

POLITECNICO DI TORINO

Corso di Laurea Magistrale in Ingegneria Gestionale



Tesi di Laurea Magistrale

QUALI SONO I DRIVERS DELLA SOPRAVVIVENZA DI UN RETAILER? *Il caso della città di Torino*

Relatori

Prof. Giulio Zotteri

Candidato

Alice Mirabella

Aprile 2018

SOMMARIO

1.	INTRODUZIONE	1
2.	RASSEGNA DELLA LETTERATURA	3
2.1	LA CONCENTRAZIONE URBANA.....	3
2.1.1	<i>Evoluzione storica della struttura urbana e costi di trasporto</i>	3
2.1.2	<i>Evoluzione della densità urbana ed espansione delle città</i>	5
2.1.3	<i>La città moderna: i poli-centri e l'agglomerazione</i>	7
2.2	L'AGGLOMERAZIONE NEL MONDO DEL RETAIL	10
2.2.1	<i>Come impatta le decisioni la tipologia di prodotto venduto</i>	14
2.2.2	<i>Come impatta le decisioni la presenza/assenza di amenità</i>	16
2.2.3	<i>Uno sguardo ai modelli di agglomerazione</i>	17
3.	IL MONDO DEL RETAIL: TRASFORMAZIONI IN CORSO	21
3.1	COSA RAPPRESENTA L'ATTIVITÀ DI RETAILING.....	21
3.2	I VARI FORMATI DEL RETAIL	22
3.2.1	<i>Giganti Chain Stores vs Mom-and-Pop Stores</i>	23
3.3	RETAIL TRENDS: PROFONDA TRASFORMAZIONE IN ATTO	27
4.	DATA SET & TOOLS	32
4.1	ORIGINE DEI DATI.....	32
4.2	QGIS.....	36
4.3	SPSS	36
5.	METODO UTILIZZATO: LA SURVIVAL ANALYSIS	38
5.1	COS'È UNA DURATION ANALYSIS?.....	38
5.1.1	<i>Perché non si utilizza un modello ols?</i>	40
5.2	BACKGROUND – LA REGRESSIONE LOGISTICA	41
5.2.1	<i>Estensione all'event history analysis – il ruolo cruciale del tempo</i>	41
5.2.2	<i>La curva di sopravvivenza e la "life table"</i>	42
5.3	TERMINOLOGIA E NOTAZIONI	44
5.3.1	<i>L'ipotesi di rischi proporzionali</i>	46
5.4	LA STIMA DEI PARAMETRI.....	47
5.4.1	<i>Verifica della Bontà del Modello in termini di adattamento</i>	48

5.4.2	<i>Analisi dei residui</i>	49
5.5	TECNICA DI REGRESSIONE SEMI-PARAMETRICA: COX REGRESSION.....	51
6.	ANALISI EMPIRICA: IL CASO DELLA CITTÀ DI TORINO.....	56
6.1	I RETAILERS A TORINO, STUDIO DEI FATTORI IMPATTANTI.....	56
6.1.1	<i>Fattori impattanti - La popolazione</i>	57
6.1.2	<i>Fattori impattanti – Il formato del Retailer</i>	59
6.1.3	<i>Fattori impattanti – Il posizionamento geografico del Retailer</i>	64
6.1.4	<i>Fattori impattanti – La Categoria Merceologica del Retailer</i>	66
6.1.5	<i>Fattori impattanti – L’Agglomerazione tra Categorie Merceologiche</i>	70
6.2	ANALISI DI SOPRAVVIVENZA	77
6.2.1	<i>Life Table e Cox Regression – Categoria Merceologica</i>	78
6.2.2	<i>Cox Regression – Drivers del generico negozio in Torino</i>	85
6.2.3	<i>Cox Regression – Drivers per le diverse Categorie Merceologiche</i>	89
7.	CONCLUSIONI	103
8.	BIBLIOGRAFIA	108
9.	APPENDICE A	111
10.	APPENDICE B.....	114
11.	APPENDICE C.....	117
12.	APPENDICE D	120
13.	APPENDICE E	125

