi di un fenomeno oggetto di studi possono portare ad avere dati di natura ad intervallo, ordinali o categoriali. Misurare la distanza tra le osservazioni dipenderà quindi dal tipo di dati di cui dispone.

Per poter identificare correttamente gruppi di unità omogenee occorre infatti considerare quanto le unità statistiche siano simili fra di loro e per poter individuare questa similarità tra le unità bisogna identificare una misura di distanza.

Sia $X_i = (X_{i1}, X_{i2}, X_{ip})$ un vettore di osservazioni della variabile i -sima unità statistica sulle p variabili oggetti di studio. La distanza $d_{ij} = d(x_i, x_j)$ è una funzione che associa a ciascuna coppia x_i ed x_j , un qualsiasi valore numerico non negativo che quantifica la differenza tra le due unità. Una misura di distanza deve soddisfare le seguenti proprietà:

- Non negatività: $d_{ij} \geq 0$
- Identicità: $d_{ii} = 0$
- Simmetria: $d_{ij} = d_{ji}$
- Disuguaglianza triangolare: $d_{ij} \leq d_{ij} + d_{jk}$

Le quattro proprietà citate non sono tra loro indipendenti: si può dimostrare infatti che se valgono le proprietà, ad esempio, di identità e di disuguaglianza triangolare, sono valide anche le proprietà di non negatività e di simmetria.

Nel caso di variabili quantitative, possono essere utilizzati vari tipi di misure di distanza come ad esempio:

- **La distanza euclidea** che risulta essere la misura più utilizzata per calcolare la dissimilarità tra due punti in un piano cartesiano qualora si sia in presenza di sole variabili quantitative. La sua funzione è riportata di seguito, dove d_{ij} è la distanza tra l' i -esima e la j -esima unità statistica.

$$d_{ij} = \sqrt{\sum_{h=1}^p (x_{ih} - x_{jh})^2}$$

- **La distanza di Manhattan**, il cui nome è suggerito dalla struttura rettangolare di molte città statunitensi, in particolare New York. La sua funzione è riportata di seguito:

$$d_{ij} = \sum_{h=1}^p |x_{ih} - x_{jh}|$$

Essa è calcolata tramite la somma del valore assoluto delle differenze delle coordinate dei punti. La distanza di Manhattan viene utilizzata quando le variabili sono: binarie, ordinali o miste, è importante che la dispersione delle variabili deve essere omogenea dividendo i dati per ciascuna di esse per il range.

- **Distanza in base alla correlazione**, con la quale è possibile calcolare la distanza tramite il coefficiente di correlazione r . La sua funzione è riportata in seguito:

$$cor d_{ij} = 1 - r_{ij}$$

$d_{ij}=1-r_{ij}$, con un campo di variazione tra 0 e 2. Se uguale a 0, ovvero distanza minima, le variabili sono positivamente correlate, se uguale a 2 quindi la distanza massima, le variabili sono negativamente correlate tra loro.

Le distanze citate sono definite quando le variabili sono quantitative, e le osservazioni dunque sono di tipo numerico. Se invece le variabili sono di tipo qualitativo, o un misto tra variabili qualitative e quantitative è consigliabile utilizzare **l'indice di dissimilarità di Gower**, la cui funzione è riportata nel seguito

$$gower d_{ij} = 1 - \frac{\sum_{h=1}^p \delta_{ijh} s_{ijh}}{\sum_{h=1}^p \delta_{ijh}}$$

Nello specifico, se la variabile è quantitativa ed $R(h)$ è il suo campo di variazione

$$s_{ijh} = 1 - \frac{|x_{ih} - x_{jh}|}{R(h)}$$

altrimenti se la variabile è qualitativa si ha S_{ijh} uguale a 1 se la i -esima variabile ha la stessa modalità per i e j altrimenti 0. Se invece ci dovessero essere dati mancanti nel dataset allora in generale si avrebbe: δ_{ijh} uguale a 1 se si conoscono i valori della h -esima variabile per i e j ; 0 in caso di accordo 0/0 per variabili binarie si avrebbe 0.

Dopo aver applicato una misura di distanza ai dati si ottiene la cosiddetta matrice delle distanze un esempio è riportato nella figura 3. nella quale sono riportate i valori massimi di dissimilarità tra ciascuna coppia di osservazioni.

$$D = \begin{bmatrix} 0 & d_{12} & \dots & d_{1n} \\ & 0 & \dots & d_{2n} \\ & & \ddots & \vdots \\ & & & 0 \end{bmatrix}$$

Figura 3.10, Esempio matrice di distanza.

APPROFONDIMENTO METODO GERARCHICO

L'obiettivo del metodo gerarchico è quello di passare da uno scenario in cui ci sono n cluster di una singola unità a una posizione in cui esiste un solo cluster che contiene tutte le n unità unendo i cluster meno distanti in una serie di passaggi.

Di conseguenza, il risultato dell'utilizzo di metodi gerarchici non è una singola partizione delle n unità, ma piuttosto una raccolta di partizioni che possono essere rappresentate graficamente utilizzando un "dendogramma" o un "diagramma ad albero", in cui il livello di separazione viene visualizzato sull'asse delle ascisse e le singole unità sono riportate sull'asse delle ordinate. È importante dunque specificare che innanzitutto durante un'analisi è importante, come descritto precedentemente, rilevare le distanze di misura tra le variabili utilizzando per esempio la distanza Euclidea, di Gower o Manhattan, permettendo di specificare il grado di somiglianza tra le variabili oggetto di analisi. Il secondo passaggio è quello di scegliere un algoritmo che mi permetta di raggruppare i cluster sulla base di un determinato criterio di distanza attraverso una procedura che può essere di tipo agglomerativo o divisivo, come precedentemente descritto.

Il terzo step per implementare il metodo di cluster gerarchico è quello di rilevare la distanza tra i cluster i metodi più comuni utilizzati sono:

Metodo del legame singolo (nearest neighbour)

In questo metodo la distanza tra due cluster è definita come la distanza tra i due membri più vicini, o vicini come riportato dalla figura 3.11. Questo metodo è relativamente semplice ma viene spesso criticato perché non tiene conto della struttura dei cluster e può causare un problema chiamato concatenamento, per cui i cluster finiscono per essere lunghi e disordinati. Tuttavia, è meglio degli altri metodi quando i grappoli naturali non sono di forma sferica o ellittica. Consente inoltre di individuare grappoli di qualsiasi forma e mette in evidenza valori anomali con miglior risultati rispetto gli altri metodi.

$$d(C, D) = \min(d_{ij}) , \forall i \in C, \forall j \in D .$$

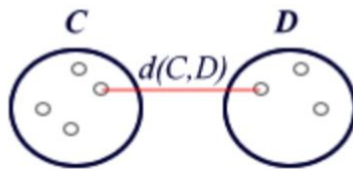


Figura 3.11, Rappresentazione dell'esempio di metodo del legame singolo. Fonte: <http://www.rescoop.com/multivariata/AnalisiCluster.htm>

Metodo del legame completo (furthest neighbour)

In questo caso la distanza tra due cluster è definita come la distanza massima tra i membri, ovvero la distanza tra i due soggetti più distanti. Questo metodo tende a produrre cluster compatti di dimensioni simili ma, come per il metodo del legame singolo, non tiene conto della struttura dei cluster. È anche abbastanza sensibile ai valori anomali.

Formula ed un esempio di rappresentazione del metodo del legame completo sono raffigurate dalla figura 3.12.

$$d(C, D) = \max(d_{ij}) \quad \forall i \in C, \forall j \in D$$

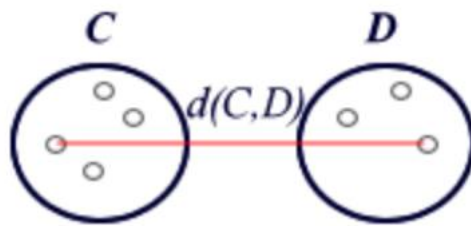


Figura 3.12, Rappresentazione del metodo del legame completo. Fonte: <http://www.rescoop.com/multivariata/AnalisiCluster.htm>

Metodo di collegamento medio (average linkage)

La distanza tra due cluster è calcolata come distanza media tra tutte le coppie di soggetti nei due cluster. Questo è considerato un metodo abbastanza robusto. Un esempio di raffigurazione del metodo del collegamento medio è riportato nella figura 3.13.



$$d(C, D) = \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=2}^{n_2} d_{ij}, \quad \forall i \in C, \forall j \in D$$

Figura 3.13, Rappresentazione del metodo del legame completo. Fonte: <http://www.rescoop.com/multivariata/AnalisiCluster.htm>

Metodo centroide

Qui viene calcolato il centroide (valore medio per ogni variabile) di ciascun cluster e viene utilizzata la distanza tra i centroidi. I cluster i cui centroidi sono più vicini tra loro vengono uniti. Questo metodo come il collegamento medio è ritenuto abbastanza robusto. Un esempio raffigurativo è riportato dalla figura 3.14.

$$d(C, D) = d(\bar{X}_C, \bar{X}_D)$$

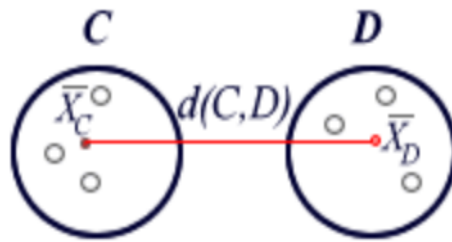


Figura 3.14, Rappresentazione del metodo cetroide. Fonte: <http://www.rescoop.com/multivariata/AnalisiCluster.htm>

Metodo di Ward

In questo metodo vengono combinate tutte le possibili coppie di cluster e viene calcolata la somma delle distanze al quadrato all'interno di ciascun cluster. Questo viene quindi sommato su tutti i cluster. Viene scelta la combinazione che dà la somma di quadrati più bassa. Questo metodo tende a produrre grappoli di dimensioni approssimativamente uguali, il che non è sempre desiderabile. È anche abbastanza sensibile ai valori anomali. Nonostante ciò, è uno dei metodi più popolari, insieme al metodo di collegamento medio. In genere è una buona idea provare due o tre dei metodi di cui sopra. Se i metodi concordano ragionevolmente bene, i risultati saranno molto più credibili. Una volta effettuata la cluster analysis è quindi necessario selezionare la soluzione di cluster “migliore”.

Il processo di aggreffazione può essere rappresentato tramite diagramma ad albero noto anche come “dendrogramma”. Questo diagramma illustra quali cluster sono stati uniti in ogni fase dell'analisi e la distanza tra i cluster al momento dell'unione.

Come riportato nella figura 3.15 nell'asse verticale sono segnalate le distanze tra i cluster, nell'asse orizzontale invece le unità del cluster. Il dendrogramma, dunque, mostra la distanza tra i gruppi o i cluster: maggiore è la distanza tra due aggregazioni, più distanti sono i due cluster.

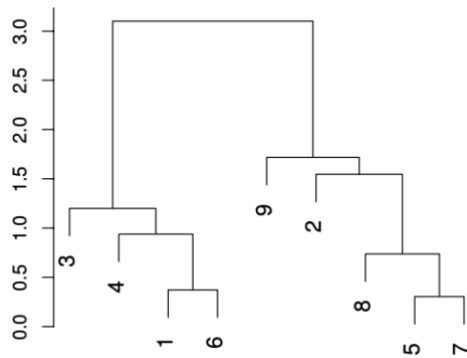


Figura 3.15, Esempio di dendrogramma. Fonte: *Introducing to statistical learning*

METODO NON GERARCHICO

I passi per lo sviluppo di un algoritmo di cluster non gerarchico sono i seguenti:

1. Scegliere i centri del cluster iniziale (essenzialmente questo è un insieme di osservazioni che sono molto distanti: ogni soggetto forma un cluster di uno e il suo centro è il valore delle variabili per quel soggetto).
2. Assegnare a ciascun soggetto il suo cluster "più vicino", definito in termini di distanza dal baricentro.
3. Trovare i centroidi dei cluster che si sono formati
4. Ricalcolare la distanza da ciascun soggetto a ciascun centroide e spostare le osservazioni che non si trovano nel cluster a cui sono più vicine.
5. Continuare finché i centroidi non rimangono relativamente stabili.

L'analisi dei cluster non gerarchica tende ad essere utilizzata quando sono coinvolti insiemi di dati di grandi dimensioni. A volte è preferito perché consente ai soggetti di spostarsi da un cluster all'altro (questo non è possibile nell'analisi dei cluster gerarchica in cui un soggetto, una volta assegnato, non può spostarsi in un cluster diverso).

Tra gli svantaggi che si possono incontrare dall'analisi dei cluster non gerarchica c'è: la difficoltà di sapere quanti cluster selezionare all'inizio della procedura, portando a ripetere più volte l'analisi per arrivare ad un risultato stabile.

Una possibile strategia da adottare è utilizzare inizialmente un approccio gerarchico per determinare quanti cluster ci sono nei dati e quindi utilizzare i centri di cluster ottenuti da questo come centri di cluster iniziali nel metodo non gerarchico.

I metodi principali non gerarchici vengono riassunti così:

- K-means clustering: ogni cluster è rappresentato dal centro o dalla media delle osservazioni appartenenti al cluster. Il metodo è sensibile ad osservazioni anomale e outliers.
- K-medoids clustering o PAM: ogni cluster è rappresentato da una delle osservazioni nel cluster stesso. Il metodo PAM è meno sensibile agli outliers rispetto all'approccio K-means
- Algoritmo CLARA (Clustering Large Applications): è una estensione del metodo PAM adattato a grandi insiemi di dati

Specificatamente, nell'approccio **K-means**. I cluster sono rappresentati dalla media delle osservazioni che sono presenti in esso. L'idea di base del metodo K-means è quello di definire i cluster in modo tale che la variazione totale del cluster WSS, *Whithin sum of squares*, ovvero la somma delle medie della variabilità al quadrato, sia la più piccola possibile. La formula di WSS è definita nello specifico come la media di tutte le distanze euclidee quadrate a coppie di osservazioni appartenenti al cluster h. Più basso è il valore di WSS più è alta la compattezza dei cluster:

Within Cluster Sums of Squares :
$$WSS = \sum_{i=1}^{N_c} \sum_{x \in C_i} d(x, \bar{x}_{C_i})^2$$

I passaggi per applicare l'approccio K-means sono dunque:

- Selezionare k unità dal set di dati come centri di cluster iniziali;
- Assegnare a ciascuna unità al più vicino dei cluster, in base alla distanza euclidea tra le unità e i centroidi;
- Per ciascuno dei k cluster, bisogna calcolare il nuovo cetroide (ovvero i valori medi delle unità assegnate ai cluster)
- Verificare che ogni unità sia assegnata al cluster con la distanza euclidea più vicina al baricentro. Se sono presenti unità che non verificano questa condizione, esse devono essere riallocate;
- Devono essere ripetuti i passaggi 2 e 3 finchè non è più possibile riallocare alcune unità;

I vantaggi di questo metodo sono inanzitutto la velocità rispetto il metodo gerarchico, perché viene calcolata solo la distanza di ciascuna unità dai centroidi al contrario del calcolo della distanza tra cluster utilizzata nel metodo gerarchico. Inoltre, è più affidabile il metodo non gerarchico con dati di grandi dimensioni.

I fattori critici invece del metodo K-means sono la scelta del numero di cluster k richiesta a priori che influisce sul risultato finale, che è necessaria per i metodi gerarchici.

Per determinare il numero di cluster ottimale per entrambi i metodi quindi sia per il gerarchico e sia per il non gerarchico si può utilizzare il metodo Elbow o il metodo Silhouette.

Metodo Elbow

Si ricorda che l'idea alla base dei metodi di partizionamento, come il clustering di K-means è definire i cluster in modo tale che la variazione totale all'interno del cluster [o la somma del quadrato totale all'interno del cluster (WSS)] sia ridotta al minimo.

Il totale WSS misura la compattezza del clustering il quale deve essere il più piccolo possibile.

Il metodo Elbow considera il WSS totale in funzione del numero di cluster: si dovrebbe scegliere un numero di cluster in modo che l'aggiunta di un altro cluster non migliori di troppo il WSS totale. Un esempio grafico è rappresentato dalla figura 3.16, dove si può notare che il numero di cluster ottimale è pari a 2, successivamente alla quale si può notare una deflessione della retta, detta “gomito”.

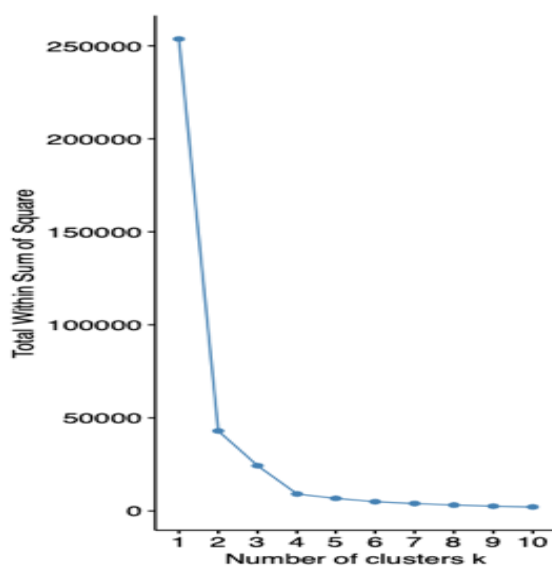


Figura 3.16, Esempio grafico Elbow. Fonte: Metodi quantitativi per la segmentazione e posizionamento, Università Ca' Foscari di Venezia

Metodo Silhouette

La Silhouette misura la distanza di ciascuna unità dal cluster più vicino rispetto al cluster a cui appartiene. Si può comprendere la separazione tra cluster utilizzando la formula

$$S_i = \frac{b_i - a_i}{\max(a_i, b_i)}$$

Dove a_i rappresenta la misura della distanza media del punto i da tutte le unità del suo cluster; b_i misura la distanza media del punto i da tutte le unità del cluster più vicino.

Un esempio di rappresentazione del metodo di Silhouette è rappresentato dalla figura 3.17 dove viene scelto il punto di massimo dunque 3.

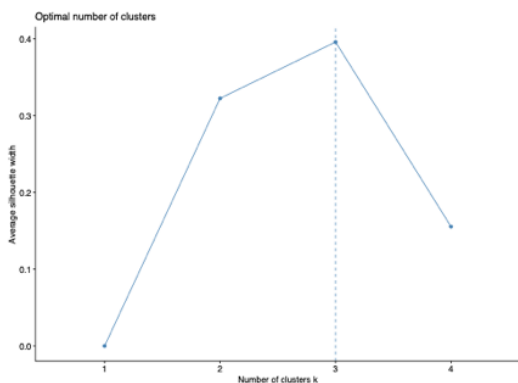


Figura 3.17, Esempio grafico Silhouette. Fonte: Metodi quantitativi per la segmentazione e posizionamento, Università Ca' Foscari di Venezia

3.5.2 MODELLO DI REGRESSIONE LOGISTICA

Un modello di regressione fornisce una funzione che descrive la relazione tra una o più variabili indipendenti e una variabile di risposta, dipendente o target. Nello specifico l'analisi di regressione è una tecnica di modellazione predittiva che analizza la relazione tra la variabile target o dipendente e la variabile indipendente in un set di dati. Esistono diversi tipi di tecniche di analisi di regressione che vengono utilizzati quando la variabile target e quella indipendente mostrano una relazione lineare o non lineare tra loro e la variabile target ha natura numerica o qualitativa.

REGRESSIONE LINEARE

La regressione lineare rappresenta un modello in cui si vuole studiare la relazione lineare tra più variabili predittivi e una variabile dipendente correlate linearmente tra loro. Se vengono coinvolte più variabili indipendenti allora la regressione lineare prende il nome di regressione lineare multipla.

Una regressione lineare multipla è definita dalla seguente formula:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \epsilon$$

Se $p=1$ si parla di regressione lineare semplice. Nella formula β_0 è il valore di y quando X_k è zero. β_p è il coefficiente di regressione, ovvero la variazione attesa di y quando X_p aumenta. X_p è la generica variabile indipendente $k=1 \dots p$, infine ϵ è l'errore stimato nella regressione. La regressione lineare semplice può essere utilizzata per trovare l'intensità della dipendenza tra due o più variabili o studiare il valore della variabile dipendente su un valore esplicito delle variabili indipendenti. La regressione lineare multipla è utilizzata principalmente per stimare quanto due o più variabili indipendenti influenzano la singola variabile dipendente o per trovare il valore delle variabili dipendenti in una condizione definita di tutte le variabili indipendenti.

I modelli di regressione si basano su alcune assunzioni fondamentali di seguito enunciate:

- Sussiste una relazione lineare tra un insieme di variabili p (esplicative) e una variabile risposta su ognuna delle n unità statistiche;
- Gli errori sono a media nulla, i.e. $E(\epsilon) = 0$, varianza costante, i.e. $V(\epsilon) = \sigma^2 I_n$, e incorrelati tra loro, i.e. $cov(\epsilon_i, \epsilon_j) = 0$. L'assunzione di varianza costante è detta omoschedasticità.
- Gli errori seguono una distribuzione multi-Normale $\epsilon \sim N(\mu, \Sigma)$, dove $\mu = 0$ e $\Sigma = \sigma^2 I_n$.
- I valori x_{ij} (per $i = \dots n$) delle variabili esplicative X_j (per $j = 1 \dots, p$) sono osservati senza errore⁴⁸.