INDICE DELLE FIGURE

<i>Figura 1- Estensione area di Torino</i>	32
<i>Figura 2- categorie merceologiche</i>	34
<i>Figura 3- categorie merceologiche selezionate</i>	35
<i>Figura 4- Esempio life table</i>	43
<i>Figura 5 - Esempio Funzione di Densità di Probabilità</i>	44
<i>Figura 6 – Esempio Funzione cumulata di probabilità</i>	45
<i>Figura 7 – Esempio Funzione di Sopravvivenza</i>	45
<i>Figura 8 - Circoscrizioni Torino</i>	56
<i>Figura 9- Zone statistiche Torino</i>	57
<i>Figura 10 - Mappa di Posizionamento</i>	62
<i>Figura 11- Mappa di posizionamento Formati</i>	62
<i>Figura 12 - Distribuzione Metri Quadri</i>	63
<i>Figura 13- Mappa di posizionamento Quartieri</i>	65
<i>Figura 14 - Distribuzione Distanza dal Centro</i>	66
<i>Figura 15- Mappa di Posizionamento Categorie Merceologiche</i>	69
<i>Figura 16 - Istantanea al 2017 Abbigliamento</i>	71
<i>Figura 17 - Istantanea al 2017 Supermercati</i>	72
<i>Figura 18 - Life Table</i>	77
<i>Figura 19 - Curva di Sopravvivenza</i>	78
<i>Figura 20 – Curva di Sopravvivenza Abbigliamento</i>	80
<i>Figura 21 - Curva di Sopravvivenza Supermercati</i>	81
<i>Figura 22 - Driver Dimensione per tutte le categorie</i>	93
<i>Figura 23 - Driver Distanza dal centro per tutte le categorie</i>	98
<i>Figura 24 - Drivers per tutte le categorie merceologiche</i>	102
<i>Figura 25 – Alimentari</i>	111
<i>Figura 26 - Articoli Animali</i>	111
<i>Figura 27 - Articoli per la Casa</i>	111
<i>Figura 28 - Articoli Funerari</i>	111
<i>Figura 29 – Calzature</i>	111
<i>Figura 30 – Cartolerie</i>	111
<i>Figura 31 – Elettronica</i>	112
<i>Figura 32 – Farmacie</i>	112
<i>Figura 33 - Fiori e Piante</i>	112

<i>Figura 34 – Giocattoli</i>	112
<i>Figura 35 – Ipermercati</i>	112
<i>Figura 36 – Librerie</i>	112
<i>Figura 37 – Minimercati</i>	112
<i>Figura 38 – Mobili</i>	112
<i>Figura 39 – Ottica</i>	113
<i>Figura 40 - Profumeria e Cosmetica</i>	113
<i>Figura 41 – Quotidiani</i>	113
<i>Figura 42- Tabacchi</i>	113
<i>Figura 43 – Telefonia</i>	113
<i>Figura 44/63 - APPENDICE B</i>	114

INDICE DELLE TABELLE

<i>Tabella 1 - Sopravvivenza Abbigliamento</i>	84
<i>Tabella 2 - Sopravvivenza Supermercati</i>	84
<i>Tabella 3 - Riassunto Sopravvivenza Categorie</i>	85
<i>Tabella 4 - Output Generico 1</i>	86
<i>Tabella 5 - Output generico 2</i>	86
<i>Tabella 6 - Metri Quadri abbigliamento</i>	90
<i>Tabella 7 - Metri Quadri Supermercati</i>	92
<i>Tabella 8 - Metri Quadri Farmacie e Parafarmacie</i>	92
<i>Tabella 9 - Distanza dal centro Abbigliamento</i>	95
<i>Tabella 10 - Distanza dal centro Supermercati</i>	96
<i>Tabella 11 - Distanza dal centro Profumeria e Cosmetica</i>	97
<i>Tabella 12 - Agglomerazione Abbigliamento</i>	99
<i>Tabella 13 - Agglomerazione Supermercati</i>	100
<i>Tabella 14 - Agglomerazione Calzature</i>	102
<i>Tabella 15/32 – APPENDICE C</i>	117
<i>Tabella 33/50 - APPENDICE D</i>	120
<i>Tabella 51/66 - APPENDICE E</i>	125

1. INTRODUZIONE

Le scelte di localizzazione degli spazi commerciali in una città non sono mai state affidate al caso e, fin dai principi dell'economia, (Marchall, 1890), hanno rappresentato un affascinante elemento di studio. Nelle città Medievali, il commercio era localizzato vicino agli edifici amministrativi in modo da capitalizzare sui movimenti di crescita; successivamente, si formarono molte strade commerciali all'interno delle città, a causa dei vantaggi economici e di attrattività a cui la prossimità conduce. Agli inizi del XX secolo, Hotelling osservò che le bancarelle di gelati sceglievano di posizionarsi, e quindi concentrarsi, al centro delle spiagge, in modo da raggiungere l'intero mercato, piuttosto che disperdersi e controllare solo una piccola porzione di esso (Hotelling, 1929).

Il mercato dei retailers, uno dei sistemi commerciali interessanti per l'Economia Industriale, è caratterizzato da tipiche dinamiche di entrata ed uscita dal mercato stesso. Prima di stabilirsi in città, i retailers prendono delle decisioni circa la loro localizzazione ottimale: questa scelta si rivela più importante e vitale di quanto ci si possa aspettare. A tal proposito, Taneja affermò: *"No matter how good its offering, merchandising or customer service, every retail company still has to contend with three critical element of success: location, location, location"* (Taneja, 1999). Alexander et al. aggiunsero che *"è il sito che crea l'acquisto"* (Alexander et al., 1999). Con il termine **location** non si intende, però, il mero posizionamento su un punto di una mappa, ma molto più: la location è una sovrapposizione di aspetti sociali, economici e morfologici della vita urbana, come centralità e accessibilità, layout e design, visibilità e popolarità, attrattività del cluster e capacità di spill-over. Una posizione che risulterà preziosa è quella che risponde ai bisogni, sia reali che percepiti, dei retailers e dei consumatori. La vicinanza e la centralità, ad esempio, non sono più solo sinonimo di "centro città"; lo dimostra il numero sempre crescente di aree commerciali al di fuori dei centri urbani, che presentano la caratteristica di "centralità", grazie alla loro accessibilità, attrattività e

concentrazione di attività su larga scala. La fedeltà ad un negozio non può essere solo influenzata dalla distanza, come affermano (Rabbanee et al. , 2012); molti studi degli ultimi decenni hanno dimostrato che i consumatori presentano delle preferenze in base ad un mix di fattori, i quali possono portare alla scelta dell'atmosfera e dei siti localizzati fuori dal centro cittadino.

Comprendere le logiche e le dinamiche dei moderni retailers può rappresentare la chiave di volta nell'interpretazione e regolazione dei processi di sviluppo urbano, sociale ed economico. Il primo passo per ridefinire le politiche pubbliche e sviluppare delle nuove strategie di sviluppo a livello cittadino è quello di valutare ed interpretare le nuove configurazioni degli spazi commerciali all'interno delle città, (M. Saraiva, P. Pinho, 2015).

Il seguente elaborato, mediante un excursus sulla letteratura in questione e una successiva analisi empirica, si propone di studiare i fattori che impattano le scelte e le dinamiche dei retailers. In particolare, si analizza se ed in che modo determinate scelte effettuate dal retailer ex-ante abbiano aumentato, o diminuito, la sua *probabilità di sopravvivere*. L'analisi empirica è applicata alle attività commerciali della città di Torino, mediante l'utilizzo della Survival Analysis; tale metodo è presentato dapprima dal punto di vista teorico e, successivamente, applicato al dataset a disposizione. Sono quindi, presentati alcuni dei drivers che, nello specifico caso torinese, conducono i retailers a maggiori probabilità sopravvivenza.

2. RASSEGNA DELLA LETTERATURA

Di seguito si propone una rassegna della letteratura che permetta di inquadrare il contesto generale in cui si eseguirà la successiva analisi. Questa parte si può suddividere in tre macro aree: nella prima si analizzano i contributi in ambito urbano che fondano le economie di agglomerazione a livello microeconomico, approfondendo i temi più significativi per la ricerca ogni qual volta questi vengono toccati. La seconda parte è dedicata allo studio di tale tema dal punto di vista dei Retail.

2.1 LA CONCENTRAZIONE URBANA

Di seguito viene analizzata l'evoluzione che ha subito la struttura urbana, fino ad arrivare alla città intesa in chiave moderna.

2.1.1 EVOLUZIONE STORICA DELLA STRUTTURA URBANA E COSTI DI TRASPORTO

L'evoluzione della struttura urbana è il principale oggetto del lavoro di (A. Anas, R. Arnett, & K. A. Small, 1998) *"Urban Spatial Structure"*, 1998. In *Urban Spatial Structure* viene effettuata una descrizione storica delle teorie di agglomerazione e un commento dei differenti modelli, partendo da quelli monocentrici fino ad arrivare quelli policentrici. Il paper mostra il mutamento della struttura urbana, in funzione dei costi di spostamento e dell'evoluzione tecnologica.