Dopo aver elencato alcune delle assunzioni più importanti sui modelli di regressione, è rilevante introdurre il concetto della stima dei parametri del modello di regressione.

⁴⁸ GARETH J., DANIELA W., TREVOR H., ROBERT T., *An Introduction to statistical learning*, Seconda edizione.

Per rilevare una stima dei coefficienti B_j definiti ignoti occorre innanzitutto un campione di n osservazioni per la variabile dipendente e per le p variabili esplicative.

Lo scopo principale è quello di ottenere le stime di parametri B_j , indicate con il termine $\hat{\beta}_j$ tali che l'iperpiano stimato passi il più vicino possibile ai dati osservati, viene dunque rappresentata la formula utilizzata:

$$\underbrace{y_i}_{\text{valore osservato}} \approx \underbrace{\hat{\beta}_0 + \hat{\beta}_1 \cdot x_{i1} + \dots + \hat{\beta}_p \cdot x_{ip}}_{\text{valore stimato}} \quad (i = 1, \dots, n),$$

Viene riportata successivamente nella figura 3.18 un esempio di retta di regressione con $p=1$.

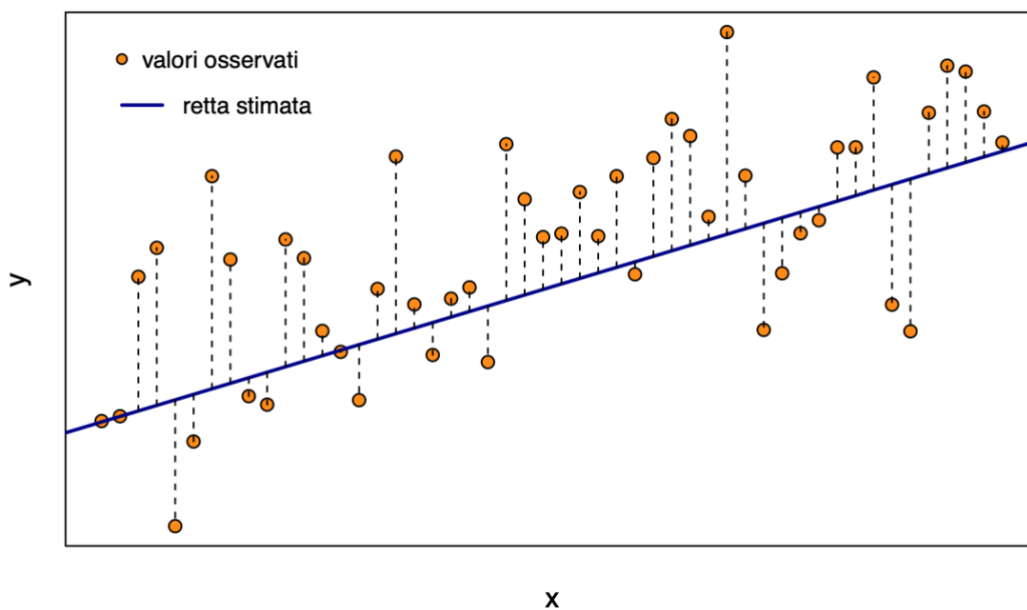


Figura 3.18, Esempio di rappresentazione di una retta lineare con valori osservati e retta stimata.
Fonte: Statistica di base, Apogeo, D.Moore

Un metodo per ottenere le stime citate pocanzi è il metodo dei minimi quadrati ordinari o meglio definito con il termine OLS (Ordinary Least Squares).

OLS presuppone che:

- Sia $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 \cdot x_{i1} + \dots + \hat{\beta}_p \cdot x_{ip}$, il valore stimato della variabile Y in corrispondenza delle i-me osservazioni delle variabili esplicative X.

- Sia $e_i = y_i - \hat{y}_i = y_i - (\hat{\beta}_0 + \hat{\beta}_1 \cdot x_{i1} + \dots + \hat{\beta}_p \cdot x_{ip})$, il residuo i -esimo
- Sia la somma dei quadrati dei residui, definito anche come RSS (*Residual sum of squares*), la cui formula è riportata

$$RSS = e_1^2 + e_2^2 + \dots + e_i^2 + \dots + e_n^2 = \sum_{i=1}^n e_i^2$$

Con il metodo dei minimi quadrati si cerca dunque di determinare i valori di $\hat{\beta}_j$ che minimizzano RSS, quest'ultima in termini generali, è una metrica utilizzata nell'analisi di regressione per determinare la distanza tra le osservazioni disponibili e le previsioni delle stesse ottenute dal modello stimato.

In un'analisi di regressione, l'obiettivo è costruire una funzione (modello) che riesca a spiegare come è stato generato un insieme di dati. RSS, dunque, è una delle metriche utilizzate per valutare quanto una funzione o un modello rappresenti bene i dati.

REGRESSIONE LOGISTICA

La regressione logistica, si introduce quando la variabile dipendente è dicotomica o binaria. La regressione viene utilizzata per descrivere i dati e per spiegare la relazione tra una variabile binaria dipendente e una o più variabili indipendenti nominali, ordinali, di intervallo o di rapporto.

Nel modello di regressione logistica, dunque, si considera una variabile di risposta binaria Y , la cui distribuzione Bernoulliana è $Y \sim \text{Ber}(p)$ e una o più variabili esplicative. L'equazione della regressione logistica per la probabilità che $P(Y = 1|X) = p$, è riportata nell'equazione seguente il caso in cui si mostra presente una sola variabile esplicativa X di seguito la formula:

$$p = \frac{e^{\beta_0 + \beta_1 X}}{1 + e^{\beta_0 + \beta_1 X}}$$

L'equazione di cui sopra è chiamata modello logistico o regressione logistica e viene utilizzata per prevedere una variabile dipendente binaria. L'equazione può essere riscritta con questa terminologia:

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X$$

Nello specifico, $p/1-p$ è definito odd, invece la funzione $\log(p/1-p)$ è definito log odd. La funzione odd è la probabilità tra il rapporto di un evento e la probabilità che esso non accada. Più p sarà elevato più alta sarà la misurazione di odd e log odd. Nella regressione logistica si può estendere il modello al caso multiplo, introducendo più variabili esplicative X_i , $i=1, \dots, p$.

In un modello di regressione logistica, β_i rappresenta la variazione media della log quota in seguito all'aumento di X_i di un'unità.

Per ottenere l'effetto sulla probabilità naturalmente basta applicare la funzione inversa del logit sul coefficiente stimato.

Indipendentemente invece dal valore di X_i , se β_i è positivo allora l'aumento di X sarà associato all'aumento di $p(X_i)$, e se β_i è negativo allora l'aumento di X sarà associato alla diminuzione di $p(X_i)$.

Concludendo, la regressione logistica è uno degli approcci di analisi più utilizzati in quanto è facile da implementare e versatile.

Per una disamina più approfondita dell'argomento, si rimanda al testo di Hastie, Tibshirani e Friedman "*The elements of statistical learning*" (2009)

CAPITOLO IV: IL SISTEMA E-COMMERCE SUPERMERCATI ED IPERMERCATI: PRESENTAZIONE CASO STUDIO

4.1: IL QUESTIONARIO

Nei capitoli precedenti sono stati introdotti concetti come la struttura delle organizzazioni di vendita di ipermercati e supermercati e l'organizzazione dei servizi offerti di E-commerce. L'obiettivo della tesi è quello di ottenere informazioni su quali siano le i fattori che possono influenzare le persone ad utilizzare uno strumento come l'E-commerce per effettuare la spesa. È importante comprendere quali siano gli svantaggi e vantaggi che derivano dal servizio online e quali siano i fattori che incentivano le persone al suo utilizzo. Come verrà poi specificato successivamente, la maggior parte degli intervistati hanno iniziato ad utilizzare il servizio E-commerce a causa della pandemia Covid-19 e per la scarsa possibilità di spostarsi con la propria autovettura nel punto vendita abituale.

Per sviluppare dettagliatamente il tema della tesi, si è ricorso ad un questionario per raccogliere informazioni di natura campionaria con finalità esplorativa. Il questionario è uno strumento utile per la raccolta di informazioni su argomenti specifici e consente un confronto tra i vari dati raccolti per realizzare ricerche di mercato. Aiuta inoltre a comprendere ciò che soddisfa maggiormente i consumatori investigando sulle attitudini di consumo.

Per costruire il questionario usato per raccogliere i dati si è proceduto nel seguente modo:

- Definizione di uno schema concettuale
- Redazione del questionario
- Verifica delle risposte ottenute

Definizione schema concettuale

Nell'figura 4.1 viene riportato uno schema semplificato del questionario, descrivendo i temi principali su cui si vuole ricavare informazioni necessarie. Il questionario è diviso principalmente in quattro parti: la prima è incentrata su domande per raccogliere informazioni sociodemografiche. Utilizzando una domanda filtro le domande per gli intervistati si dividono in due sezioni separate: una parte riguardante solo a coloro che acquistano online da supermercato e l'altra parte invece riguarda a coloro che non acquistano online da supermercato ma preferiscono effettuare la propria spesa

fisicamente. L'ultima parte è indirizzata all'intera popolazione che ha compilato il questionario, ed è relativa a domande generiche sugli acquisti effettuati in rete ed i servizi che vengono usufruiti.



Figura 4.1: Mappa descrittiva del questionario E-commerce supermercati ed ipermercati. Fonte: Elaborazione personale

Redazione del questionario

La redazione del questionario è strutturata tramite una “successione a imbuto”: nella prima parte vengono utilizzati quesiti di interesse generale come le caratteristiche sociodemografiche di una persona e nelle parti restanti vengono poste domande più mirate che hanno come obiettivo la valutazione delle sensazioni e atteggiamenti dell'intervistato. Nella prima parte vengono inserite domande di tipo sociodemografico, nello specifico si cerca di raccogliere informazioni investigando su: sesso, cittadinanza, regione di residenza, età, stato civile, domande relative alla composizione familiare, titolo di studio conseguito ed infine se l'individuo svolge un'attività lavorativa oppure no. Le domande di tipo sociodemografiche sono essenziali, per indagare su quale segmento appartiene la persona rispetto il campione generale. È necessario specificare che alla fine della prima parte del questionario dopo aver risposto all'ultima domanda filtro: “Ha mai acquistato online da un supermercato?”, il questionario si divide in altre due parti: una per coloro che hanno affermato l'acquisto online da supermercato ed un'altra per chi ha risposto in modo negativo alla domanda in questione.

Nella seconda parte, si vuole indagare l'attitudine al consumo di un individuo che effettua acquisti online da supermercato o ipermercato. In questo caso vengono utilizzate domande a risposta aperta, a risposta chiusa, a risposta multipla e risposte utilizzando

scala di valutazione delle preferenze. Nello specifico nel seguito vengono presentate le domande di questa sezione.

“Dove acquista maggiormente i prodotti?”, le possibili risposte al quesito sono due: supermercato o ipermercato, in questo caso si cerca di investigare su quale servizio l'intervistato utilizza maggiormente per procurarsi la propria spesa. Qui si cerca di raccogliere dati qualitativi nominali, sul dataset la domanda viene rinominata ACQUISTO_PRODOTTI_GEN;

“Con che frequenza acquista i propri prodotti in supermercato/ipermercato (fisicamente)?”, la domanda è a risposta multipla e sono presenti risposte come: quotidianamente, settimanalmente, mensilmente, annualmente, mai. Si cerca di comprendere quante volte una persona acquista la propria spesa. I dati raccolti sono qualitativi ordinali, sul dataset la domanda viene rinomata FREQUENZA_ACQUISTO_SUP;

“Con che frequenza online da supermercato/ipermercato?”, la formattazione della domanda è a risposta multipla, al rispondente viene proposto le stesse risposte presenti nella domanda precedente. L'obiettivo è quello di raccogliere i dati sulla frequenza di acquisto da parte delle persone che acquistano online. I dati raccolti sono qualitativi ordinali e sul dataset la domanda viene rinominata: FREQUENZA_ONLINE_SUPERMERCATO;

“Quale è la spesa media di un acquisto online da supermercato?”, la domanda è a risposta multipla: sotto i 20 euro, dai 20 ai 50 euro, dai 50 ai cento euro, dai 100 ai 200 euro, dai 200 ai 300 euro ed infine dai 300 in poi. Le risposte servono a raccogliere dati sull'acquisto medio da parte dei rispondenti. Nel dataset la raccolta di dati prende la denominazione di SPESA_MEDIA;

“In che fornitori acquista i propri prodotti online?”, in questa domanda è possibile scegliere più di una risposta. Vengono presentati una serie di supermercati ed ipermercati, in modo tale da ricavare dati sui supermercati più utilizzati per la propria spesa online. Qui si ricavano dati qualitativi nominali e nel dataset viene rinominato con il termine FORNITORI_SUPERMERCATO;

“Dove ritira la propria spesa?”, vengono elencate risposte come: direttamente dal punto vendita, direttamente a casa e ritiro della spesa nel LockerBox. Si vuole indagare su come la spesa online viene ritirata dal rispondente. Vengono anche qui ricavati dati qualitativi nominali, nel dataset la domanda viene rinominata con RITIRO_SPESA_ONLINE;

“Che dispositivo utilizza per acquistare la merce online?”, le risposte disponibili sono: tablet, smartphone/telefono e computer. Si vuole evidenziare quale mezzo viene utilizzato maggiormente per effettuare la spesa online. Si ricavano dunque dati qualitativi, nominali e nel dataset i dati vengono raccolti sotto nomina di DISPOSITIVO_SPESA;

“Utilizza un’applicazione o un sito Internet nell’acquistare i prodotti?”, le risposte disponibili sono: app del telefono o ipad e sito internet. Qui si vuole identificare il software utilizzato per la spesa online. I dati ricavati sono qualitativi nominali sul dataset viene identificato con il nominativo di CANALE_ONLINE;

“Quale pagamento online predilige quando acquista online?”, vengono proposte una serie di pagamenti che si possono effettuare acquistando online. Si vuole raccogliere informazioni sul metodo di pagamento che prediligono i rispondenti quando acquistano online. I dati raccolti sono qualitativi nominali, nel dataset prende il nome di PAGAMENTO_ONLINE;

“Quanto dista mediamente il supermercato/ipermercato da casa sua?”, vengono indicate una serie di intervalli di chilometri che misurano la distanza. I dati raccolti sono qualitativi ordinati. Nel dataset i dati vengono raggruppati sotto nominativo di DISTANZA_SUPERMERCATO;

“Come è venuto a conoscenza dell'e-commerce da parte del supermercato/ipermercato?”, vengono elencate alcune delle campagne pubblicitarie utilizzate dai supermercati per incentivare l’uso del servizio E-commerce. I dati raccolti sono qualitativi nominali ed i dati sono raccolti sotto nominativo di: CONOSCENZA_ONLINE;

“Perché acquista online? (1=per nulla d’accordo, 5 totalmente d’accordo)”, le risposte elencate sono ad esempio: posso farmi recapitare gli acquisti a casa e i prezzi sono più bassi. La domanda è utilizzata per ricavare dati quantitativi discreti raccolti

tramite scala Linkert. La domanda ha il fine di evidenziare e quantificare il pregio nell'utilizzo di un servizio E-commerce da parte dei rispondenti;

“Quale fattore può incidere di più sull'acquisto online? (1=per nulla importante, 5=estremamente importante)”. Un esempio di risposte sono prezzi di vendita e a causa della pandemia Covid-19. Qui si vuole indagare sul fattore che incide e che incentiva gli acquisti online. I dati raccolti sono quantitativi discreti raccolti tramite scala Linkert;

“Scelga fra queste risposte quali siano i punti di forza sulla consegna più importanti (1=per nulla importante, 5=estremamente importante)”. Un esempio di risposte a cui il rispondente deve indicare quanto ritiene importante questi fattori di consegna: recapito la sera, recapito di mattina e recapito entro due ore. I dati raccolti sono come i precedenti quindi quantitativi discreti e viene utilizzata una scala Linkert per raccogliarli;

“Quali prodotti prevalentemente acquista online?”. Il rispondente deve scegliere tra prodotti di genere chimico, genere alimentare confezionati, prodotti freschi e prodotti surgelati. Si vuole indagare su quale genere di prodotto il rispondente acquista online prevalentemente. I dati raccolti sono qualitativi nominali. Nel dataset vengono raccolti con il nominativo di PRODOTTI_ONLINE;

“Utilizzavate l'e-commerce del punto vendita in questione prima della pandemia Covid-19?”. Il rispondente deve scegliere tra sì e no. Si vuole indagare se i rispondenti hanno iniziato ad utilizzare il servizio dell'E-commerce di supermercati ed ipermercati pre o post Covid-19. Qui si ricavano dati qualitativi nominali. Nel dataset vengono raccolti sotto il nome di: COVID;

“Quali forme di assistenza alla clientela sono importanti per lei?”. Vengono elencate alcune delle possibili modalità di assistenza al cliente. Vengono raccolti dunque dati qualitativi nominali e nel data set vengono raccolti sotto la voce di ASSISTENZA;

“Quali secondo lei possono essere gli svantaggi sugli acquisti online?”. Il rispondente deve scegliere tra alcuni dei fattori che sfavoriscono l'utilizzo dell'E-commerce. I dati raccolti sono qualitativi nominali e nel dataset vengono raccolti sotto il nome di SVANTAGGI_ONLINE;

“Da una scala da 1 a 5 quanto vi ritenete soddisfatti dei vostri acquisti online?”. Si indaga dunque quanto mediamente i rispondenti sono soddisfatti del servizio offerto dell’E-commerce dei supermercati ed ipermercati. I dati raccolti sono quantitativi discreti e vengono raccolti tramite scala Linkert. Nel dataset vengono raccolti sotto la denominazione di SODDISFAZIONE_ONLINE;

La terza parte del questionario è indirizzata a coloro che non effettuano acquisti online da supermercato ed ipermercato. Si vuole dunque indagare su quali siano le variabili che influenzano la loro scelta di non fruizione del servizio. Le prime due domande sono le medesime della seconda parte, si vuole indagare dove il consumatore acquista la propria spesa e con che frequenza. Nel dataset i dati sono raccolti rispettivamente sotto il nominativo di: ACQUISTO_PRODOTTI_GEN e FREQUENZA_SUP. Successivamente le domande sono diverse.

“Perché non ha mai acquistato online da supermercato?”. Il rispondente può scegliere tre delle risposte elencate ad esempio: non forniscono il servizio e-commerce nel suo paese, obbligo di registrazione, scarsa sicurezza nell’acquisto online e preferisce acquistare direttamente al punto vendita. Si vuole indagare sul mancato utilizzo del servizio E-commerce da parte delle persone. I dati raccolti sono qualitativi nominali e nel dataset vengono rinominati NO_ACQUISTO_ONLINE;

“Cosa ne pensa della spesa online offerta dai supermercati/ipermercari?”. Vengono raccolte direttamente risposte da parte dei rispondenti, perchè la domanda in questione è aperta. È difficile trasformare le risposte riportate dagli intervistati in dati; quindi, verranno utilizzate per capire direttamente che pensiero hanno sul servizio E-commerce.

“Secondo lei acquisterà online dai supermercati in futuro?”. Il rispondente deve rispondere sì o no alla domanda. È importante capire se il servizio online verrà utilizzato o no. Le risposte sono raccolte nel dataset con il nominativo di ACQUISTO_ONLINE_FUTURO. I dati raccolti sono qualitativi nominali.

“Quale opzione servirebbe per convincerla ad utilizzare la spesa online?”. Vengono elencate, alcuni dei vantaggi che possono emergere con l’utilizzo dell’E-commerce come: spedizione gratuita, reso gratuito, promozioni e sconti e raccolta punti. I dati ricavati sono

qualitativi nominali e nel dataset prendono il nome di CONVINZIONE_SPESA_ONLINE.

Con la quarta parte si vuole indagare sulla fruizione dei servizi ed acquisti online generici da parte delle persone. Il questionario in questo caso viene compilato sia da coloro che utilizzano il servizio E-commerce e sia da chi non l'utilizza. Vengono poste domande relative all'utilizzo di: Social Media, servizi online maggiormente utilizzati, quali generi di prodotti vengono acquistati maggiormente online, dove ricercano le informazioni per acquistare online ed infine viene posta una domanda relativa alla soddisfazione da parte del rispondente nella fruizione dei vari servizi online offerti dalla rete. I dati raccolti sono qualitativi nominali. Nel dataset vengono indicate con i nomi di: SOFFISFAZIONE_ACQUISTO_ONLINE_GEN, TEMPO_SOCIAL_MEDIA, SERVIZIO_ONLINE, SPEDIZIONE_ACQUISTI_ONLINE E DOVE_CERCHI_ONLINE.