Secondo gli autori la forma delle città moderne è il risultato delle evoluzioni nel campo dei trasporti e della comunicazione. Infatti, prima del 1840 le città erano ubicate in prossimità delle sorgenti d'acqua ed in questa maniera risultava semplice lavorare le merci in quantità sufficiente a garantire l'ottenimento delle economie di scala. In quell'epoca, gli spostamenti tra le città avvenivano prevalentemente mediante cavalli, quindi per minimizzare i trasporti veniva creato un unico centro produttivo vicino alle zone di maggiore interesse, come porti o miniere. Alla fine dell'800, grazie all'avvento del telegrafo, i flussi

informativi tra le città vennero velocizzati fortemente, ma i costi comunicativi erano talmente elevati che venne confermato il modello di unico centro manifatturiero nodale; infatti dato che, in quel periodo, gli spostamenti all'interno della città avvenivano maggiormente a piedi, sia i ricchi, che i poveri, erano maggiormente propensi a vivere in prossimità del centro. Nei primi anni del 900 avvennero i cambiamenti più rilevanti: l'invenzione del telefono e la diffusione del motore a combustione interna. Grazie a queste due innovazioni i costi e i tempi di trasporto subirono una riduzione notevole; grazie a ciò, gli affari iniziarono ad espandersi e ad avvenire anche al di fuori dei centri cittadini. L'innovazione tecnologica più impattante fu costituita dall'automobile. L'automobile divenne il mezzo di trasporto che permise la grande movimentazione di massa, caratterizzata da velocità, privatezza e convenienza. Grazie a ciò, si verificò una grande espansione delle città e la crescita delle **“edge cities”** al di fuori dei centri cittadini, costituite da una grande presenza di zone dedicate al settore Retail (vendita al dettaglio) e da una fitta concentrazione di uffici.

Un altro paper analizzato è quello di (W. Alonso, 1960) *“A Theory of Urban Land Market”*, 1960, che studia il contributo dei costi di trasporto all'evoluzione della struttura urbana. Viene studiato il paradosso delle città americane, secondo il quale le persone più benestanti vivono nelle periferie delle città, mentre i meno abbienti nei centri cittadini. Il paradosso consiste nella differenza dei costi della superficie, infatti questi sono nettamente superiori nei centri città rispetto alle zone periferiche. Secondo W. Alonso quando qualcuno acquista della terra “sta in realtà acquistando due beni tramite una sola transazione: la terra e l'ubicazione”. Esistono due tipologie di beni che vanno bilanciati: quelli negativi (la distanza), che hanno un costo positivo (costo di spostamento) e, al contrario, beni positivi (l'accessibilità) con un costo negativo (il risparmio in termini di costo di spostamento). Sulla base del bilancio di questi costi vengono esaminati differenti modelli, tra cui quello agricolo e quello residenziale. All'equilibrio si

ha che, a seconda del tipo di attore, le curve di indifferenza¹ dell'individuo avranno pendenze differenti. I cittadini più ricchi avranno curve con pendenza minore dei cittadini meno abbienti. Questo avviene perché, in ogni parte della città, i cittadini più facoltosi possono acquistare più terra e, per tale motivo, sono più influenzati dalle variazioni di prezzo dei terreni rispetto al costo dello spostamento, che cresce con la distanza dal centro. Viceversa, i cittadini meno ricchi, hanno la possibilità di comprare meno terra rispetto ai più benestanti e, dato che si stanno considerando piccoli quantitativi di terreno, le variazioni di prezzo della terra non saranno tanto rilevanti quanto il crescente costo del fastidio derivante dallo spostamento; pertanto, si troveranno a vivere in delle zone più costose di quelle in cui risiederanno i più ricchi.

2.1.2 EVOLUZIONE DELLA DENSITÀ URBANA ED ESPANSIONE DELLE CITTÀ

Con lo sviluppo tecnologico, come visto sopra, è stato reso possibile il processo di decentralizzazione, che ha portato ad avere città sempre più disperse. La tendenza alla dispersione urbana può portare a delle situazioni di policentrismo, ovvero casi in cui si ha l'affermarsi di sotto-centri: differenti centri di residenza e occupazione circostanti ad un **“central business district”** (CBD). Ciò che si osserva con la nascita delle città policentriche, è il transito da cittadine interamente attraversabili a piedi in poco tempo, a grandi unità urbane contraddistinte da densità sempre decrescenti.

(S. Angel, J. Parent, D.L. Civico, & A. M. Blei, 2010) nell'articolo *“The Persistent Decline in Urban Densities: Global and Historical Evidence of “Sprawl”*, 2010 trattano il tema dei cambiamenti della densità urbana nel tempo. Quello che gli autori si

¹ Si hanno infatti delle curve di indifferenza in cui si mappa la relazione fra il prezzo di un bene (la terra) con la quantità di un altro bene (distanza dal centro città). Il prezzo che l'attore sarà disposto a pagare diminuirà con l'aumentare della distanza dal centro a un tasso che sia sufficiente a garantirgli un effetto reddito tale da compensare l'aumento del costo di spostamento e il fastidio di un lungo viaggio.

domandano è se esistano, o meno, delle densità urbane ottimali e quali sono i fattori che ne determinano il cambiamento. Per rispondere a queste domande sono state analizzate delle misure di espansione urbana (sprawl). Gli autori hanno confermato la presenza di fenomeni di decentramento urbano, infatti esistono città che non sono caratterizzate da un reale confine, altre che hanno delle densità di popolazione veramente basse, altre ancora nelle quali i limiti tra città e zone rurali non sono distinti ed infine, città policentriche. Lo scopo degli scrittori è quello di comprendere se il fenomeno della decongestione, ovvero del declino della densità urbana (misurato come la diminuzione nel tempo del rapporto fra il totale della popolazione urbana ed il totale dell'area che essa occupa), e della decentralizzazione vadano interrotti o meno, studiare il valore scientifico e i risvolti politici di questa decisione. Per svolgere questo lavoro è stato analizzato il fenomeno della decongestione in venti città statunitensi differenti; grazie a questa analisi vengono proposte una serie di implicazioni rilevanti per la policy di un paese. Fra queste si nota il fatto che "Le fatiche necessarie per rendere le città più dense richiedono di invertire una tendenza piuttosto forte e sostenuta a livello globale, ovvero quella di riduzione di densità" (p. 111) e che "Il tasso di diminuzione della crescita si è ridotto nel tempo, e la densità nelle città dei paesi sviluppati e con territori ricchi presto raggiungerà una stabilità" (p. 116).

Il tema della espansione urbana era già stato affrontato da (A. Anas, R. Arnett, & K. A. Small, 1998), che hanno analizzato la concentrazione spaziale dal punto di vista dell'occupazione e della popolazione urbana. Si ha una distinzione tra concentrazione a livello di città (attività centralizzate o decentralizzate a seconda della vicinanza ad un CBD) o a livello locale (agglomerazione che porta al policentrismo o dispersione). Ciò che deriva dalle analisi è che, con il crescere del reddito disponibile e con la riduzione dei costi di trasporto dovuti al progresso tecnologico, i fenomeni di decentralizzazione sono sempre più frequenti.

2.1.3 LA CITTÀ MODERNA: I POLI-CENTRI E L'AGGLOMERAZIONE

Con l'evoluzione della struttura urbana vi è sempre più la tendenza da parte delle attività economiche ad agglomerarsi tra loro e a posizionarsi in dei centri che ne favoriscano l'interazione. Con il termine "centro" si fa riferimento sia al centro della città, che a dei sotto-centri. (G. Giuliano & K. A. Small, 1991) nel lavoro "*Subcenters in the Los Angeles Region*", 1991, pongono la loro attenzione sull'identificazione dei sotto-centri in una città, nello specifico Los Angeles. La loro analisi parte con la constatazione del fatto che il modello monocentrico, già alla fine del ventesimo secolo, non poteva più essere lo schema urbano che al meglio rappresentava le aree metropolitane; infatti, esistono delle evidenze empiriche che si focalizzano sui conflitti esistenti tra le forze agglomerative e i fenomeni di congestione. Se le prime sono deboli, quando la congestione raggiunge dei valori elevati, le attività si spostano e si decentralizzano, tendendo a disperdersi in tutto lo spazio che hanno a disposizione. Invece, se sono forti, le attività si espanderanno, ma si concentreranno in degli agglomerati secondari o in nodi, chiamati sotto-centri.