Verifica dei dati

Per verificare la composizione corretta del questionario oggetto di raccolta dei dati, si è adottato un pre-test, per valutare la reazione da parte degli intervistati sull'idea proposta e raccogliere suggerimenti su come possono essere impostate le domande. Il campione oggetto di analisi è stato autoselezionato tramite questionario, di conseguenza al suo interno sussistono degli errori di campionamento. L'obiettivo dell'analisi è di tipo descrittivo e non inferenziale. L'indagine è cominciata precisamente il 15 Febbraio 2022 e terminata il 28 Giugno 2022. Il campione raccolto è composto da 252 unità. Per l'analisi statistica vengono considerate solo le risposte complete e senza dati mancanti raccolte dagli intervistati. Vengono dunque eliminate tutte le unità che non hanno risposte esaustive o che sono state effettuate per il pre-test del questionario. Le risposte effettive che verranno analizzate sono 245.

4.2: CARATTERISTICHE SOCIODEMOGRAFICHE

Gli intervistati sono principalmente provenienti dalla regione Veneto con il 90%, le altre regioni con un tasso di risposta minore sono: Lombardia (4%), Friuli-Venezia Giulia (1,6%), Campania (0,8%) ed infine con lo 0,4%, Umbria, Toscana, Sicilia, Piemonte, Molise, Calabria e Trentino-Alto Adige. Nel questionario per raccogliere informazioni sono utilizzate scale nominali con la quale si possono valutare i soggetti che appartengono alla stessa o differenti classi come: il genere, lo stato lavorativo, la regione di provenienza

o lo stato civile. Con i dati sociodemografici si possono evidenziare le percentuali di peso che hanno sul campione in oggetto.

Vengono presentati i dati raccolti effettuando un'analisi sociodemografica sulle unità che hanno risposto al questionario.

Nella figura 4.2 viene riportata la distribuzione di frequenze della variabile genere. Emerge la prevalenza del sesso femminile con un 64% rispetto il 36% del genere maschile.

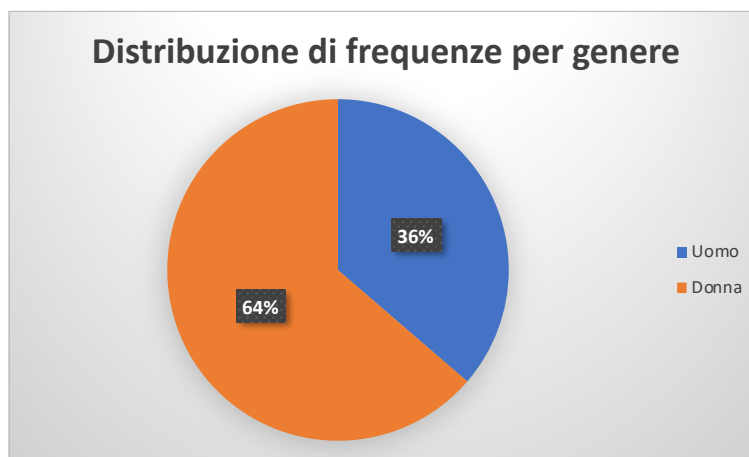


Figura 4.2, Percentuale di risposte in base al genere. Fonte: Google questionare

FASCIA D ETA'	PESO SUL CAMPIONE	FEMMINE	MASCHI
Età 18-24	52,7%	38,5%	14,2%
Età 25-35	27,8%	13,2%	14,6%
Età 36-50	7,3%	4,0%	3,3%
Età 51-65	11,0%	7,2%	3,8%
Età maggiore di 65	1,2%	1,2%	0%

Figura 4.3, Tabella distribuzione d'età del campione. Fonte: Elaborazione personale

Nella tabella 4.3 viene riportato la distribuzione in percentuale del campione in base alla fascia d'età. Le unità che hanno risposto con maggior misura al questionario sono comprese tra i 18 e 35 anni. La fascia che contiene un'età compresa tra i 36 e 50 e la fascia d'età maggiore di 65 sono quelle che hanno un tasso di risposta minore al questionario, con rispettivamente il 7% e 1%. Infine, la fascia tra i 50-65 anni ha una percentuale di risposta del 11%. Dalla tabella 4.3 viene reso evidente come il genere femminile abbia un peso maggiore di risposta tre la fascia d'età tra i 18 e i 24 anni,

identica invece la percentuale di risposta per il genere maschile per la fascia d'età tra i 18-24 e 25-35 anni con un 15%. Non viene rilevata alcuna risposta da parte della fascia d'età maggiore di 65 anni di genere maschile.

Il confronto viene reso più evidente con la figura 4.4 dove il genere femminile domina il campione analizzato soprattutto nella fascia d'età 18-24, dunque il peso di questa fascia avrà rilevanza sull'analisi dei dati.

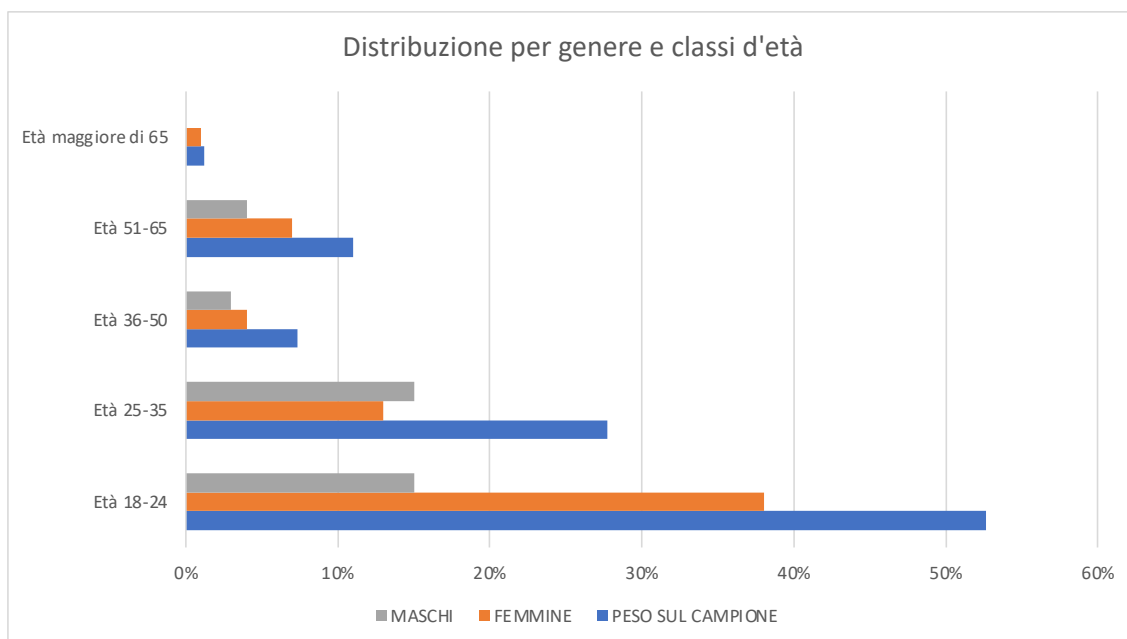


Figura 4.4, Distribuzione per genere e classi d'età. Fonte: Google questionare

È importante ricordare che i dati raccolti dal questionario sono rappresentativi come riportato nella figura 4.4 di soggetti giovani, questo è dovuto al fatto che il metodo di somministrazione del questionario è autoselezionato.

Analizzando i dati sul titolo di studio nella figura 4.5 ricavati dal questionario oggetto di analisi, i titoli con peso maggiore sul campione sono: diploma o qualifica media superiore con il 49,8% e laurea con il 42%. Anche in questa situazione il genere femminile domina sul genere opposto. Titoli di studio come dottore di ricerca o specializzazione post-laurea hanno una percentuale di peso sul campione del 2,4%, di cui 1,6% sono maschi e 0,8% sono donne.

	PESO SUL CAMPIONE	MASCHIO	FEMMINA
TITOLO DI STUDIO			
Licenza elementare	0,4%	0,0%	0,4%
Licenza di scuola media inferiore	5,3%	0,8%	4,5%
Diploma o qualifica di scuola media superiore	49,8%	18,8%	31,0%
Laurea	42,0%	15,1%	26,9%
Dottore di ricerca o specializzazione post-laure	2,4%	1,6%	0,8%

Figura 4.5, Tabella distribuzione titolo di studio. Fonte: Elaborazione personale

Vengono successivamente analizzati i dati attinenti all'occupazione lavorativa da parte del campione oggetto di studio. Come viene evidenziato nella tabella 3.5 la maggior parte degli intervistati hanno un impiego, nello specifico il 64,5%, al contrario il 33,5% del campione non lavora ed infine il 2% è disoccupato. Il dato sulla fatispecie dei lavoratori fa emerge che il 38% sono di genere femminile, invece il 26,5% sono di genere maschile.

STATO LAVORO	PESO SUL CAMPIONE	MASCHIO	FEMMINA
Lavoratore	64,5%	26,5%	38,0%
Non Lavoratore	33,5%	9,8%	23,7%
Disoccupato	2,0%	0,0%	2,0%

Figura 4.6, Tabella distribuzione stato lavoro, Fonte. Elaborazione personale

Dall'analisi delle caratteristiche sociodemografiche del campione, vengono raccolte anche le caratteristiche riguardanti la distribuzione sulla composizione familiare e sociale delle unità. Innanzitutto, come riportato dalla figura 4.7 la maggior parte delle unità del campione sono celibe o nubile con una percentuale del 77%; con un peso minore si rileva che il 20,6% delle unità sono coniugate; infine, l'1,6% delle unità sono separate o divorziate e lo 0,8% risultano vedovi.

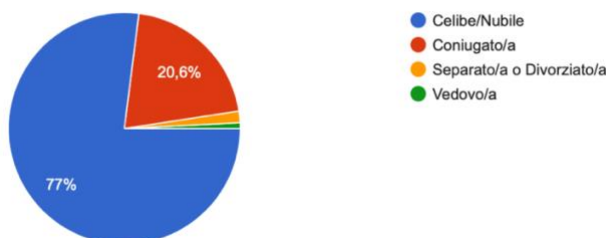


Figura 4.7, Grafico a torta sulla distribuzione della composizione familiare e sociale delle unità, Fonte: Google questionare

Per quanto riguarda la convivenza nella stessa abitazione, prevale come evidenziato nella figura 4.8, la percentuale di coabitazione con i genitori con una percentuale del 30%. Le frequenze percentuali con maggior peso sono poi rappresentate dalle modalità raccolte: convivenza con marito o moglie (compagno o compagna) con il 14,3% e 13,1% con coniuge e figli. Successivamente con minore tasso di risposta le unità convivono solo con i figli 1,2% o da soli 4%. È rilevante ancora rimarcare il fatto che i dati stati raccolti con obiettivo descrittivo e la compilazione è stata somministrata con una autoselezione.

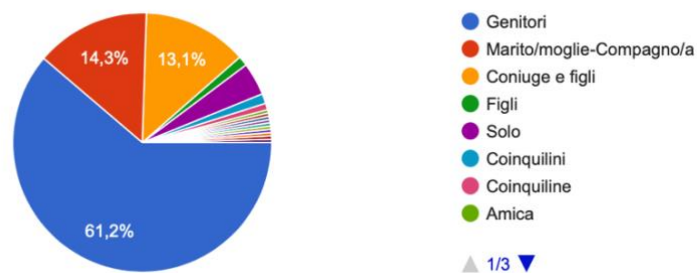


Figura 4.8 Grafico a torta sulla distribuzione della coabitazione familiare, Fonte: Google questionare

In conclusione, è importante ora evidenziare il tasso di risposta da parte del campione sul tema principale della tesi di laurea, l'utilizzo da parte dei consumatori, dell'E-commerce di supermercati o ipermercati per effettuare la spesa. È stato rilevato tramite il questionario che il 71% delle risposte indicano che il campione non utilizza la spesa online; invece, il 29% si come riportato dalla figura 4.9. Con più precisione, 70 unità del campione totale effettuano acquisto online da supermercato ed ipermercato, invece il restante quindi 175 unità non acquistano tramite E-commerce. Si capisce dunque che il mercato virtuale, non è ancora un servizio abituale utilizzato dalle persone.

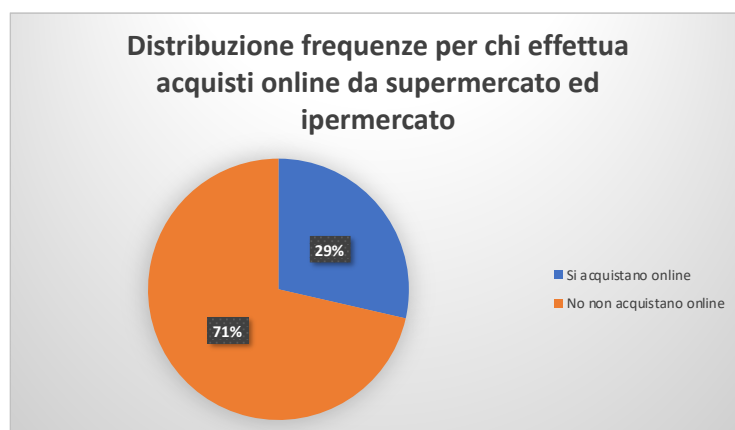


Figura 4.9, Grafico a torta sulla distribuzione frequenze per chi effettua acquisti online da supermercato ed ipermercato, Fonte: Google questionare

4.3 CARATTERISTICHE DEI CONSUMATORI CHE UTILIZZANO L'E-COMMERCE

Come riportato dalla figura 4.9, viene reso noto che il 29% degli intervistati utilizza il servizio E-commerce offerto da supermercati ed ipermercati. In questo capitolo verrà effettuato semplice studio sui dati ricavati. Precisamente 70 unità su 245 utilizza la spesa online. Come viene descritto dalla tabella 4.10, la fascia d'età che utilizza maggiormente il servizio E-commerce è tra i 18 e 24 anni d'età con una percentuale pari al 47%. Di seguito il 30% delle unità acquista la propria spesa online con un'età compresa tra i 25 e 35 anni. Infine, rilevante è che solamente l'1% del campione delle 70 unità che utilizzano la spesa online ha un'età maggiore di 65. È importante sottolineare come nel questionario emerge che il genere femminile ha una presenza maggiore rispetto il genere maschile per la fascia d'età tra i 18 e 24 anni d'età. Questo dato è reso noto anche dalla figura 4.3 riportata precedentemente sulla distribuzione di genere per fascia d'età, dove spicca la prevalenza del genere d'età femminile.

FASCIA D ETA'	PESO SUL CAMPIONE	FEMMINE	MASCHI
Età 18-24	47,1%	69,6%	30,4%
Età 25-35	30,0%	42,8%	57,1%
Età 36-50	12,8%	55,6%	44,4%
Età 51-65	8,5%	50,0%	50,0%
Età maggiore di 65	1,4%	0	100%

Figura 4.10, Tabella distribuzione della fascia d'età e genere su coloro che utilizzano il servizio E-commerce ipermercato e supermercato. Fonte: Elaborazione personale

Tramite il questionario vengono rilevati dati sulla frequenza di acquisto online da supermercato/ipermercato da parte dei consumatori. Viene riscontrato che la maggioranza delle unità 33 precisamente il 47,1%, acquista da supermercato/ipermercato con il servizio online con una cadenza annuale, di seguito 24 quindi il 34,3% degli intervistati acquistano mensilmente, 6 unità invece acquistano settimanalmente con 8,6%. Infine, viene rilevato che solo una persona, quindi l'1,4% acquista online effettua la propria spesa online quotidianamente. I dati riportati vengono rappresentati dal grafico a

barre 4.11, rappresentazione utile per le variabili ordinali. Non è da considerare la variabile mai, perché è ritenuto impossibile che, chi effettua acquisti online potesse rispondere con “Mai”.

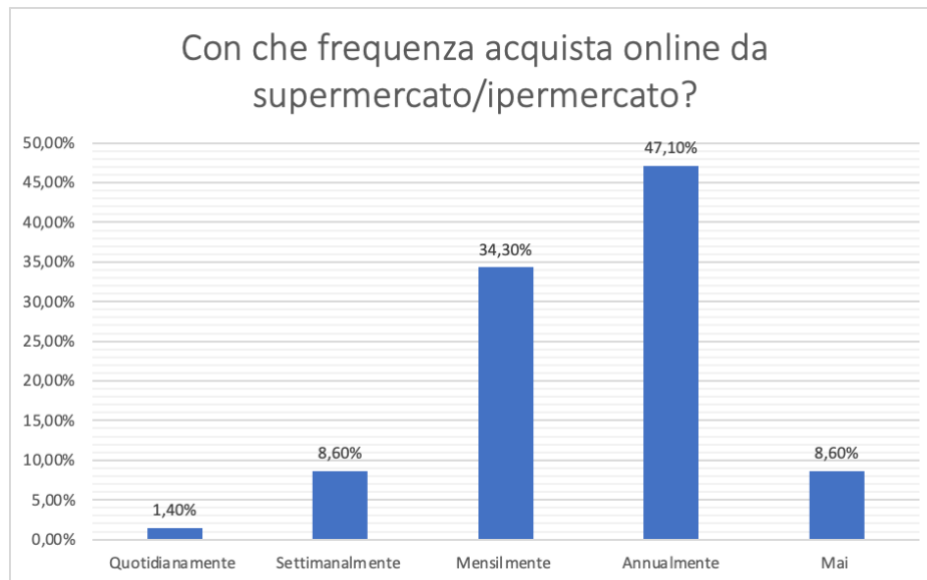


Figura 4.11, Grafico a barre relativo alla distribuzione di frequenza acquisto online da supermercato ed ipermercato. Fonte: Elaborazione personale

Dal questionario vengono raccolti dati sulla spesa media effettuata dal campione che effettua la spesa online. La maggioranza delle unità con il 33,8% effettua una spesa compresa tra i 50 e 100 euro; successivamente il 28,2% degli intervistati realizza una spesa tra i 20 e 50 euro; il 21,1% delle unità che acquistano online effettua una spesa tra i 100 e 200 euro ed infine il 15,5% effettua una spesa al di sotto dei 20 euro. Secondo i dati emerge dunque che la maggior parte degli intervistati mediamente spendono non oltre ai 100 euro di spesa e solo una persona effettua una spesa al di sopra dei 200 euro.

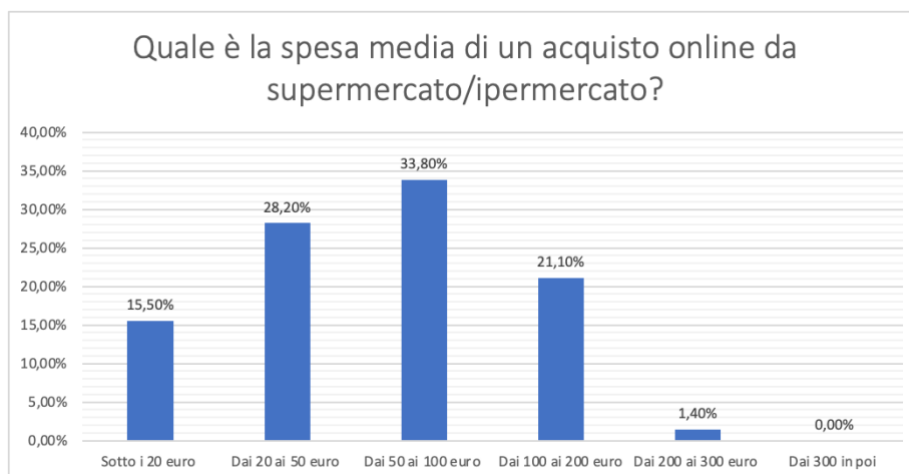


Figura 4.12, Grafico a barre relativo alla distribuzione di frequenza spesa media di un acquisto online supermercato/ipermercato. Fonte: Elaborazione personale

Non rilevante per l'analisi dei dati ma comunque apprezzabile a livello conoscitivo sono i supermercati o ipermercati utilizzati maggiormente dagli intervistati per effettuare la propria spesa online. Il primato è detenuto dall'Ipertosano, in seconda posizione il Conad ed in terza la Coop. Come citato dai capitoli precedenti la maggioranza delle unità dell'intero campione proviene principalmente dal Veneto ed i principali punti vendita dell'Ipertosano sono localizzati in questa regione. Il motivo, inoltre, per cui supermercati ed ipermercati come Conad e Ipertosano, hanno un maggior numero di voti è derivato dal fatto che il questionario è stato somministrato tramite un metodo di autoselezione. Dunque, era prevedibile che Conad ed Ipertosano risultassero i più votati.