Secondo (McDonald, 1987) un sotto-centro è "una zona, nella quale la misura di concentrazione dell'occupazione è superiore rispetto alle zone circostanti". (G. Giuliano & K. A. Small, 1991) concordano sul fatto che non è la concentrazione della popolazione a determinare la formazione dei centri urbani, ma quella dei lavoratori. Secondo gli autori, dunque, un centro è "un insieme continuo di zone, ciascuna con una densità superiore ad una soglia \bar{D} , che abbiano cumulativamente almeno \bar{E} lavoratori totali e, tale per cui, tutte le zone immediatamente adiacenti al di fuori del sotto-centro abbiano densità minore di \bar{D} ".

Dal lavoro di (G. Giuliano & K. A. Small, 1991) si deduce che la struttura urbana è stata rimodellata dal fenomeno agglomerativo in sotto-centri, che ha permesso il passaggio da città monocentriche, molto dense e congestionate, a città

policentriche suddivise in sotto-centri, che hanno portato ad avere uno spazio urbano eterogeneo.

Il fenomeno dell'eterogeneità spaziale viene ripreso da (A. Anas, R. Arnett, & K. A. Small, 1998) secondo i quali, "i luoghi differiscono per diversi fattori quali il terreno, il clima o l'accesso a fonti d'acqua". L'inomogeneità spaziale, può generare sotto-centri o nuovi CBD grazie a specifiche specializzazioni produttive legate alle caratteristiche dell'ubicazione geografica. L'agglomerazione policentrica può essere auspicabile, inoltre, perché porta al raggiungimento di economie di scala, interne ed esterne. Le prime possono emergere in qualche particolare processo produttivo o fase di questo; le seconde riguardano ad esempio la riduzione dei costi di spedizione o logistici. Quando sono presenti economie di scala esterne tra delle imprese si parla di economie di localizzazione; se sono presenti fra più settori si parla invece di economie di urbanizzazione.

Negli ultimi anni sono stati proposti nuovi micro-fondamenti che spiegano le economie di agglomerazione; (Duranton & Puga, 2004) propongono l'esistenza di tre meccanismi alla base delle economie di agglomerazione:

- **meccanismi di condivisione, sharing:** "riguarda la condivisione di entità indivisibili che potrebbero avere elevati costi fissi e che talvolta sono anche beni escludibili²". Lo sharing determina la creazione di una struttura urbana ben definita; la dimensione ottimale, dal punto di vista dell'efficienza, della città è il risultato del tradeoff tra le economie di agglomerazione urbana e l'affollamento urbano: all'equilibrio città simili hanno dimensioni simili e, queste sono efficienti;
- **meccanismi di accoppiamento, matching:** in questo contesto vi sono due fonti di agglomerazione, la prima per cui "Una crescita del numero degli

² bene il cui godimento è riservato solamente ad alcuni individui e non ad altri.

agenti che cercano di eseguire un accoppiamento (match) migliora la qualità attesa di ciascuno di essi” (p. 2086), mentre la seconda dice che “Una maggiore competizione aiuta a risparmiare in costi fissi rendendo l’aumento del numero di imprese meno che proporzionale rispetto alla forza lavoro” (p. 2086). Il matching porta a svariati vantaggi, come ad esempio la possibilità di trovare migliori lavoratori o a quella di ridurre il rischio di holdup³ derivante dall’incompletezza contrattuale;

- **meccanismi di apprendimento, learning:** l’apprendimento non è attività solitaria ma è piuttosto una comunicazione fra svariate persone. Per questo motivo “la città è in grado di collegare un numero significativo di individui facilitando il processo di apprendimento, di generazione della conoscenza e della sua diffusione”.

Un ulteriore spunto di riflessione riguardo il tema della congestione e della dispersione delle città è fornito da (E. L. Glaeser & J. D. Gottlieb, 2006) in “*Urban Resurgence and the Consumer City*”, 2006. Gli autori analizzano la rinascita delle città attraverso l’esame di alcune variabili⁴, considerando, però, a differenza di quanto affermato prima, l’area urbana come luogo di consumo e non più di produzione. Focalizzano, quindi, la loro attenzione sugli aspetti sociali correlati alla vita nei centri cittadini, che portano ad interazioni dapprima negative, come i crimini e la diffusione di malattie, e poi positive risultanti dall’incontro e l’interazione tra le persone.

³ Il rischio di comportamento opportunistico della controparte in presenza di investimento specifico; ovvero il tentativo di appropriazione della quasi rendita altrui. (la quasi rendita è definita come la differenza tra il valore di un bene nel suo uso presente e il suo valore nella migliore alternativa, ovvero il suo costo opportunità)

⁴ la crescita della popolazione (numero di persone che vogliono vivere in città), la crescita del reddito (e quindi della produttività), la crescita del prezzo delle abitazioni (vista come proxy del desiderio che le persone hanno di pagare per ricevere dei servizi)

Vedere la città come luogo di consumo e non di produzione fu possibile a partire dagli anni '80, dato che prima di allora le zone urbane erano un insieme di disamenità, a causa della facilità di diffusione di malattie o la presenza di criminalità. A partire dall'ultimo ventennio del XX secolo, le amenità a disposizione dei cittadini iniziarono ad aumentare (in seguito ad investimenti fatti al miglioramento della qualità della vita) e la tendenza si invertì: iniziò la ripresa dei grandi centri. Il desiderio di pagare per godere dei servizi urbani aumentò più che proporzionalmente rispetto a quello per le amenità nelle periferie; le città divennero luoghi in cui i costi di interazione tra le persone erano inferiori; nacquero nuove forme di consumo, come ad esempio ristoranti, cinema, bar, musei, favorite dalle grandi densità urbane, in grado di coprire i costi fissi di queste attività, dato che generano clienti per queste attrazioni. Era nata la città moderna!

2.2 L'AGGLOMERAZIONE NEL MONDO DEL RETAIL

Si rivolge adesso l'attenzione all'aggregazione all'interno delle città; in particolare si analizza quello che è il comportamento del retailer nelle scelte di localizzazione dentro il contesto cittadino.

Due forze contrastanti guidano i fenomeni di agglomerazione geografica, infatti: da una parte la paura di una eccessiva competizione diretta porta gli attori ad allontanarsi l'un l'altro, dall'altra le imprese tendono ad avvicinarsi tra loro, godendo dello spillover dei concorrenti vicini. Lo spillover, secondo (Marchall, 1890) è il vantaggio che le persone ottengono dallo stare vicino ad altri soggetti con le stesse capacità commerciali, questo fenomeno è talmente elevato che "i misteri delle attività commerciali non divengono più un mistero, al punto che questi paiono essere nell'aria, ed i bambini imparano molti dei segreti in maniera quasi inconscia" (p.225).

Il tema del posizionamento delle imprese è l'argomento centrale dello scritto di (Hotelling, 1929), il quale analizza le scelte di posizionamento di due imprese A e B, che si trovano in una città lineare⁵ di lunghezza l . Le assunzioni alla base del modello della città lineare di Hotelling sono: distribuzione uniforme lungo l degli abitanti della città, e quindi dei possibili compratori; esiste un costo di trasporto, per unità di distanza, pari a c che ogni cittadino deve sostenere per effettuare l'acquisto; la domanda è inelastica e il prodotto acquistato è omogeneo. Da questo modello emerge una posizione ottimale, nella quale le due aziende dovrebbero trovarsi (nel caso in cui vendessero il bene allo stesso prezzo p), che è rispettivamente a $\frac{l}{4}$ e a $\frac{3l}{4}$. La scelta ottimale, però non si verifica, in quanto l'incertezza sul prezzo che il concorrente farà, porterà entrambi gli attori a posizionarsi al centro della città lineare, ovvero in una posizione socialmente subottimale.

La scelta dell'ubicazione della propria attività commerciale, al fine di attrarre il maggior numero di consumatori, è un tema trattato anche in altri scritti, come *"The Law of Retail Gravitation"* di (W. J. Reilly, 1931) e *"A Probabilistic Analysis of Shopping Center Trade Areas"* di (D. L. Huff, 1963). I modelli proposti vengono definiti modelli gravitazionali, nei quali si analizzano i negozi e le rispettive attrattività. Vengono definiti gravitazionali in quanto affermano che la location di un negozio è funzione due variabili, tipiche della teoria gravitazionale: la massa e la distanza. La prima variabile, la dimensione del negozio, viene utilizzata come proxy dell'attrattività dello stesso, infatti maggiori dimensioni del negozio garantiscono una maggiore varietà e, quindi, una più ampia possibilità di scelta. La seconda variabile è rappresentata dalla distanza esistente tra il consumatore e il negozio stesso. Secondo il modello di Reilly maggiore è l'attrattività del negozio (o centro), maggiore è la distanza che il consumatore è