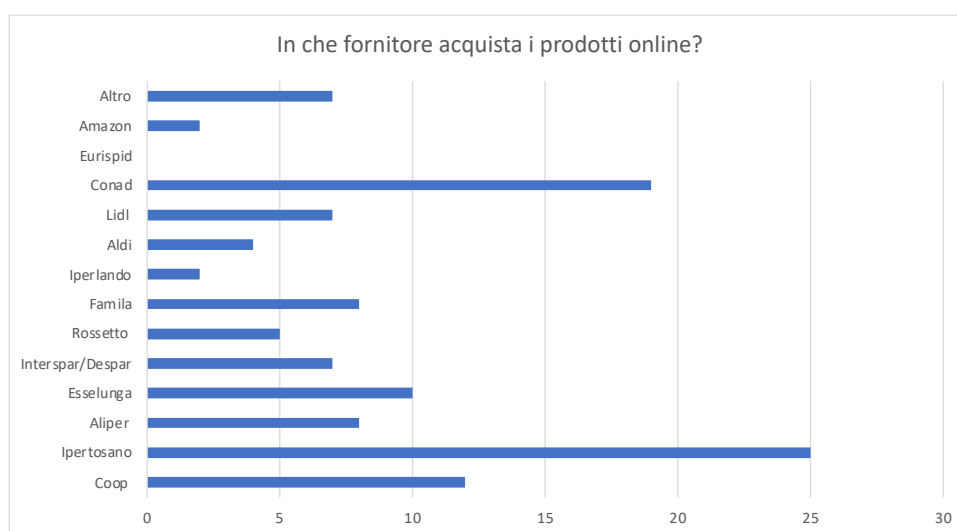


Figura 4.13, Grafico a barre sulla distribuzione dei fornitori principali che forniscono prodotti online. Fonte: Elaborazione personale

Nel questionario, inoltre, si vuole indagare su quale dei servizi di consegna vengono utilizzati maggiormente dagli intervistati. Come riportato dal grafico 4.14, il 69,9% delle unità che effettuano la spesa online ritirano la propria spesa direttamente a casa, invece il 28,8% ritira la propria spesa in punto vendita, risparmiando dunque le spese di spedizione. Ed infine, solamente un intervistato utilizza il servizio locker box offerto dai punti vendita.

Dove ritira la sua spesa?

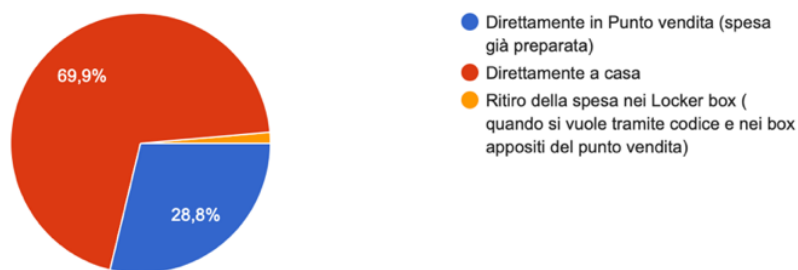


Figura 4.14, Grafico a torta sulla distribuzione del ritiro della spesa online; Fonte: Google questionare

Ai fini dell'analisi, è importante comprendere come le unità del campione sono venute a conoscenza del servizio online offerto da ipermercati e supermercati. Nel questionario gli intervistati potevano elencare più di una risposta come riportato nella figura 4.15. Il primato è detenuto da "Internet" con un tasso di risposte di 35,6% di risposte ed in posizione secondaria da "amici e parenti" con un tasso di risposta del 34,2%. Successivamente, con una tendenza di risposta quasi analoga le due modalità di pubblicità come "direttamente dal punto vendita" e dai "social media", con rispettivamente il 23,3% e 21,9% di percentuale di risposta. Per quanto riguarda invece a pubblicità fisiche come volantini e manifesti il tasso di risposta è nettamente inferiore rispetto le modalità citate pocanzi. Il "volantino" ha un tasso di risposta del 11%, invece "da manifesti" il 6,8%.

Come è venuto a conoscenza dell'e-commerce da parte del supermercato/ipermercato (è possibile elencare più di una risposta)

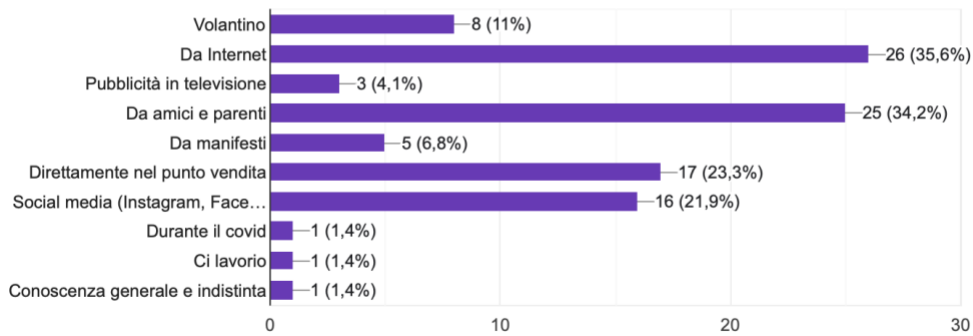


Figura 4.15, Grafico a barre sulla distribuzione dei metodi di conoscenza E-commerce. Fonte: Google questionare

Gli intervistati durante la compilazione del questionario devono indicare tramite una scala Linkert quanto siano d'accordo su una serie di motivazioni di acquisto online tramite supermercato ed ipermercato. Dalla Figura 4.16, che raffigura i dati raccolti, emerge che l'acquisto online è dovuto principalmente da: epidemia Covid-19; velocità di acquisto dei prodotti tramite app o sito senza recarsi direttamente al punto vendita; comodità di ritirare direttamente la spesa nella propria abitazione. È inoltre importante evidenziare come gli intervistati non danno per nulla importanza all'eventualità di trovare prezzi più bassi sul sito internet.

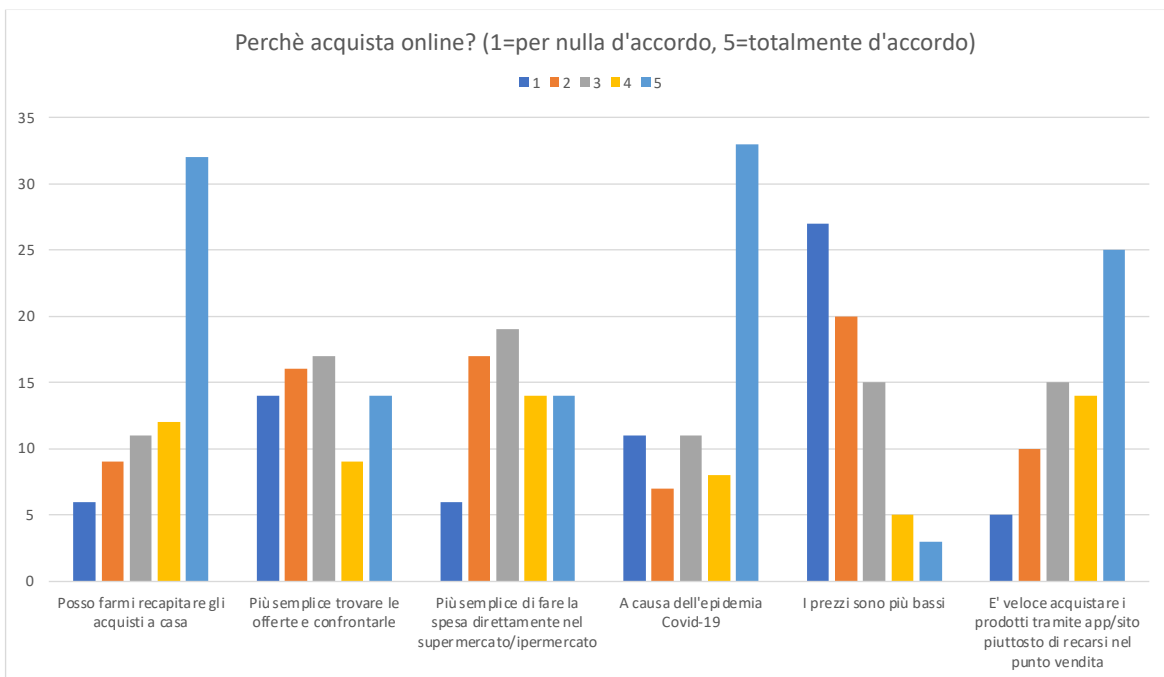


Figura 4.16, Grafico sulla motivazione di acquisto online. Fonte: Elaborazione personale

Interessante è osservare quale siano i fattori che incidono sull'acquisto online dai supermercati ed ipermercati. Come riportato dalla figura 4.17 per gli utilizzatori del servizio E-commerce è estremamente importante usufruire: del servizio della spedizione gratuita e l'invio di ritorno gratuito in caso di reso. Non è importante invece per coloro che acquistano online, la raccomandazione di amici e recensioni sul sito E-commerce.

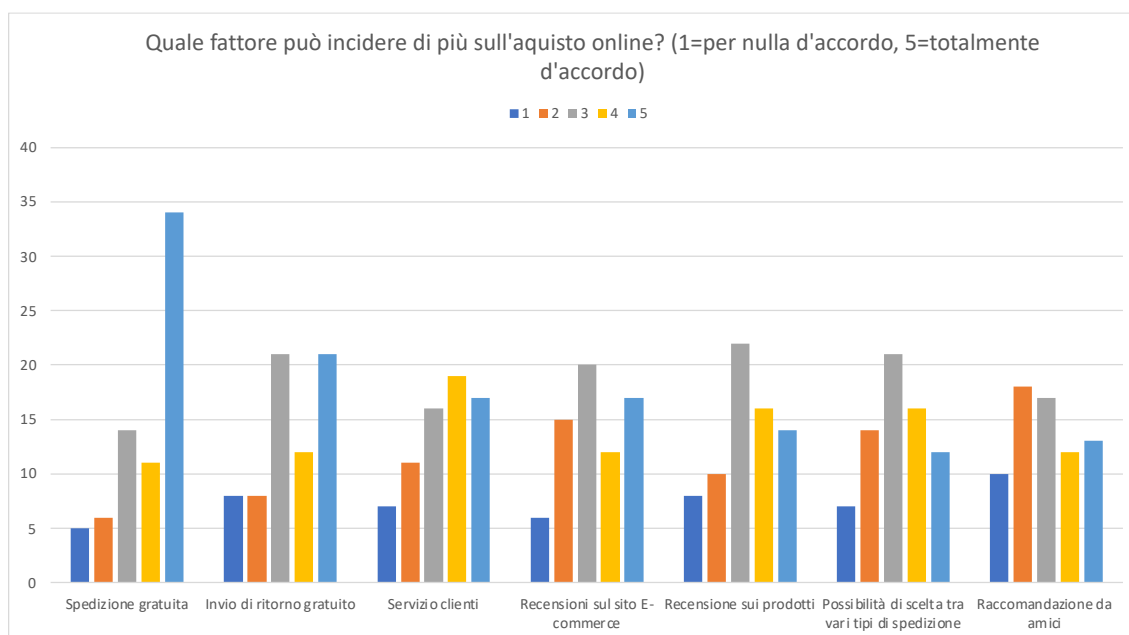


Figura 4.17, Grafico a barre sui fattori che incidono sull'acquisto online. Fonte: Elaborazione personale

Come emerso dalla figura 4.18, uno dei fattori che incentivano le persone ad utilizzare il servizio E-commerce è la spedizione gratuita. Dal questionario vengono inoltre rilevati i punti di forza favoriscono la consegna della spesa online, come riportate dalla figura 4.8 le più importanti sono: recapito della spesa lo stesso giorno dell'ordine o entro due ore, il tracciamento dell'ordine quando viene spedito dal supermercato o ipermercato ed infine il preavviso della ricezione degli invii. Non è relativamente importante per gli intervistati il recapito alla mattina o alla sera ed il recapito al termine desiderato.

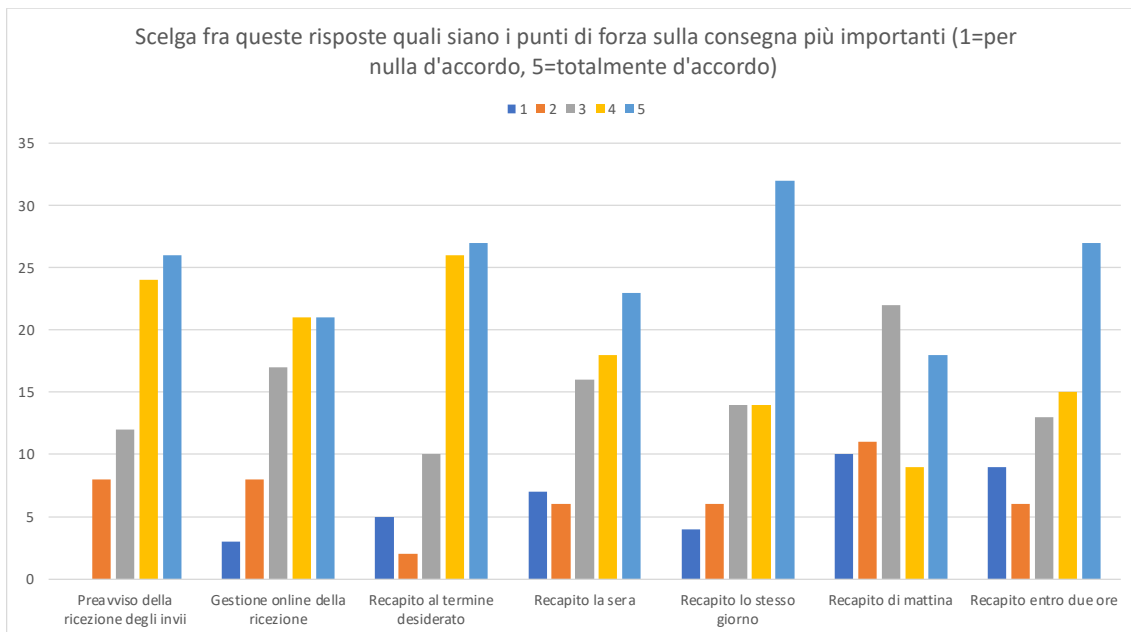


Figura 4.18, Grafico a barre sui punti di forza della consegna tramite E-commerce. Fonte: Elaborazione personale

I dati indicano che la maggior parte degli intervistati non utilizzava il servizio E-commerce offerto dai vari punti vendita prima dell'arrivo della pandemia Covid-19. Come riportato dalla figura 4.19, infatti il 18,1% degli intervistati usufruiva dei servizi online.

Utilizzavate l'e-commerce del punto vendita in questione prima della pandemia Covid 19?

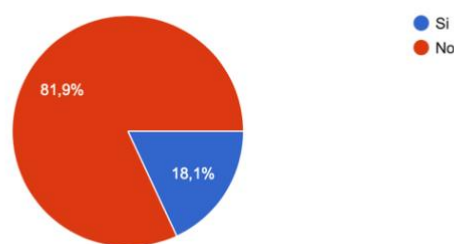


Figura 4.19, Grafico a torta sulla distribuzione di utilizzo del servizio E-commerce prima della Pandemia Covid-19. Fonte: Google questionare

È interessante comprendere quali siano i vantaggi che derivano dalla fruizione del servizio E-commerce. I dati raccolti ed illustrati dalla figura 4.20 dal questionario manifestano come la maggiore facilità d'uso del sito sia uno dei punti di forza offerto dal servizio online. In seconda posizione invece con il 31,9% di tasso di risposte confrontare prezzi, prodotti e promozioni è un fattore che favorisce l'utilizzo dell'E-commerce.

Che cosa la induce ad acquistare sui mercati virtuali?

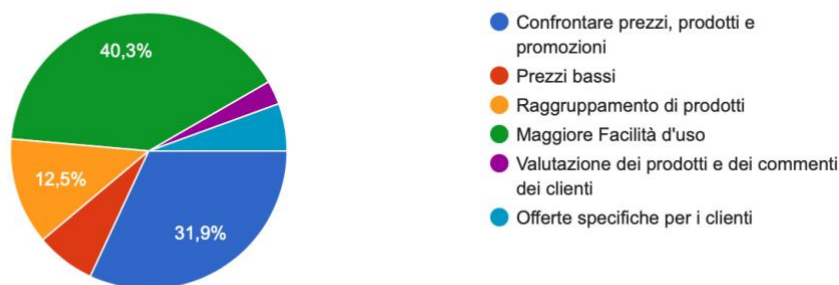


Figura 4.20, Grafico a torta sulla distribuzione delle motivazioni che inducono i consumatori ad utilizzare il mercato online. Fonte: Google questionare

Vengono proposte inoltre delle domande relative gli svantaggi del mercato online offerto dai supermercati ed ipermercati. Come riportato nel grafico 4.21, lo svantaggio principale del servizio online offerto dai supermercati ed ipermercati sono “la sostituzione dei prodotti senza il proprio consenso” e al secondo posto “la scadenza dei prodotti reperibili a breve termine”. Bisogna inoltre considerare che il 33,8% del tasso di risposta sottolinea come “la spedizione troppo costosa” sia uno svantaggio rilevante che si aggiunge alle due risposte citate prima.

Quali secondo lei possono essere gli svantaggi sull'acquisto online? (massimo 2 risposte)

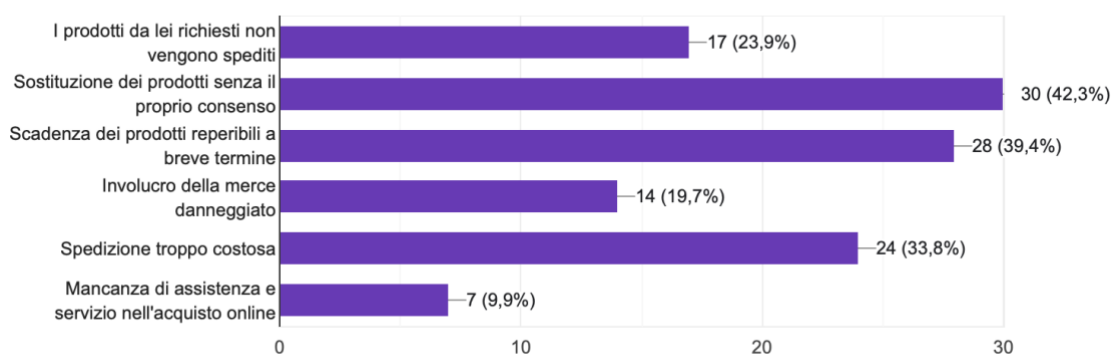


Figura 4.21, Grafico a barre sulla distribuzione delle risposte riguardante gli svantaggi sull'acquisto online. Fonte: Google questionare

Infine, viene posta un'ultima domanda, tramite scala Linkert, sulla soddisfazione da parte dei consumatori sugli acquisti online. Si rilevano dati positivi sull'utilizzo dell'E-

commerce, infatti come riportato dalla figura 4.22, la maggior parte degli intervistati precisamente il 62,5%, è soddisfatto del servizio offerto dai mercati online.

Infine, da una scala da 1 a 5 quanto vi ritenete soddisfatti dei vostri acquisti online?

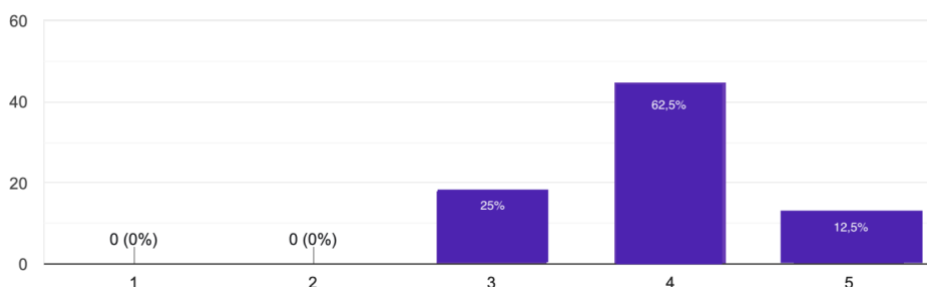


Figura 4.22, Grafico a barre sulla distribuzione della soddisfazione sugli acquisti online. Fonte: Google questionare

Vengono riportate inoltre le tabelle riguardanti la media e la deviazione standard dei dati raccolti tramite scala Linkert, tramite le domande “Perché acquista online?” e “Quale fattore può incidere di più sull’acquisto online” e “Scelga fra queste risposte quali siano i punti di forza sulla consegna più importanti”.

Dalla tabella 4.23, si può notare un valore della media simile tra le risposte: Consegna a casa, Semplicità, Covid e Velocità utilizzo. Al contrario la media delle risposte Confronto_offerte e Prezzi risultano avere valori più bassi. Ciò sta a significare che le persone tendono più ad utilizzare l’E-commerce per la consegna della merce a casa, per la semplicità del suo utilizzo, per la sicurezza sanitaria dovuto alla problematica Covid-19 ed infine per la velocità di utilizzo. Infine, esaminando la deviazione standard di tutte le variabili riportate nella figura 4.23, riportano quasi tutte lo stesso livello, tranne per il caso dei Prezzi dove la deviazione standard è pari a 1,13.

Variabili	Consegna_casa	Confronto_offerte	Semplicità	Covid	Prezzi	Velocità_utilizzo
Media	3,786	2,900	3,186	3,643	2,100	3,657
Deviazione Standard	1,371	1,405	1,254	1,533	1,131	1,306

Tabella 4.23, Tabella sulla media e deviazione standard (“Perché acquista online?”). Fonte: Elaborazione Personale

Per quando riguarda la seconda domanda: “Quale fattore può incidere di più sull’acquisto online”, si può notare dalle risposte riportate nella tabella 4.24, che la media delle risposte

è simile per tutte le risposte, lo stesso vale per la deviazione standard non c'è quindi un fattore che sembra incidere più degli altri nella scelta dell'acquisto online. È importante ricordare che la domanda è stata sviluppata con la stessa metodologia della precedente; dunque, con scala Linkert dove gli intervistati devono rispondere con una valutazione che tra 1 e 5.

Variabili	spedizione_gratuita	reso_gratuito	servizio_clienti	servizio_sito_E.commerce	Recenzioni	scelta_spedizione	raccomandazione_amici
Media	3,900	3,443	3,400	3,271	3,257	3,171	3,000
Deviazione standard	1,298	1,326	1,290	1,284	1,259	1,227	1,330

Tabella 4.24, Tabella media e deviazione standard sulla domanda “Quale fattore può incidere di più sull'acquisto online”. Fonte: Elaborazione Personale

Infine, l'ultima domanda la cui risposta è stata impostata come una scala Linkert è “Scelga fra queste risposte quali siano i punti di forza sulla consegna più importanti”. Le risposte sono riportate nella tabella 4.25, dove spiccano delle risposte con medie più elevate rispetto le altre, come la risposta “Preavviso_ricezione_invi”, “Recapito_al_termine_desiderato” e “Recapito_stesso_giorno”. Ciò sta ad evidenziare come le persone preferiscano un servizio di consegna a casa lo stesso giorno in cui viene effettuato l'ordine o scegliere quando farsi recapitare la spesa a casa.

Variabili	Preavviso_ricezione_invi	Gestione_online_ricezione	Recapito_al_termine_desiderato	Recapito_sera	Recapito_stesso_giorno	Recapito_mattina	Recapito_entro_due_ore
Media	3,971	3,700	3,971	3,629	3,914	3,200	3,643
Deviazione Standard	1,007	1,147	1,142	1,299	1,236	1,368	1,404

Tabella 4.25, Tabella media e deviazione standard sulla domanda “Scelga fra queste risposte quali siano i punti di forza sulla consegna più importanti”. Fonte: Elaborazione personale

Infine, si è effettuata un'analisi bivariata per evidenziare le eventuali correlazioni tra le risposte ottenute tramite scala Linkert dalle domande: “Scelga fra queste risposte quali siano i punti di forza sulla consegna più importanti”, “Quale fattore può incidere di più sull'acquisto online” e “Perché acquista online?”. Dal grafico si può notare una correlazione molto alta tra le variabili: “Confronto offerte” e “Semplicità di utilizzo” entrambe posizionate all'interno della batteria: “Perché acquista online”. Proseguendo si può notare una correlazione tra le variabili “Spedizione gratuita” e “Reso gratuito”,

all'interno della batteria "Quale fattore può incidere di più sull'acquisto online". Infine, si rileva un'ulteriore correlazione tra "Recapito entro due ore", "Entro lo stesso giorno" e "Recapito entro la mattinata", all'interno della batteria "Scelga fra queste risposte quali siano i punti di forza sulla consegna più importanti". Da questa correlazione emerge come le unità, desiderano una consegna da parte di supermercati ed ipermercati veloce ed all'interno dell'arco della giornata da quando si è effettuato l'ordine della spesa. Nella figura 4.26, viene riportato il grafico sulla correlazione tra le varie variabili raccolte tramite scala Likert.

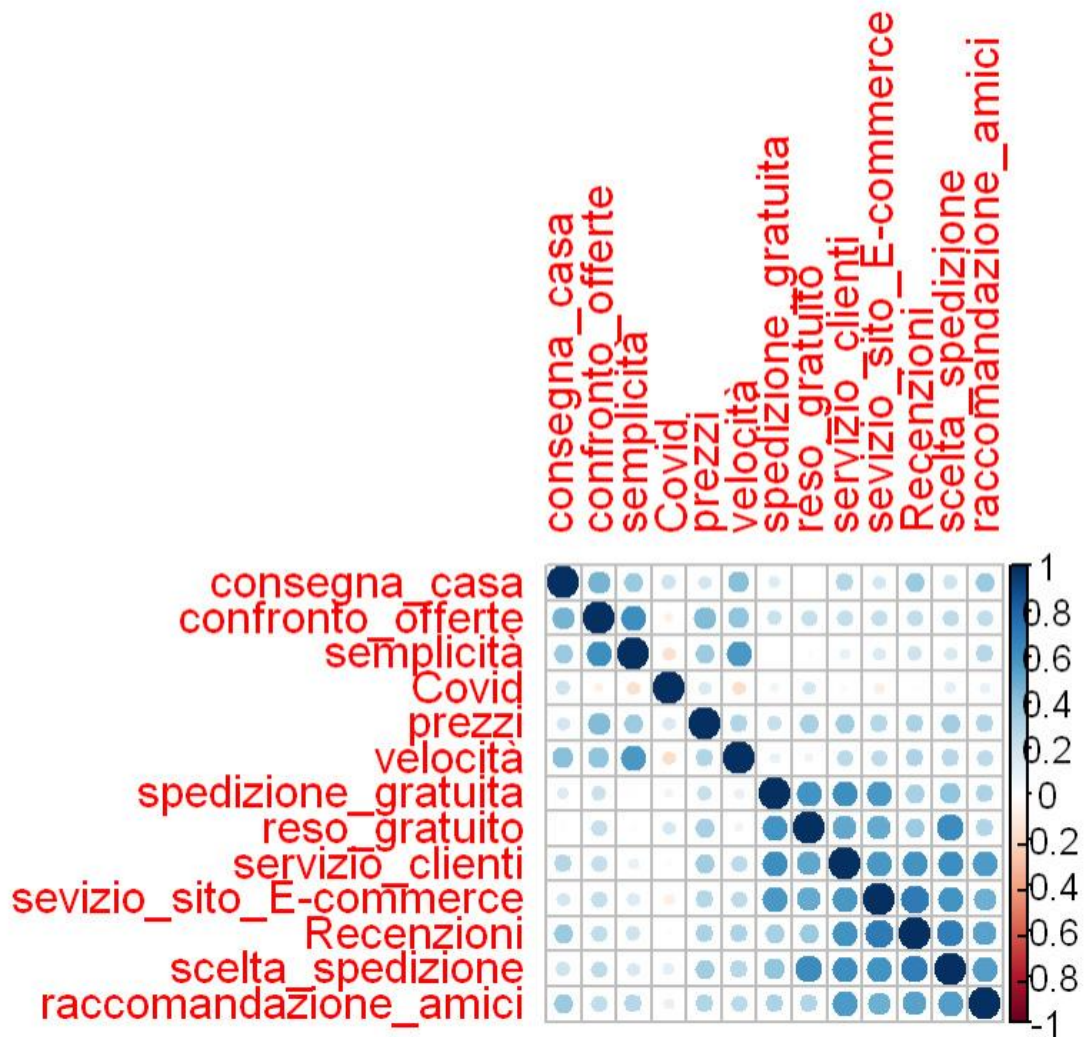


Figura 4.26, Grafico correlazione scala Likert. Fonte: Elaborazione personale

4.4: CARATTERISTICHE DEI CONSUMATORI CHE NON UTILIZZANO L'E-COMMERCE

Tramite il questionario “Indagine E-commerce in Italia”, sono stati raccolti dati relativi alle persone che non utilizzano l'E-commerce di ipermercati e supermercati. Nel campione totale quindi di 245 unità, sono state prelevate 175 unità. È importante dunque elencare quali siano le principali caratteristiche raccolte.

La maggior parte degli intervistati, hanno risposto che preferiscono acquistare direttamente i propri prodotti della spesa direttamente dal punto vendita, infatti il tasso di risposta è del 82,4%, quasi la totalità di tutte le unità che non effettuano acquisti online da supermercato o ipermercato. Viene rilevato inoltre che, alcuni supermercati, non forniscono il servizio E-commerce per effettuare la spesa online; infatti, la seconda percentuale con il maggior numero di risposte è “non forniscono il servizio e-commerce” con una percentuale di tasso di risposta del 23,3%, come riportato nella figura 4.27.

Perché non ha mai acquistato online da un supermercato? (massimo 3 risposte)

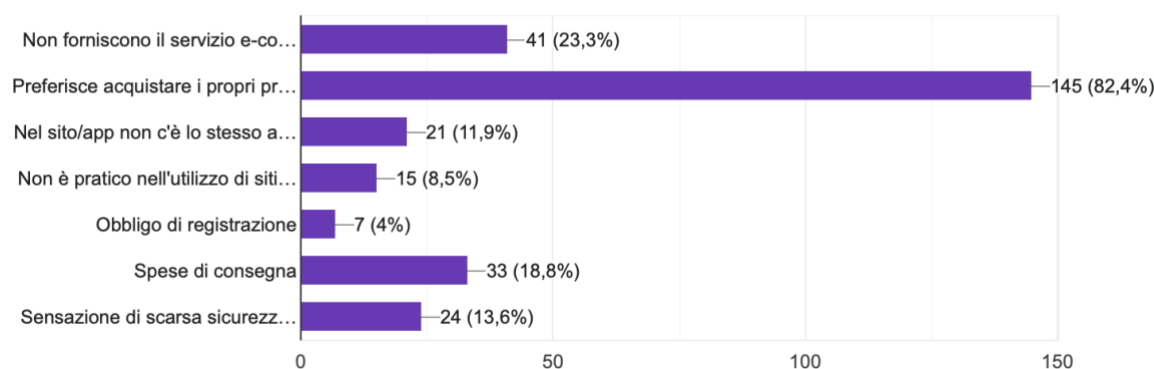


Figura 4.27, Grafico a barre in percentuale sulla motivazione per la quale i consumatori non acquistano online da supermercato. Fonte: Google questionare

Nel questionario vengono inoltre raccolte informazioni tramite la domanda a risposta aperta: “Cosa ne pensa della spesa online offerta da supermercati/ipermercati?” sulla percezione che le persone hanno sulla spesa online offerta da ipermercati e supermercati. Le risposte più esaustive raccolte sono:

- “Non c'è garanzia che arrivi tutto quello che si ordina e che arrivi nel formato o con le caratteristiche richieste, comodo, ma magari per categorie di persone che non possono recarsi nei punti vendita

fisici, gli altri ritengo che possano ritagliarsi un'ora la settimana per fare la spesa”

- “Penso che sia un buon servizio soprattutto per chi non ha tempo materiale per fare la spesa fisicamente come i genitori con i bimbi piccoli o chi che per qualsiasi motivo fisico non riesca”
- “È un ottimo servizio che può venire incontro a chi ha poco tempo, o difficoltà di qualche tipo, per effettuare la spesa”
- “Possono essere utili per alcune nicchie di mercato, ma non offrono la funzione (di commercializzazione) psico-sociale, inoltre le consegne a domicilio, se non gestite tramite micro-hubs e mezzi di trasporto alternativi, causano un incremento delle emissioni di CO₂”
- “Buon servizio, ma non è possibile scegliere la qualità/caratteristiche come se andassi fisicamente (ed esempio sulla frutta fisicamente posso fare attenzione a prendere quella migliore o sugli affettati fare richieste specifiche di taglio)”
- “Per prodotti a lunga conservazione e di media grande dimensione è un'idea utile l'acquisto online (es. acqua, latte, detersivi, carta igienica, prodotti in lattine etc.). Per quanto riguarda frutta, verdura, prodotti da banco, pasta e generi alimentari ad uso quotidiano invece scelgo di acquistare direttamente al supermercato”

Dalle risposte raccolte emerge che l'E-commerce anche se non utilizzato è ritenuto un buon servizio, per esempio, per coloro che hanno poco tempo a disposizione, ma che ha ancora dei limiti per esempio: prodotti da banco o frutta ad uso quotidiano è preferibile acquistarli direttamente dal punto vendita. Viene rilevata una problematica ambientale come le immissioni di CO₂ per il trasporto della spesa nella propria abitazione.

Viene chiesto agli intervistati se in un futuro adotteranno il servizio online offerto dai supermercati, come riporta la figura 4.28, più della metà, quindi il 61,1% vorrà usufruire del servizio online.

Secondo lei acquisterà online dai supermercati in futuro?

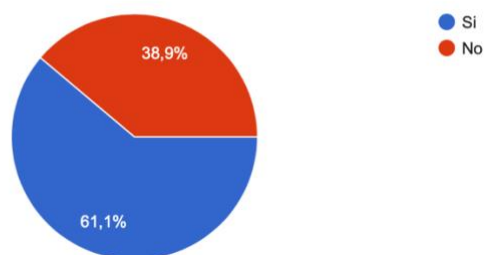


Figura 4.28, Grafico a torta sulla distribuzione di coloro che effettueranno in futuro un acquisto online. Fonte: Google questionare

Infine, viene posta una domanda su quale siano i fattori che possono incidere le persone ad utilizzare l'E-commerce dei supermercati ed ipermercati per effettuare la propria spesa. Dalla figura 4.29 si nota che il dato con il maggior numero di risposte è “promozioni e sconti solo su acquisto online” con 101 risposte totali. Con un numero lievemente inferiore “la spedizione gratuita” è ritenuta anche essa uno dei fattori che possono incentivare le persone ad effettuare la spesa online, come analizzato precedentemente sul capitolo 4.3 sulle caratteristiche di coloro che acquistano online.

Quale opzione sceglierebbe per convincerla ad utilizzare la spesa online? (massimo 2 risposte)

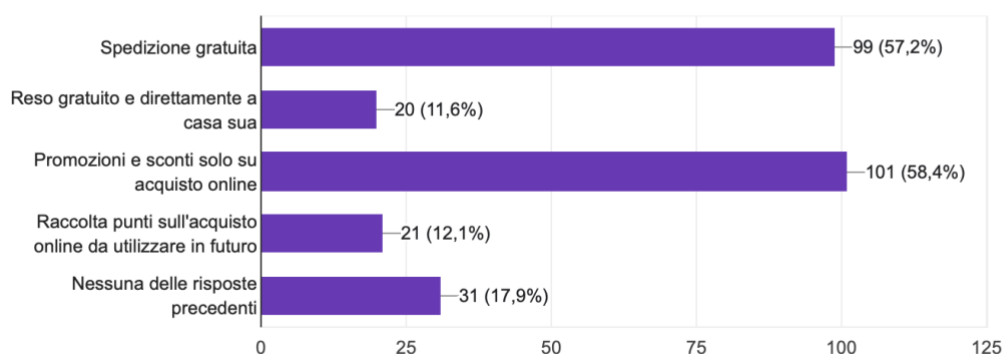


Figura 4.29, Grafico a barre sulle opzioni che possono convincere coloro che non utilizzano la spesa online ad usufruire di quest'ultima. Fonte: Google questionare

4.5: APPLICAZIONE DEI METODI STATISTICI AI DATI

Nei capitoli seguenti verranno presentate le analisi effettuate sul questionario oggetto di tesi.

4.5.1: CLUSTER ANALYSIS SUI SOGGETTI CHE ACQUISTANO ONLINE

Riguardo le persone che effettuano acquisti online si è scelto di procedere con la ricerca di gruppi che rappresentassero comportamenti di acquisto simili, sulla base delle scale Linkert raccolte dal questionario dalle domande “Perché acquista online?” e “Quale fattore può incidere di più sull’acquisto online” e “Scelga quale risposte siano i punti di forza sulla consegna più importanti”. Prima di tutto si è anche qui proceduto con un pre-processing dei dati. Avendo osservato un valore mancante solo per la variabile reso gratuito si è scelto di procedere imputando la mediana della stessa, evitando quindi di eliminare l’osservazione vista la già esigua numerosità del campione. Si è scelto quindi di standardizzare le variabili per preparare i dati alla ricerca dei cluster. Per la scelta dei metodi con cui effettuare l’analisi si è scelto di affidarsi sia al K-means che ad un clustering gerarchico agglomerativo in modo da compararne i risultati e prendere decisioni sulla divisione dei dati in gruppi con maggiore sicurezza.

Per quanto riguarda l’algoritmo K-means si è optato per selezionare il numero di cluster k da considerare tramite il metodo Elbow che, unito al principio di parsimonia, ha portato all’individuazione di tre cluster, come riportato nella figura 4.30.

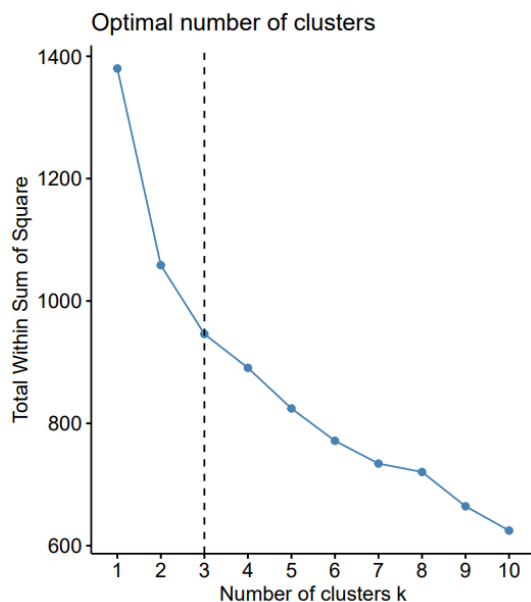


Figura 4.30, Rappresentazione del grafico Elbow. Fonte: Elaborazione personale

Per quanto riguarda il clustering gerarchico il dendrogramma evidenzia la netta presenza di tre gruppi anche se la scelta di 5 o 6 gruppi poteva sembrare ragionevole. Tuttavia,

tenendo sempre a mente il principio di parsimonia, anche in questo caso $k=3$ sembra sia la scelta più ragionevole. Viene rappresentato nella figura 4.31 il dendrogramma.

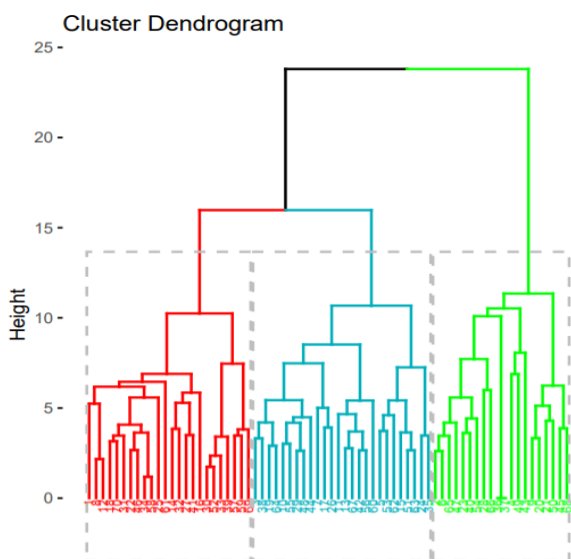


Figura 4.31, Rappresentazione del dendrogramma. Fonte: Elaborazione personale

Viene svolta successivamente una valutazione in base alle medie e deviazioni standard per ciascun gruppo sulle variabili che hanno permesso il raggruppamento dei tre cluster.

CLUSTER 1

Il cluster 1 rappresenta i clienti “Meno fedeli”. Si può notare infatti dalla tabella 4.32, sulle variabili utilizzate per il processo di cluster che le valutazioni medie raccolte dalle opinioni degli intervistati durante la compilazione del questionario, non sono alte. Si può dunque ipotizzare che il servizio E-commerce non sia così rilevante per le unità che formano il cluster 1.

Motivazione spesa online	consegna_casa	confronto_offerte	semplicità	Covid	prezzi	velocità	
Media	2,889	2,556	3,222	3,056	1,444	3,389	
Deviazione Standard	1,278	1,381	1,309	1,626	0,705	1,501	
Incentivo dell'E-commerce	spedizione_gratuita	reso_gratuito	servizio_clienti	servizio_sito_E-commerce	Recensioni	scelta_spedizione	raccomandazione_amici
Media	2,389	2,278	2,000	2,278	2,333	2,167	1,889
Deviazione Standard	1,243	1,320	0,970	1,179	0,970	0,985	0,676
Fattori di successo consegne online	Preavviso_ricezione_invii	Gestione_online_ricezione	Recapito_al_termine_desiderato	Recapito_sera	Recapito_stesso_giorno	Recapito_mattina	Recapito_entro_due_ore
Media	3,111	2,500	2,833	2,222	3,000	2,278	2,833
Deviazione Standard	1,132	0,857	1,295	1,060	1,495	1,274	1,618

Tabella 4.32, Tabella sulla media e deviazione standard delle variabili dei cluster. Fonte: Elaborazione Personale

CLUSTER 2

Il cluster 2 rappresenta i clienti “Più fedeli”. La motivazione della nomina del cluster 2 è data dalle medie che sono nettamente più alte rispetto il cluster 1. Se nello specifico vengono descritte le medie delle tabelle riportate nella tabella 4.33, si rende noto che la media è al di sopra

di 4, ricordando che il massimo della valutazione proposta alle unità era 5. Dunque, si può confermare che il cluster 2 sia motivato nell'utilizzare l'E-commerce offerto dai supermercati ed ipermercati.

Motivazione spesa online	consegna_casa	confronto_offerte	semplicità	Covid	prezzi	velocità	
Media	4,455	4,000	4,045	3,591	3,000	4,455	
Deviazione Standard	1,057	1,234	1,090	1,593	1,234	0,858	
Incentivo dell'E-commerce	spedizione_gratuita	reso_gratuito	servizio_clienti	servizio_sito_E-commerce	Recensioni	scelta_spedizione	raccomandazione_amici
Media	4,682	4,182	4,364	4,091	4,045	4,091	3,955
Deviazione Standard	0,568	1,053	0,953	0,971	0,950	0,921	1,090
Fattori di successo consegne online	Preavviso_ricezione_invi	Gestione_online_ricezione	Recapito_al_termine_desiderato	Recapito_sera	Recapito_stesso_gi	Recapito_mattina	Recapito_entro_due_ore
Media	4,636	4,636	4,591	4,364	4,636	4,364	4,364
Deviazione Standard	0,581	0,581	0,590	1,002	0,727	0,902	1,217

Tabella 4.33, Tabella sulla media e deviazione standard delle variabili dei cluster. Fonte: Elaborazione Personale

CLUSTER 3

Il cluster 3, rappresenta i clienti “Fedeli incerti”. Il nome scelto per etichettare questo gruppo è influenzato da una media della variabile Covid, rispetto le altre, ciò: può significare che l'utilizzo del mercato online offerto da supermercati ed ipermercati sia stato incentivato principalmente dalla pandemia Covid in atto. Ciò non giustifica che in futuro se quest'ultima terminerà, le unità che compongono questo cluster utilizzeranno ancora il servizio online. Confrontando i dati raccolti dai cluster citati prima, le valutazioni raccolte dalle unità che compongono questo cluster si discostano di molto da quelle evidenziate nel cluster 2. È da evidenziare inoltre che le medie riportate sulla tabella 4.34 sono comunque maggiori rispetto le medie del cluster 1.

Motivazione spesa online	consegna_casa	confronto_offerte	semplicità	Covid	prezzi	velocità	
Media	3,833	2,300	2,533	4,033	1,833	3,233	
Deviazione Standard	1,367	1,055	0,937	1,351	0,834	1,223	
Incentivo dell'E-commerce	spedizione_gratuita	reso_gratuito	servizio_clienti	servizio_sito_E-commerce	Recensioni	scelta_spedizione	raccomandazione_amici
Media	4,233	3,600	3,533	3,267	3,233	3,100	2,967
Deviazione Standard	0,935	1,037	0,900	1,143	1,251	1,062	1,273
Fattori di successo consegne online	Preavviso_ricezione_invi	Gestione_online_ricezione	Recapito_al_termine_desiderato	Recapito_sera	Recapito_stesso_giorno	Recapito_mattina	Recapito_entro_due_ore
Media	4,000	3,733	4,200	3,933	3,933	2,900	3,600
Deviazione Standard	0,788	0,944	0,847	0,944	1,015	1,125	1,133

Tabella 4.34, Tabella sulla media e deviazione standard delle variabili dei cluster. Fonte: Elaborazione Personale

Dopo aver effettuato una possibile segmentazione della clientela si passa quindi allo studio della composizione dei cluster in termini di variabili sociodemografiche e di comportamento d'acquisto.

Per quanto riguarda il sesso si può vedere dalla tabella riportata nella tabella 4.35 di seguito come per il gruppo tre la percentuale di uomini e donne sia la stessa mentre per i primi due gruppi vi è un 63% circa di donne, questo è dovuto dalla presenza maggiore di risposta del genere femminile nel campione come riportato nell'analisi univariata eseguita nei capitoli precedenti.

	Meno fedele	Più fedele	Fedeli incerti
Donna	61,11%	63,64%	50%
Uomo	38,89%	36,36%	50%

Tabella 4.35, Tabella sulla segmentazione di genere per i vari cluster. Fonte: Elaborazione personale

Nella tabella 4.36 viene rappresentata la tabella sulla distribuzione tra i vari cluster delle unità su alcune delle variabili all'interno della batteria "TITOLO DI STUDIO". Si prendono in esame soprattutto i dati relativi ai laureati e diplomati in quanto gli altri sono poco numerosi e quindi eventuali associazioni non sarebbero troppo attendibili.

Dalla tabella si può notare come all'interno del Cluster 1 non siano presenti dottori di ricerca o che abbiano effettuato una specializzazione post-laurea. Le unità si distribuiscono maggiormente all'interno delle variabili diploma o qualifica di scuola superiore e Laurea. Per quanto concerne invece al Cluster 2 dalla tabella 4.36 si può notare che molte unità sono presenti all'interno della variabile diploma o qualifica di scuola superiore. Non è presente nessuna unità invece nella variabile licenza scuola media inferiore. Infine, per il Cluster 3 come per il Cluster 1 le unità si distribuiscono maggiormente nelle variabili diploma o qualifica di scuola superiore e Laurea. Rispetto agli altri due cluster, è presente una maggiore presenza sulla variabile dottore di ricerca o specializzazione post-laurea.

	Diploma superiore	Dottorato o Specializzazione	Laurea	Licenza media
Meno fedele	9	0	8	1
Più fedele	17	1	4	0
Fedeli incerti	13	3	12	2

Tabella 4.36, Tabella di distribuzione sulla batteria TITOLO STUDIO confrontata con i cluster.

Fonte: Elaborazione personale.

In aggiunta vengono rilevate le percentuali di frequenza sulla distribuzione delle unità sulla batteria TITOLO DI STUDIO, come riportato dalla tabella 4.37. Si può notare come vi siano differenze in termini di gruppo d'appartenenza; infatti, se il cluster 2 è caratterizzato soprattutto da diplomati e solo il 16% dei laureati vi appartiene mentre questi sono per il 50% nel gruppo 3 e per il rimanente nel gruppo 1.

	Diploma superiore	Dottorato o Specializzazione	Laurea	Licenza media
Meno fedele	23,1%	0,0%	33,3%	33,3%
Più fedele	43,6%	25,0%	16,7%	0,0%
Fedeli incerti	33,3%	75,0%	75,0%	66,7%

Tabella 4.37, Tabella di distribuzione percentuale sulla batteria TITOLO STUDIO confrontata con i cluster. Fonte: Elaborazione personale.

Per quanto riguarda le variabili sociodemografiche sullo stato lavorativo della batteria LAVORO, delle unità appartenenti al cluster si può notare dalla figura 4.38, avendo la maggior parte dei soggetti che cade sotto la categoria "Si" (attualmente lavorano) occorre essere cauti nelle interpretazioni. Tuttavia, si può vedere come per i non lavoratori il Cluster 1 è il più presente, mentre la variabile "Si", è la meno frequente rispetto gli altri Cluster.

	Disoccupato	No	Si
Meno fedele	1	5	12
Più fedele	0	4	18
Fedeli incerti	0	5	25

Tabella 4.38, Tabella di distribuzione sulla batteria LAVORO confrontata con i cluster. Fonte: Elaborazione personale.

Continuando le analisi sulle variabili riguardanti il comportamento di acquisto, per la batteria FREQUENZA_ ONLINE_SUPERMERCATO, si mostra nella tabella 4.39, effettuando le medesime precisazioni del caso precedente, si può vedere che effettuare acquisti online da supermercato ed ipermercato con frequenza quotidiana e settimanale vi sia rispettivamente prevalenza nel Cluster 2 e 3. Infine, nel Cluster 2 e 3 vi è una presenza maggiore di risposta sulla frequenza di spesa online mensile rispetto il Cluster

1. Non viene considerata la variabile “Mai” che era presente sul Cluster dei “Fedeli incerti”, in quanto è impossibile da valutare dato che si parla di unità che acquistano online.

	Annuale	Mensile	Settimanale	Giornaliero
Meno fedele	1	1	12	4
Più fedele	1	4	14	3
Fedeli incerti	0	3	25	2

Tabella 4.39, Tabella di distribuzione sulla batteria FREQUENZA_ONLINE_SUPERMERCATO confrontata con i cluster. Fonte: Elaborazione personale.

Viene eseguita un’ulteriore analisi sul ritiro della propria spesa online. Si nota dalla tabella 4.40, che vengono considerate solo due modalità: direttamente a casa e direttamente in punto vendita. Si osservano dunque questi risultati. Per il Cluster 3 si ha una netta prevalenza di persone che si fanno consegnare la spesa a casa e questa prevalenza è meno netta per gli altri due gruppi. Infine, nel cluster 2 è presente una maggiore risposta rispetto gli altri due Cluster sul ritiro della spesa già preparata nel punto vendita.

È importante sottolineare che la risposta Locker Box è stata eliminata per la bassa frequenza, evidenziando come metodi alternativi della consegna della merce non siano ancora entrati nella quotidianità di utilizzo.

	Direttamente a casa	Direttamente in punto vendita
Meno fedele	10	7
Più fedele	14	8
Fedeli incerti	25	5

Tabella 4.40, Tabella di distribuzione sulla batteria FREQUENZA_ONLINE_SUPERMERCATO confrontata con i cluster. Fonte: Elaborazione personale.

Per quanto riguarda la distanza dal supermercato non emergono grosse differenze tra i cluster, perché la maggior parte delle osservazioni ricadono nella fascia tra 5 e 10 chilometri, tranne per il gruppo 1 e 3 dove quasi il 35% delle osservazioni abita ad una distanza dal supermercato tra i 10 e 20 chilometri come riportato dalla tabella 4.41.

Distanza	Meno fedele	Più fedele	Fedeli incerti
5-10 km	0,61	0,64	0,53
10-20 km	0,39	0,18	0,33
20-30 km	0,00	0,14	0,07
più di 40 km	0,00	0,05	0,07

Tabella 4.41, Tabella sulla distanza dal supermercato o ipermercato dei cluster. Fonte: Elaborazione personale

Tuttavia, dal grafico figura 4.42 seguente dove i pallini rappresentano la media delle distanze nei gruppi, si può vedere come il primo gruppo sia decisamente il più vicino al supermercato mentre il terzo sia leggermente più distante in media del secondo.

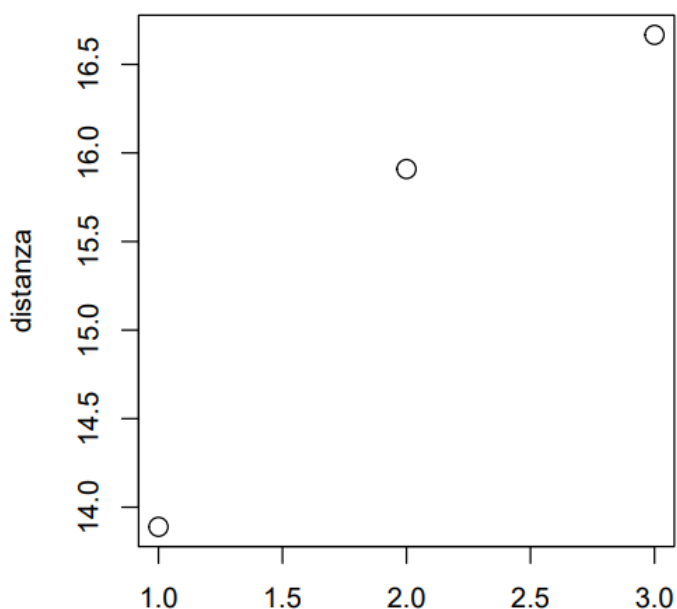


Figura 4.42, Plot sulla media distanza dal supermercato dei Cluster. Fonte: Elaborazione personale

La tabella 4.43, riporta i dati ottenuti per costruire il plot sulla media della distanza da supermercato o ipermercato.

	Meno fedele	Più fedele	Fedele incerto
Distanza media	13,889	15,909	16,667

Tabella 4.43, Tabella sulle medie della distanza da supermercato o ipermercato. Fonte: Elaborazione personale

Di seguito invece nella figura 4.44 si possono vedere i boxplot per l'età per ciascun gruppo. Si può notare come il terzo gruppo sia caratterizzato da persone con un range di età ampio che va dai 24 ai 34 circa e con un paio di osservazioni anomale attorno ai 50 anni. Il primo gruppo sembra presentare uno scarto interquantilico meno ampio e con il 50 per cento delle osservazioni tra i 24 e 32 anni circa, presentando tuttavia tre outlier tra i 50 e i 60 anni. Il secondo gruppo sembra il più regolare in termini di numero di osservazioni anomale con un unico outlier rappresentato da una persona sopra i 70 anni. Il range di età della maggior parte delle osservazioni sembra più ampio rispetto al primo gruppo con il 50 per cento delle osservazioni tra circa i 22 e i 34 anni. Da notare come in tutti i gruppi l'età presenta una distribuzione non simmetrica, con coda a destra.

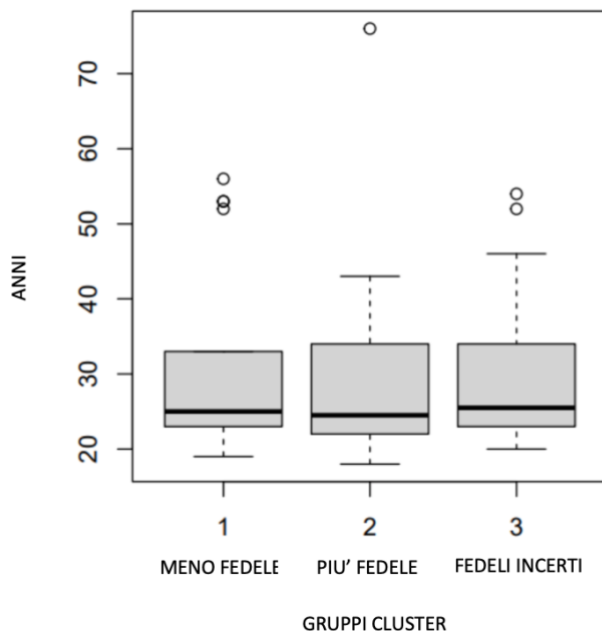


Figura 4.44, Rappresentazione dei boxplot sull'età dei cluster, Fonte: Elaborazione personale

Vengono poi confrontati i gruppi in termini di ragioni che portano a preferire l'utilizzo della spesa online, come rappresentato dalla figura 4.45. Il primo gruppo sembra prevalentemente definito da persone che desiderano confrontare prezzi prodotti e promozioni, motivo che tuttavia risulta abbastanza presente in tutti e 3 i gruppi. Il secondo gruppo è il più eterogeneo in termini di motivazioni tanto che nessuna ragione sembra prevalere sulle altre. Infine, il terzo gruppo invece è definito per il 50 per cento da persone che apprezzano la maggiore facilità di utilizzo, aspetto ancora più interessante se si pensa che questo era anche il cluster con età mediana maggiore.

	Meno fedele	Più fedele	Fedeli incerti
Confronto prezzi e prodotti	55,56%	22,72%	26,67%
Maggiore facilità d'uso	38,89%	22,72%	50,00%
Offerte specifiche per i clienti	0,00%	18,18%	0,00%
Prezzi bassi	0,00%	18,18%	3,33%
Raggruppamento prodotti	5,00%	13,64%	16,67%
Valutazione dei prodotti e commenti dei clienti	0,00%	4,55%	3,00%

Figura 4.45, Tabella sulla percentuale di risposta vantaggi acquisto online da supermercato o ipermercato dei cluster, Fonte: Elaborazione Personale

I dati raccolti possono essere confrontati con altri elaborati. Si è dunque cercato in letteratura, ricerche simili a quelle effettuate in questa tesi. È emerso da uno studio effettuato sull'evoluzione dell'E-commerce durante la pandemia Covid-19 in Portogallo⁴⁹ che i risultati descritti sono simili. Dal campione di 358 unità raccolto dai ricercatori dell'articolo *"Evolution of the Online Grocery Shopping Experience during the COVID-19 Pandemic: Empiric Study from Portugal"* è emerso che consumatori più giovani influenzano positivamente l'esperienza di acquisto di generi alimentari online durante la pandemia di COVID-19. Anche nelle analisi effettuate in questo capitolo si può comprendere come i giovani siano i principali utilizzatori del servizio online. Un'altra tematica discussa dai ricercatori portoghesi sono le principali determinanti che portano all'acquisto di generi alimentari online che risulta essere: la facilità di ordinazione e la velocità di consegna della merce.

Come riportato dalla Cluster analysis effettuata in questa tesi emerge come la semplicità d'uso dell'E-commerce offerto da supermercati ed ipermercati sia una delle motivazioni di utilizzo principale del servizio, come riportato nella tabella 4.50, ma non solo, anche la variabile "Confrontare prezzi, prodotti e promozioni", ha un peso rilevante per tutti e tre i cluster.

⁴⁹ GOMES SOFIA, JOAO M. LOPEZ. "Evolution of the Online Grocery Shopping Experience during the COVID-19 Pandemic: Empiric Study from Portugal" *Journal of Theoretical and Applied Electronic Commerce Research*, 2022

4.5.2 ANALISI REGRESSIONE LOGISTICA APPLICATA AI NON ACQUIRENTI

Per effettuare l'analisi dei dati sul campione che non effettua acquisto online, viene utilizzato un modello di regressione, nello specifico il modello logit. Il vantaggio di questo modello è la possibilità di confrontare una variabile binaria con variabili di tipo qualitativo. È stato necessario procedere con un pre-processing dei dati al fine di diminuire eventuali variabili non informative per l'analisi e trattare i dati mancanti. La prima operazione è stata quella di eliminare le variabili che risultavano costanti, in quanto non informative. Viene affrontata dunque la questione NA (dati mancanti), vedendo prima di tutto come si distribuivano nelle variabili osservate. Si è osservato che la variabile SPEDIZIONE_ACQUISTI_ONLINE (che nel dataset rappresentava dove i clienti compravano prodotti online in generale) presentava molti valori mancanti (fatto compatibile con il campione che si sta analizzando, supponendo che chi non acquista online in supermercato o ipermercato sia comunque un soggetto che non acquista beni online a prescindere). Si è scelto quindi di procedere con l'eliminazione di quest'ultima variabile. Per quanto riguarda gli altri valori mancanti invece si è scelto di utilizzare tutte e sole le osservazioni complete.

Per la stima del modello di regressione logistica della variabile "ACQUISTO ONLINE FUTURO" del campione dei non acquirenti online, si è dunque definito un modello per la selezione delle variabili.

Dato il numero di possibili variabili esplicative, si è deciso di ricercare un modello ottimale con un procedimento iterativo noto come selezione stepwise. Questo approccio parte dalla stima di un modello che considera tutti i regressori ed elimina dal modello una variabile alla volta, scegliendo passo per passo quale sia il regressore migliore da considerare nel modello sulla base di alcuni indicatori di bontà.

Si è scelto di utilizzare il criterio BIC, concludendo che il modello migliore è quello che considera un'unica variabile "ANNI". È stato utilizzato il BIC e non un altro criterio quale l'AIC nella speranza di ottenere un modello più parsimonioso e quindi di più semplice interpretazione.

I risultati del modello stimato sono riportati di seguito nella tabella 4.46.

Coefficienti	Stima	Errore standard	Z-test	P-value
Intercetta	1,4578	0,4254	3,4270	0,0006
Anni	-0,0324	0,0133	-2,4310	0,0151

Tabella 4.46, Risultati modello regressione logistica con variabile ANNI. Fonte: Elaborazione personale

Come si può vedere dalla tabella l'effetto della variabile ANNI risulta significativamente diverso da zero si è scelto quindi di interpretarlo adeguatamente applicando alle stime riportate sopra nella la funzione inversa della logistica. Si è ottenuto una intercetta pari a 0.8111 e la variabile ANNI 0.4919⁵⁰. Si deduce che al crescere dell'età la probabilità di essere propensi ad acquisti online in futuro decresca. Quindi è importante sottolineare come gli acquisti online siano effettuati nella maggior parte dei casi da individui giovani. In particolare, ad ogni aumento unitario degli anni (rispetto al livello medio rappresentato dall'intercetta) ci si aspetta che la probabilità di essere propensi ad acquisti online futuri si dimezzi.

La formula ottenuta dall'utilizzo della variabile "ANNI" per la stima del modello della variabile "ACQUISTO ONLINE FUTURO" è riportata di seguito.

$$\log \left(\frac{p(x)}{1 - p(x)} \right) = 0.889 + 0.491 * ANNI + \epsilon$$

Durante le analisi dei dati, si è tentato di aggiungere una nuova variabile al modello di regressione logistica ovvero la variabile NO_ACQUISTI_ONLINE dove i rispondenti potevano scegliere un massimo di tre risposte tra: preferisce acquistare i propri prodotti direttamente al punto vendita, spese di consegna, sensazione di scarsa sicurezza nell'acquisto online e non forniscono il servizio E-commerce nel Paese. Tuttavia, per semplificare l'analisi data la natura a scelta multipla della domanda nel questionario, si è tenuta in considerazione per semplicità solo la prima delle risposte disponibili che risulta un numero nettamente maggiore rispetto le altre. Si è dunque costruita una variabile dummy che rappresenta la modalità: "Preferisce acquistare i propri prodotti direttamente al punto vendita" accorpendo le altre risposte, sotto la categoria "Altro".

Dall'analisi si può osservare che al netto degli anni l'effetto (sulla probabilità di aderire in futuro all'online) dell'appartenenza al gruppo che ha dato come risposta "Preferisce acquistare i propri prodotti direttamente al punto vendita" è negativo. Dunque, all'aumentare della variabile ACQUISTO_ONLINE_FUTURO, diminuirà la variabile NO_ACQUISTO_ONLINE, dove per la maggioranza dell'osservazione ha scelto come

⁵⁰ Risultati ottenuti da R-studio, applicando la funzione inversa logistica

risposta “Preferisce acquistare i propri prodotti direttamente al punto vendita”. Vengono riportati i risultati nella tabella 4.47.

Coefficienti	Stima	Errore standard	Z-test	P-value
Intercetta	2,0798	0,4810	4,3240	1,53E-05
Anni	-0,0364	0,0130	-2,7880	0,0053
Preferenze sull' acquisto	-0,7803	0,3695	-2,1120	0,0347

Tabella 4.47, Risultati regressione logistica con variabile ANNI e NO_ACQUISTO_ONLINE, Fonte: Elaborazione Personale

La formula ottenuta dalla variabile ANNI e NO_ACQUISTI_ONLINE, per la stima della variabile ACQUISTO_ONLINE_FUTURO è la seguente:

$$\log\left(\frac{p(x)}{1-p(x)}\right) = 2.080 - 0.036 * ANNI - 0.780 * PUNTOVENDITA + \epsilon$$

Infine, come analizzato con la variabile ANNI si è utilizzata la stima inversa della logistica. Si è ottenuto una intercetta pari a 0,888, la variabile ANNI pari a 0,4909 ed infine la variabile NO_ACQUISTI_ONLINE FUTURO pari a 0.314⁵¹. Coloro che preferiscono acquistare i propri prodotti direttamente al punto vendita la probabilità è più bassa di circa il 70 per cento.

4.6: ALCUNE CONSIDERAZIONI FINALI

Anche se il campione oggetto di analisi è stato auto selezionato e la presente analisi consente di avere un’idea su quelle che possono essere i fattori rilevanti per l’utilizzo di servizi E-commerce dei supermercati o ipermercati.

Innanzitutto, è importante evidenziare come il modello di regressione logistica si adatti bene alle variabili scelte per l’analisi. Tramite il modello di regressione logistica sul campione che non effettua acquisti online, viene evidenziato come all’aumentare del tasso di età si abbia una minore probabilità utilizzo futuro del servizio E-commerce offerto da supermercati ed ipermercati, evidenziando come le persone più “anziane” non siano propense ad utilizzare il servizio online in futuro.

Si è evidenziato inoltre come al crescere della variabile “Preferisce acquistare i propri prodotti direttamente al punto vendita”, difficilmente le unità adotteranno il servizio online offerto da supermercati ed ipermercati in futuro.

⁵¹ Risultati ottenuti tramite R-studio, applicando la funzione della logistica inversa

Nella Cluster analysis effettuata sul campione che effettua acquisti da supermercato ed ipermercato emergono 3 cluster, dove si racchiudono individui con comportamenti di acquisto simili, sulla base delle risposte alle domande “Perché acquista online?” e “Quale fattore può incidere di più sull’acquisto online” e “Scelga quale risposte siano i punti di forza sulla consegna più importanti”. È importante ricordare che la Cluster analysis è utile per effettuare analisi di marketing e segmentare gli individui secondo comportamento di acquisto simili, ed analizzare sotto alcuni aspetti rilevanti per il rivenditore. L’analisi effettuata sul comportamento di acquisto da parte dei consumatori, fa emergere che in prevalenza si preferisce utilizzare il ritiro della spesa a casa e non direttamente in punto vendita. Ciò significa dunque che le unità oggetto di analisi, hanno preferenza nell’utilizzare il servizio comune offerto dall’E-commerce di supermercati ed ipermercati (ritiro della spesa a casa), rispetto altre modalità.

Viene introdotta infine una variabile riguardante i vantaggi derivati dall’utilizzo del servizio online offerto da ipermercati e supermercati, rivelando che prevalentemente per i tre gruppi si utilizza l’E-commerce anche per confrontare prezzi prodotti e promozioni. Emerge che il secondo gruppo “Più fedeli” è il più eterogeneo in termini di motivazioni tanto che nessuna ragione sembra prevalere sulle altre. Il terzo gruppo invece è definito per il 50 per cento da persone che apprezzano la maggiore facilità di utilizzo.

In conclusione, dunque è importante notare come tramite la Cluster analysis si sia potuto effettuare una segmentazione del comportamento di acquisto da parte degli utenti e notare come all’interno dei gruppi ci sia omogeneità di risposta.

CONCLUSIONI

L'e-commerce si sta affermando come uno dei canali di vendita più importanti per la crescita delle aziende.

Le enormi opportunità che la tecnologia digitale può offrire sono in continua evoluzione, tant'è che il mercato digitale mette a disposizione servizi, che mettono addirittura in discussione il più delle volte i mercati tradizionali.

Indubbiamente, la crescita dell'e-commerce nel suo complesso è un fattore chiave nell'evoluzione dell'economia globale. È uno dei fenomeni economici più significativi degli ultimi dieci anni con grosse capacità di espansione e risulta essere in grado di affermarsi durante lunghi periodi di condizioni di mercato stabili o in calo.

L'utilizzo di Internet e l'aumento dell'accesso delle persone alle nuove soluzioni tecnologiche hanno trasformato il modo di vivere e comportamenti di consumo della popolazione.

Entrando più nello specifico, per quanto riguarda il mercato alimentare online, è in continua crescita. Come riportato dall'Osservatorio eCommerce B2C Netcom- School of Management del Politecnico, il settore, Food & Grocery è in crescita del 17% rispetto al 2021 e continua a crescere sopra la media, con un totale di 4,8 miliardi di euro. Il settore è caratterizzato da un andamento favorevole, soprattutto grazie all'effetto delle iniziative online su tutto il territorio italiano e dal miglioramento dell'infrastruttura logistica e della capacità di recapito da parte delle aziende. È da sottolineare però che il 2022 risulta essere uno degli anni più complessi, sia dal lato economico che sociale dato la pandemia Covid-19 tutt'ora in atto e la recente guerra in Ucraina. Infatti, come viene descritto da Valentina Pontiggia, Direttrice dell'Osservatorio E-Commerce B2C Netcomm del politecnico di Milano: "L'instabilità geopolitica, l'inflazione, il rincaro dei beni energetici e delle materie prime, la crisi della supply chain, così come il ritorno alla piena attività dei negozi fisici, stanno influenzando la dinamica dei consumi, totali e online. L'e-commerce di prodotto, dopo due anni di crescita "straordinaria", si trova in una fase di evoluzione più strutturata e controllata. L'attenzione dei retailer online è rivolta all'ottimizzazione dei processi, non solo in chiave omnicanale, e alla sperimentazione di soluzioni marketplace, quick Commerce, second hand market, subscription, social Commerce) in grado di assecondare le nuove tendenze di consumo"⁵².

⁵² Digital24: Cresce acquisto prodotti online, <https://www.digital4.biz/marketing/ecommerce/e-commerce-b2c-cresce-acquisto-di-prodotti-online-ma-a-ritmo-piu-moderato/>

In questa tesi, ci si focalizza sul servizio E-commerce offerto da strutture di vendita al dettaglio come supermercati ed ipermercati, cercando di contribuire ed evidenziare fattori positivi sia per i consumatori e sia per le stesse imprese.

Per fare ciò, si è proposto capire quali motivazioni un questionario per la raccolta di dati, al fine di evidenziare quanto e come il settore E-commerce offerto da ipermercati e supermercati, possa influenzare positivamente sul quotidiano delle persone.

Ricordando dunque che il campione è stato auto selezionato e limitato in numerosità, i dati raccolti non hanno la stessa valenza di un campione che comprende tutta la popolazione di riferimento. Si possono però comunque tramite i metodi adottati in questa tesi trarre delle informazioni utili al fine di dimostrare ciò che si è discusso nell'elaborato. Grazie all'utilizzo del questionario come strumento di raccolta dati, si sono potuti raccogliere un totale di 252 osservazioni di cui 245 utili all'analisi.

Nello specifico sul campione è stato poi possibile individuare due gruppi distinti: chi utilizza il servizio E-commerce di ipermercati e supermercati e chi non lo usufruisce.

La numerosità del campione che dichiara di non utilizzare il servizio online fa riflettere: sono presenti 175 unità rispetto le 70 che invece dichiarano di adottare il servizio online. Al fine di identificare quali fattori possono contribuire alla scelta di usare servizi online di supermercati ed ipermercati.

Da quest'ultimo è emerso che all'aumentare dell'età, diminuirà la probabilità di adottare in futuro un servizio online offerto da supermercati ed ipermercati. Si può dunque ipotizzare che probabilmente, i clienti futuri saranno la Generazione Z (i nati dal 1995 ad oggi) e i Millennials, cresciuti negli anni '90. Queste giovani generazioni sono cresciute vedendo i cambiamenti che l'era digitale ha apportato alla vita quotidiana e sono arrivate a comprenderla e rispettarla.

Dunque, già da questa analisi su dati sociodemografici, si può comprendere che le aziende dovrebbero incentivare l'uso dell'E-commerce a generazioni più "veterane".

Passando all'analisi dei comportamenti di acquisto e focalizzandosi sul perché le unità del campione non effettuano acquisti online nel mercato alimentare online, si è riscontrata una maggiore risposta sulla variabile "Preferisce acquistare i propri prodotti direttamente al punto vendita". Come dimostrato dall'analisi di regressione logistica, all'aumentare di questa variabile, al netto della variabile età, diminuisce la probabilità di acquistare online in futuro. Probabilmente questo risultato è dovuto dal fatto che la maggior parte degli italiani preferisce ancora l'interazione umana e vuole poter scegliere e toccare con mano le cose che acquista quando fa la spesa, che è alla base del patrimonio e della cultura

gastronomica italiana. Tuttavia, lo sviluppo visto negli ultimi anni dimostra come vi sia un'esigenza latente di un servizio che fa risparmiare tempo e che aspetta semplicemente di essere soddisfatto.

Le imprese che investono sull'utilizzo dell'E-commerce come canale di vendita devono riuscire a far percepire l'utilità di questo servizio anche alle generazioni più anziane che ancora non ne fanno uso, utilizzando pubblicità o iniziative che incentivano il suo impiego. Dal questionario e tramite analisi univariata si è mostrato che la spedizione gratuita e le promozioni e sconti solo su acquisto online possono convincere le persone ad adoperare in futuro il servizio online.

Inoltre, trarre informazione anche sui fattori che incidono sull'uso del servizio E-commerce è stato condotto una Cluster analysis sulle osservazioni di quella parte del campione che ha dichiarato di far uso del servizio di E-commerce. Grazie all'analisi si è potuto segmentare le unità in tre gruppi che rappresenta comportamenti di acquisto simili, sulla base delle domande raccolte "Perché acquista online?" e "Quale fattore può incidere di più sull'acquisto online" e "Scelga quale risposte siano i punti di forza sulla consegna più importanti". Tramite questa analisi si è potuto descrivere i dati raccolti tramite il questionario.

Dopo aver identificato i gruppi di utilizzatori del servizio con comportamento simili, si rilevano informazioni su come le unità preferiscano il ritiro della propria spesa sugli acquisti effettuati online da supermercato ed ipermercato. Le unità prediligono la consegna direttamente a casa e non ritirata direttamente in punto vendita. Ciò sta a significare che la maggior parte degli intervistati, non sono propensi a spostarsi per la consegna della propria spesa e adottare nuove metodologie che possano sostituire la consegna a casa.

Infine, si raccolgono informazioni importanti sulle ragioni riguardanti il perché le unità siano propense all'utilizzo dell'E-commerce. I tre gruppi sembrano avere delle preferenze diverse, ma le motivazioni principali che emergono sono la possibilità di confrontare promozioni, offerte, prodotti e la maggiore facilità d'uso. È da qui che le imprese devono ripartire per incentivare le persone ad utilizzare il servizio online. Ottimizzare il sito online e renderlo semplice per tutti i clienti, può portare ad un incentivo al suo utilizzo.

In conclusione, l'E-commerce offerto da ipermercati e supermercati è in via di sviluppo, come riportato in questa tesi di laurea.

Le imprese che durante la pandemia Covid-19, hanno investito in maniera precoce sul mercato online, si stanno ancora assestando, sia per soddisfare le richieste da parte dei

clienti e sia a livello logistico. È importante, dunque, che si cerchi di interagire il più possibile con i clienti, i quali dopo aver utilizzato il servizio ritornino ad effettuare acquisti continui, portando ad essere la spesa online un'attività abituale. È fondamentale sfruttare l'omnicanalità ossia la centralizzazione del consumatore che deve vivere la stessa esperienza su tutti i punti di contatto che l'azienda offre quindi sia fisico tramite negozio che online tramite E-commerce.

APPENDICE:

QUESTIONARIO UTILIZZATO:

1. Lei è: *

- Uomo
- Donna

2. Cittadinanza:

- Italiana
- Estera

3. In che regione abita? *

- Abruzzo
- Basilicata
- Calabria
- Campania
- Emilia-Romagna
- Friuli Venezia Giulia
- Lazio
- Liguria
- Lombardia
- Marche
- Molise
- Piemonte
- Puglia
- Sardegna
- Sicilia
- Toscana
- Trentino Alto-Adige
- Umbria
- Valle d'Aosta
- Veneto

4. Quanti anni ha? (in numero) *

5. Stato Civile:

- Celibe/Nubile
- Coniugato/a
- Separato/a o Divorziato/a Vedovo/a

6. Ha figli?

- Sì
- No

7. Attualmente abita con qualcuno?

- Si
- No

8. Se si con chi?

- Genitori
- Marito/moglie-Compagno/a
- Coniuge e [gli
- Figli
- Solo

9. Titolo più alto da lei conseguito

- Nessun Titolo
- Licenza elementare
- Licenza di scuola media inferiore
- Diploma o qualifica di scuola media superiore
- Laurea
- Dottore di ricerca o specializzazione post-laurea

10. Attualmente lavora? *

- Si
- No
- Disoccupato
- Altro:

11. Ha mai acquistato online prodotti da supermercato? *

- Si
- No

12. Dove acquista maggiormente i prodotti? *

- Supermercato
- Ipermercato
- Altro:

13. Con che frequenza acquista i propri prodotti in supermercato/ipermercato (fisicamente)? *

- Quotidianamente
- Settimanalmente
- Mensilmente
- Annualmente
- Mai

14. Con che frequenza acquisti online da supermercato/ipermercato? *

- Quotidianamente
- Settimanalmente
- Mensilmente
- Annualmente
- Mai

15. Quale è la spesa media di un acquisto online da supermercato/ipermercato

- Sotto i 20 euro
- Dai 20 ai 50 euro
- Dai 50 ai 100 euro
- Dai 100 ai 200 euro
- Da 200 ai 300 euro
- Dai 300 in poi

16. In che fornitore acquisti i propri prodotti online? (è possibile elencare più risposte, in altro specificare il nome del fornitore)

- Coop
- Iperosano
- Aliper
- Esselunga
- Interspar/Despar
- Rossetto
- Famila
- IperLando
- Aldi
- Lidl
- Conad
- Eurospin
- Altro:

17. Dove ritira la sua spesa? *

- Direttamente in Punto vendita (spesa già preparata)
- Direttamente a casa
- Ritiro della spesa nei Locker box (quando si vuole tramite codice e nei box appositi del punto vendita)

18. Che dispositivo utilizza per acquistare la merce online? *

- Tablet
- Smartphone/Telefono
- Computer
- Altro:

19. Utilizza un'applicazione o un sito internet nell'acquistare i propri prodotti?

- App del telefono o ipad
- Sito Internet
- Altro:

20. Quale pagamento predilige quando acquista online? *

- Pagamento online tramite carta di credito/debito
- Pagamento in contanti alla consegna della merce
- Pagamento posticipato al momento del ritiro in negozio/presso centro ritiro
- Paypal
- Pagamento Anticipato
- Buoni pasto cartacei ed elettronici

21. Quanto dista mediamente il Supermercato/ipermercato da casa sua? *

- 5-10 km
- 10-20 km
- 20-30 km 30-40 km più di 40

22. Come è venuto a conoscenza dell'e-commerce da parte del supermercato/ipermercato (è possibile elencare più di una risposta)

- Volantino
- Da Internet
- Pubblicità in televisione
- Da amici e parenti
- Da manifesti
- Direttamente nel punto vendita
- Social media (Instagram, Facebook e LinkedIn)
- Altro:

23. Perché acquista online? (1=per nulla d'accordo, 5=totalmente d'accordo)

Posso farmi recapitare gli acquisti a casa	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Più semplice trovare le offerte e confrontarle	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Più semplice di fare la spesa direttamente nel supermercato/ipermercato	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A causa dell'epidemia Covid-19	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I prezzi sono più bassi	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
E' veloce acquistare i prodotti tramite app/sito piuttosto di recarsi nel punto vendita	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

24. Quale fattore può incidere di più sull'acquisto online? (1=per nulla importante, 5= estremamente importante)

Spedizione gratuita	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Invio di ritorno gratuito	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Servizio clienti	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Recensioni sul sito E-commerce	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Recensione sui prodotti	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Possibilità di scelta tra vari tipi di spedizione	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Raccomandazione di amici	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

25. Scelga fra queste risposte quali siano i punti di forza sulla consegna più importanti (1=per nulla importante, 5=estremamente importante)

Preavviso della ricezione degli invii	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Gestione online della ricezione	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Recapito al termine desiderato	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Recapito la sera	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Recapito lo stesso giorno	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Recapito di mattina	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Recapito entro due ore	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

26. Quali Prodotti prevalentemente acquista online? (massimo 2 risposte)

- Prodotti genere chimico (prodotti per la casa)
- Prodotti genere alimentare confezionati (biscotti, pasta, sughi pronti)
- Prodotti Freschi/da frigo (formaggi, yogurt e latticini in generale)
- Prodotti Surgelati

27. Utilizzavate l'e-commerce del punto vendita in questione prima della pandemia Covid 19?

- Si
- No

28. Quali forme di assistenza alla clientela sono importanti per lei? (massimo 3 risposte)

- E-mail
- Modulo di contratto

- Telefono
- Consulenza personalizzata
- FAQ
- Chat
- Messenger
- Social Media
- Chatbot

29. Che cosa la induce ad acquistare sui mercati virtuali? *

- Confrontare prezzi, prodotti e promozioni
- Prezzi bassi
- Raggruppamento di prodotti
- Maggiore Facilità d'uso
- Valutazione dei prodotti e dei commenti dei clienti Offerte speci[che per i clienti

30. Quali secondo lei possono essere gli svantaggi sull'acquisto online? (massimo 2 risposte)

- prodotti da lei richiesti non vengono spediti
- Sostituzione dei prodotti senza il proprio consenso
- Scadenza dei prodotti reperibili a breve termine
- Involucro della merce danneggiato
- Spedizione troppo costosa
- Mancanza di assistenza e servizio nell'acquisto online

31. Infine, da una scala da 1 a 5 quanto vi ritenete soddisfatti dei vostri acquisti online? *

1 2 3 4 5

Per nulla Moltissimo

32. Dove acquista i prodotti? *

- Supermercato
- Ipermercato
- Altro:

33. Con che frequenza acquista i propri prodotti in supermercato/ipermercato (fisicamente)? *

- Quotidianamente
- Settimanalmente
- Mensilmente

- Annualmente
- Mai

34. Perché non ha mai acquistato online da un supermercato? (massimo 3 risposte) *

- Non forniscono il servizio e-commerce nel suo Paese
- Preferisce acquistare i propri prodotti direttamente al punto vendita
- Nel sito/app non c'è lo stesso assortimento come in negozio
- Non è pratico nell'utilizzo di siti o app per l'acquisto online
- Obbligo di registrazione
- Spese di consegna
- Sensazione di scarsa sicurezza nell'acquisto online

35. Che cosa ne pensa della spesa online offerta dai supermercati/ipermercati? (una breve descrizione)

36. Secondo lei acquisterà online dai supermercati in futuro? *

- Sì
- No

37. Quale opzione sceglierebbe per convincerla ad utilizzare la spesa online? (massimo 2 risposte)

- Spedizione gratuita
- Reso gratuito e direttamente a casa sua
- Promozioni e sconti solo su acquisto online
- Raccolta punti sull'acquisto online da utilizzare in futuro
- Nessuna delle risposte precedenti

38. Quanto tempo trascorre nei social Media? (es. Facebook, Instagram, Twitter ecc..)

1 2 3 4 5

Non utilizzo social Utilizzo assiduo social

39. Quali di questi servizi utilizza maggiormente? (massimo 2 risposte)

- Tv tradizionale
- Netflix
- Sky
- Amazon Prime
- Radio
- Altro:

40. Quali prodotti acquista prevalentemente online? (massimo 3 risposte) *

- Viaggi
- Libri, musica, [Im
- Biglietti per eventi
- Abbigliamento
- Computer, Apparecchi elettronici

- Articoli sportivi
- Giocattoli
- Cosmetici
- Mobili
- Altro

41. Dove inizia a ricercare generalmente i prodotti quando acquista online? (Massimo 2 risposte)

- Motori di ricerca (google)
- Social Media (Instagram, Facebook ecc..)
- Forum/ piattaforme di discussione
- Pubblicità in televisione Giornale
- Altro:

41. I prodotti che lei acquista generalmente da dove vengono spediti Italia

- Francia
- Belgio
- Spagna
- Germania
- Altri

42. Generalmente quanto siete soddisfatti nei vostri acquisti online? *

1 2 3 4 5

Poco Moltissimo

CODICI UTILIZZATI PER L'ANALISI

CLUSTER ANALYSIS

```
library(readxl)
dati<- read_excel("C:/Users/check/Downloads/FINALE (3).xlsx")
dim(dati)
names(dati)
dati1=dati[,38:57]
names(dati1)=c("consegna_casa","confronto_offerte","semplicità","Covid","prezzi","ve
locità","spedizione_gratuita","reso_gratuito","servizio_clienti","servizio_sito_E-
commerce","Recensioni","scelta_spedizione","raccomandazione_amici","Preavviso_ri
cezione_invii","Gestione_online_ricezione","Recapito_al_termine_desiderato","Recapito
_sera","Recapito_stesso_giorno","Recapito_mattina","Recapito_entro_due_ore")
summary(dati1)
sum(is.na(dati1))
which(is.na(dati1))
apply(dati1,2,function(x) sum(is.na(x)))
which(is.na(dati1$reso_gratuito))
dati1$reso_gratuito[66]=median(dati1$reso_gratuito,na.rm = T)
library(factoextra)
library(cluster)
df<- scale(dati1)
set.seed(123)
fviz_nbclust(df, kmeans, method = "wss") +
  geom_vline(xintercept = 3, linetype = 2)
gap_stat<-clusGap(df,
  FUN = kmeans,
  nstart = 3,
  K.max = 10,
  B = 500)
fviz_gap_stat(gap_stat)

km <- kmeans(dati1, centers = 3, nstart = 5)
clust <- agnes(df, method = "ward")

#produce dendrogram
pltree(clust, cex = 0.6, hang = -1, main = "Dendrogram")

fviz_dend(clust, k = 3, # Cut in four groups
  cex = 0.5, # label size
  k_colors = c("red", "#00AFBB", "green"), color_labels_by_k = TRUE, rect =
TRUE)
```

```
groups <- cutree(clust, k=3)
tt=table(groups,km$cluster)
sum(apply(tt,1,max))/length(groups)
dati2=dati[,1:38]
names(dati2)
dati2$ANNI
dati2$SESSO
```

```

prop.table(table(dati2$SESSO,km$cluster),2)
boxplot(dati2$ANNI~km$cluster)
prop.table(table(dati2$DISTANZA_SUPERMERCATO,km$cluster),2)
dati$DISTANZA_SUPERMERCATO[dati$DISTANZA_SUPERMERCATO=="5-10
km"]=10
dati$DISTANZA_SUPERMERCATO[dati$DISTANZA_SUPERMERCATO=="10-20
km"]=20
dati$DISTANZA_SUPERMERCATO[dati$DISTANZA_SUPERMERCATO=="20-30
km"]=30
dati$DISTANZA_SUPERMERCATO[dati$DISTANZA_SUPERMERCATO=="più di
40 km"]=40
dati$DISTANZA_SUPERMERCATO=as.numeric(dati$DISTANZA_SUPERMERCATO),pch=3)
#boxplot(as.numeric(dati2$DISTANZA_SUPERMERCATO)~km$cluster)
boxplot(dati$DISTANZA_SUPERMERCATO~km$cluster)
plot(1:3,tapply(dati$DISTANZA_SUPERMERCATO,km$cluster,mean),pch=1,cex=1.5
,ylab="distanza")
prop.table(table(dati$WEB,km$cluster),2)

```

REGRESSIONE LOGISTICA

```

dati<- read_excel("C:/Users/check/Downloads/FINALE.xlsx",sheet=2)
names(dati)
summary(dati)
table(dati[,2])
ind=sapply(1:29,function(x)length(unlist(unique(dati[,x]))))#per eliminare variabili
costanti
dati=dati[,-which(ind<2)]
apply(dati,2,function(x) sum(is.na(x)))
table(dati$SPEDIZIONE_ACQUISTI_ONLINE)

dati=dati[,-which(apply(dati,2,function(x) sum(is.na(x)))>9)]#per eliminare la var
SPEDIZIONE_ACQUISTI_ONLINE

dati=dati[-which(apply(dati,1,function(x) sum(is.na(x)))>0),]
m1=glm(as.factor(ACQUISTO_ONLINE_FUTURO)~.,family="binomial",data=dati)
#summary(m1)

m2=step(m1,k=log(dim(dati)[1]))#BIC
coef(m2)
levels(as.factor(dati$NO_ACQUISTO_ONLINE))
m1=glm(as.factor(ACQUISTO_ONLINE_FUTURO)~ANNI,data=dati,family="binomial")
summary(m1)
plogis(coef(m1))
plogis(confint(m1))

```


ELENCO FIGURE E TABELLE:

Figura 1.1 Gli acquisti eCommerce B2c tra prodotti e servizi. Fonte: Dipartimento di ingegneria, Politecnico di Milano.

Figura 1.2: Distribuzione per classi d'età. Fonte: We are Social. Fonte: Digital Report Italia 2021

Figura 1.3: E-commerce in Italia 2021. Fonte: Casaleggio

Figura 1.4: Mobile: fatturato e crescita. Fonte: Casaleggio Associati 2021

Figura 2.1 Tipologie di vendita al dettaglio in base alla dimensione. Fonte: InsideMarketing

Figura 2.2, Totale Largo Consumo Confezionato. Previsioni IRI condotte a livello di macrocategoria ECR prima dell'evento pandemico da Covid-19. Elaborazioni RemLab su dati GFK panel famiglie. Fonte: IRI Online.

Figura 2.3: Coop Fiorfood.

Figura 2.4: Quote delle vendite in base ai servizi Home delivery e Click&Collect. Fonte: IRI Liquid Data

Figura 2.5, Contributo alla crescita dei canali e la loro quota di mercato. Fonte: IRI Liquid Data

Figura 2.6, Crescita fatturato E-commerce alimentari. Fonte: Dipartimento di ingegneria gestionale, Politecnico di Milano

Figura 3.1, Tecniche di rilevazione di dati. Fonte: Dipartimento di Scienze Statistiche, Università di Padova

Figura 3.2, Errori nell'indagine Campionaria. Fonte: Teoria e tecnica dell'indagine statistica e del campionamento, Dipartimento di Scienze Statistiche, Università di Padova

Figura 3.3, I diversi tipi di scale di misurazione. Fonte: Metodi quantitativi per la segmentazione e posizionamento, Università Ca' Foscari di Venezia

Figura 3.4 Tecnica di valutazione non comparative. Fonte: Survey-measurement and scaling

Figura 3.5 Esempio di scala di valutazione continua non comparativa. Fonte: Survey-measurement and scaling

Figura 3.6 Esempio scala Linkert; Fonte: Metodi quantitativi per la segmentazione e posizionamento, Università Ca' Foscari di Venezia

Figura 3.7, Esempio di Semantic Differential; Fonte: Metodi quantitativi per la segmentazione e posizionamento, Università Ca' Foscari di Venezia

Figura 3.8, Esempio Stapel scale. Fonte: Survey-measurement and scaling

Tabella 3.9, Vantaggi e svantaggi tecniche di valutazione comparative. Fonte: Elaborazione personale

Figura 3.10, Esempio matrice di distanza.

Figura 3.11, Raffigurazione dell'esempio di metodo del legame singolo. Fonte: Rescoop

Figura 3.14, Rappresentazione del metodo cetroide. Fonte: Rescoop

Figura 3.15, Esempio di dendogramma. Fonte: Introducing to statistical learning

Figura 3.16, Esempio grafico Elbow. Fonte: Metodi quantitativi per la segmentazione e posizionamento, Università Ca' Foscari di Venezia

Figura 3.17, Esempio grafico Silhouette. Fonte: Metodi quantitativi per la segmentazione e posizionamento, Università Ca' Foscari di Venezia

Figura 3.18, Esempio di rappresentazione di una retta lineare con valori osservati e retta stimata. Fonte: Statistica di base, Apogeo, D.Moore

Figura 4.1: Mappa descrittiva del questionario E-commerce supermercati ed ipermercati. Fonte: Elaborazione personale

Figura 4.2, Percentuale di risposte in base al genere. Fonte: Google questionare

Figura 4.3, Tabella distribuzione d'età del campione. Fonte: Elaborazione personale

Figura 4.4, Distribuzione per genere e classi d'età. Fonte: Google questionare

Figura 4.5, Tabella distribuzione titolo di studio. Fonte: Elaborazione personale

Figura 4.6, Tabella distribuzione stato lavoro, Fonte. Elaborazione personale

Figura 4.7, Grafico a torta sulla distribuzione della composizione familiare e sociale delle unità, Fonte: Google questionare

Figura 4.8 Grafico a torta sulla distribuzione della coabitazione familiare, Fonte: Google questionare

Figura 4.9, Grafico a torta sulla distribuzione frequenze per chi effettua acquisti online da supermercato ed ipermercato, Fonte: Google questionare

Figura 4.10, Tabella distribuzione della fascia d'età e genere su coloro che utilizzano il servizio E-commerce ipermercato e supermercato. Fonte: Elaborazione personale

Figura 4.11, Grafico a barre relativo alla distribuzione di frequenza acquisto online da supermercato ed ipermercato. Fonte: Elaborazione personale

Figura 4.12, Grafico a barre relativo alla distribuzione di frequenza spesa media di un acquisto online supermercato/ipermercato. Fonte: Elaborazione personale

Figura 4.13, Grafico a barre sulla distribuzione dei fornitori principali che forniscono prodotti online. Fonte: Elaborazione personale

Figura 4.14, Grafico a torta sulla distribuzione del ritiro della spesa online; Fonte: Google questionare

Figura 4.15, Grafico a barre sulla distribuzione dei metodi di conoscenza E-commerce. Fonte: Google questionare

Figura 4.16, Grafico sulla motivazione di acquisto online. Fonte: Elaborazione personale

Figura 4.17, Grafico a barre sui fattori che incidono sull'acquisto online. Fonte: Elaborazione personale

Figura 4.18, Grafico a barre sui punti di forza della consegna tramite E-commerce. Fonte: Elaborazione personale

Figura 4.19, Grafico a torta sulla distribuzione di utilizzo del servizio E-commerce prima della Pandemia Covid-19. Fonte: Google questionare

Figura 4.20, Grafico a torta sulla distribuzione delle motivazioni che inducono i consumatori ad utilizzare il mercato online. Fonte: Google questionare

Figura 4.21, Grafico a barre sulla distribuzione delle risposte riguardante gli svantaggi sull'acquisto online. Fonte: Google questionare

Figura 4.22, Grafico a barre sulla distribuzione della soddisfazione sugli acquisti online. Fonte: Google questionare

Tabella 4.23, Tabella sulla media e deviazione standard ("Perché acquista online?"). Fonte: Elaborazione Personale

Tabella 4.24, Tabella media e deviazione standard sulla domanda "Quale fattore può incidere di più sull'acquisto online". Fonte: Elaborazione Personale

Tabella 4.25, Tabella media e deviazione standard sulla domanda "Scelga fra queste risposte quali siano i punti di forza sulla consegna più importanti". Fonte: Elaborazione personale

Figura 4.26, Grafico correlazione scala Likert. Fonte: Elaborazione personale

Figura 4.27, Grafico a barre in percentuale sulla motivazione per la quale i consumatori non acquistano online da supermercato. Fonte: Google questionare

Figura 4.28, Grafico a torta sulla distribuzione di coloro che effettueranno in futuro un acquisto online. Fonte: Google questionare

Figura 4.29, Grafico a barre sulle opzioni che possono convincere coloro che non utilizzano la spesa online ad usufruire di quest'ultima. Fonte: Google questionare

Figura 4.30, Rappresentazione del grafico Elbow. Fonte: Elaborazione personale

Figura 4.31, Rappresentazione del dendogramma. Fonte: Elaborazione personale

Tabella 4.32, Tabella sulla media e deviazione standard delle variabili dei cluster. Fonte: Elaborazione Personale

Tabella 4.33, Tabella sulla media e deviazione standard delle variabili dei cluster. Fonte: Elaborazione Personale

Tabella 4.34, Tabella sulla media e deviazione standard delle variabili dei cluster. Fonte: Elaborazione Personale

Tabella 4.35, Tabella sulla segmentazione di genere per i vari cluster. Fonte: Elaborazione personale

Tabella 4.36, Tabella di distribuzione sulla batteria TITOLO STUDIO confrontata con i cluster. Fonte: Elaborazione personale.

Tabella 4.37, Tabella di distribuzione percentuale sulla batteria TITOLO STUDIO confrontata con i cluster. Fonte: Elaborazione personale.

Tabella 4.38, Tabella di distribuzione sulla batteria LAVORO confrontata con i cluster. Fonte: Elaborazione personale.

Tabella 4.39, Tabella di distribuzione sulla batteria FREQUENZA_ONLINE_SUPERMERCATO confrontata con i cluster. Fonte: Elaborazione personale.

Tabella 4.40, Tabella di distribuzione sulla batteria FREQUENZA_ONLINE_SUPERMERCATO confrontata con i cluster. Fonte: Elaborazione personale.

Tabella 4.41, Tabella sulla distanza dal supermercato o ipermercato dei cluster. Fonte: Elaborazione personale

Figura 4.42, Plot sulla media distanza dal supermercato dei Cluster. Fonte: Elaborazione personale

Tabella 4.43, Tabella sulle medie della distanza da supermercato o ipermercato. Fonte: Elaborazione personale

Figura 4.44, Rappresentazione dei boxplot sull'età dei cluster, Fonte: Elaborazione personale

Figura 4.45, Tabella sulla percentuale di risposta vantaggi acquisto online da supermercato o ipermercato dei cluster, Fonte: Elaborazione Personale

Tabella 4.46, Risultati modello regressione logistica con variabile ANNI. Fonte: Elaborazione personale

Tabella 4.47, Risultati regressione logistica con variabile ANNI e NO_ACQUISTO_ONLINE, Fonte: Elaborazione Personale

BIBLIOGRAFIA:

- TIZIANO VESCOVI: Il marketing e la rete: la gestione integrata del Web nel business: comunicazione, e-commerce, sales management, business to business, Milano: Il sole 24 ore, 2007, pag.29
- UNIVERSITA' DEGLI STUDI DI BRESCIA: Gli elementi maggiormente impattanti sulla propensione all'acquisto online, Ricerca realizzata per conto di Aicel, 2019
- ENRICO COLLA: Gli ipermercati: sviluppo e maturità delle grandi superfici di vendita in Italia e in Europa, Milano, ETAS libri, 1992, pag. 145
- ENRICO COLLA: La grande distribuzione in Europa: evoluzione delle formule distributive, strategie e strutture aziendali, rapporti con l'industria, Milano, ETAS libri, 1995, pagina 96
- M.J. KOTLER: Marketing management. Prentice Hall, 1984.
- QUESTIONARE DESIGN: Handbook of Recommended Practices for Questionnaire Development and Testing in the European Statistical System, A.A.V.V., 2006, pag.25
- PIERGIORGIO CORBETTA: Social Research, Theory Methods and Techniques, Sage Publication, London, 2003, pag. 218
- GOMES SOFIA, JOAO M. LOPEZ. "*Evolution of the Online Grocery Shopping Experience during the COVID-19 Pandemic: Empiric Study from Portugal*" *Journal of Theoretical and Applied Electronic Commerce Research*, 2022

SITOGRAFIA:

- IL CORRIERE: Come è nato l'E-commerce, la storia i primi passi di Amazon, Ebay e Alibaba: <https://www.corriere.it/tecnologia/cards/ecommerce-prime-informazioni-scambiate-digitalmente-nascita-colossi-commercio-elettronico/nascita-ebay.shtml>
- IL SOLE 24 ORE: I 25 anni di Amazon: dal garage di Bezos al negozio globale di ogni cosa: https://www.ilsole24ore.com/art/dal-garage-bezos-mille-miliardi-storia-amazon-che-ora-vuole-superare-apple-AE4ZzfjF?refresh_ce=1
- CORRIERE DELLA SERA: Il primo acquisto online: https://www.corriere.it/tecnologia/cards/ecommerce-prime-informazioni-scambiate-digitalmente-nascita-colossi-commercio-elettronico/primo-acquisto-digitale_mobile.shtml
- ANSA: La crescita dell'e-commerce traina ogni mercato: https://www.ansa.it/pressrelease/economia/2020/11/20/la-crescita-virtuosa-delle-commerce-traina-ogni-mercato_e41da430-cao2-4811-a77d-d4e932de14ca.html
- ENGAGE: Cresce l'adozione delle-commerce B2B tra le aziende italiane. I dati di NETCOM: <https://www.engage.it/dati-e-ricerche/cresce-l-adozione-dell-ecommerce-b2b-tra-le-aziende-italiane-i-dati-di-netcomm.aspx>
- BITMAT: Boom e-commerce: mercato da 39,4 miliardi di euro nel 2021: <https://www.bitmat.it/blog/internet/ecommerce/boom-e-commerce-mercato-da-394-miliardi-di-euro-nel-2021-21-rispetto-al-2020/>
- PUNTOINFORMATICO: Un italiano su due acquista online la svolta della pandemia: <https://www.punto-informatico.it/un-italiano-su-due-acquista-online-la-svolta-della-pandemia/>
- GOODCOM: Digital report 2021: I dati italiani: <https://www.goodcom.it/digital-report-2021-i-dati-italiani/>
- CASALEGGIO: E-commerce in Italia 2021: <https://www.casaleggio.it/focus/rapporto-e-commerce-in-italia-2021/>
- INSIDE MARKETING: GDO: <https://www.insidemarketing.it/glossario/definizione/gdo/>

- RETAIL E GDO: trend, innovazione, cambiamenti e best practice:
<https://www.tuttofood.it/news/news/retail-e-gdo--trend--innovazione--cambiamenti-e-best-practice.html>
- GDO E DIGITALIZZAZIONE COME SARA' IL MERCATO DEL FUTURO:
<https://www.meiko.it/it/mondo-meiko/blog/gdo-e-digitalizzazione-come-sara-il-supermercato-del-futuro;>
- ALFACOD: Coop punta sul “supermercato del futuro”: coinvolgimento ed experience: <https://www.alfacod.it/blog-supermercato-futuro-coop-customer-experience;>
- IL SUPERMERCATO DEL FUTURO TUTTO UN ESPERIMENTO:
[https://ledehors.it/news/48-il-supermercato-del-futuro_%26egrave%3B_tutto_un_esperimento;](https://ledehors.it/news/48-il-supermercato-del-futuro-%26egrave%3B-tutto-un-esperimento;)
- FOODWEB: <https://www.foodweb.it/2020/08/amazon-fresh-il-primo-store-di-freschi-targato-amazon-fresh-apre-in-california/>
- IQUII: I Retail dal primo supermercato all’omnicanalità: La storia, l’evoluzione e le opportunità di crescita: <https://iquii.com/2018/04/18/retail-omnicanalita-sfide-opportunita/>
- PORTLAND RESEARCH GROUP:
<https://www.portlandresearch.com/about/what-is-market-research>
- QUESTION STAR: Survey: Measurement and Scaling:
<https://www.questionstar.com/textbook-principles-of-survey-research/survey-measurement-and-scaling/non-comparative-scales/>
- DIGITAL24: Cresce acquisto prodotti online,
<https://www.digital4.biz/marketing/ecommerce/e-commerce-b2c-cresce-acquisto-di-prodotti-online-ma-a-ritmo-piu-moderato>
- FOODAFFAIRS: La spesa online in Italia,
<https://www.foodaffairs.it/2021/07/07/trend-nel-2021-la-spesa-online-in-italia-nel-settore-foodgrocery-superera-i-4-miliardi-di-euro-38-sul-2020-dati-osservatorio-netcomm-politecnico-di-milano-su-ecommerce-food/>