⁵ nella strada principale di una cittadina o lungo una ferrovia "trans-continentale"

disposto a percorrere per raggiungerlo. Il modello di Huff, che rappresenta una variante di quello di Reilly, ammette l'esistenza di più negozi e non solo di due, come nel modello di partenza. Tramite questo modello è possibile analizzare la possibilità che un cliente si rechi in un punto vendita in funzione della presenza di altri negozi adiacenti e pertanto, è possibile effettuare una previsione più accurata delle decisioni dei consumatori.

Il comportamento del retailer è influenzato dalle motivazioni che guidano le scelte di posizionamento. A tal proposito, sono stati studiati alcuni papers che analizzano l'argomento.

(A. Wolinsky, 1983) in *"Retail Trade Concentration due to Consumers' Imperfect Information"* afferma che le scelte dei consumatori sono affette da informazione imperfetta, ovvero non conoscono realmente l'offerta a loro disposizione. A causa di ciò, ogni singolo attore sente l'esigenza di "scoprire il mercato" e a seguito di questa azione occorre sostenere, ovviamente, un costo diverso da zero. Il consumatore, allora, cercherà di massimizzare il beneficio che si attende di ottenere e questo può portare alla preferenza di effettuare il consumo in un luogo in cui sono presenti più negozi, seppur più distanti per il soggetto (A. Wolinsky, 1983) con il suo scritto confermò quello che (D. L. Huff, 1963) aveva affermato nel ventennio precedente). Questa scelta, dipende dal fatto che il cliente è consapevole di dover "scoprire il mercato" e, questo, lo porta a preferire un luogo dove la probabilità di ricercare ciò che desidera è maggiore; questo avviene in centri più ampi e più lontani. I retailers, quindi, consapevoli di questi atteggiamenti dei consumatori, tenderanno a posizionarsi l'uno vicino all'altro. Il luogo dove si crea questa concentrazione spesso risulta essere quello non ottimale per la minimizzazione dei costi di tutti gli attori in gioco.

Anche (M. Dudey, 1990) affronta l'argomento, affermando che "le decisioni di posizionamento delle imprese sono influenzate dal comportamento dei consumatori, i quali sono attratti dalle zone occupate da un numero

relativamente grande di aziende perché essi si aspettano una più alta competizione e di conseguenza prezzi più bassi". Quindi, la concentrazione geografica è di maggiore attrattività per i consumatori, in quanto permette la comparazione di prezzo. Egli propone un modello, secondo cui "c'è un equilibrio nel quale tutte le imprese si posizionano nello stesso centro" (p. 1095). È paradossale, quasi, il fatto che le imprese decidano di aggregarsi nello stesso centro in modo tale da semplificare la comparazione di prezzo da parte dei clienti. La scelta della non concentrazione sarebbe quella seguita solo nel caso in cui la competizione di prezzo fosse troppo elevata.

Il lavoro di (Fox, et al., 2007) "*The Impact of Retail Location on Retailer Revenues: An Empirical Investigation*" analizza le scelte di posizionamento dei retailers in una maniera differente da coloro che lo hanno preceduto. Secondo (Fox, et al., 2007) "il fenomeno di concentrazione geografica può assumere due forme diverse: se si considera l'agglomerazione di differenti tipologie di negozi si parla di agglomerazione **inter-type**, quando invece si osservano cluster di negozi dello stesso tipo ci si riferisce ad agglomerazione **intra-type**⁶". Inoltre, il paper afferma che le scelte di localizzazione dei negozi possono essere analizzate da due punti di vista differenti, infatti: i retailers possono decidere di posizionarsi vicino ai clienti o vicino agli altri negozi. Vi sono altre variabili che il negoziante deve analizzare: l'influenza sul suo reddito di variabili demografiche, dei tempi di percorrenza necessari e la presenza di agglomerazione. Dall'analisi degli autori su differenti tipologie di negozi, emerge che "il tempo necessario perché un consumatore raggiunga un negozio è forte perditore della sua spesa in tale

⁶La co-localizzazione di attività deriva dal bilanciamento di due forze tra loro contrastanti: la simbiosi e il darwinismo. La simbiosi è intesa come "l'attrattività incrementale che riflette la riduzione dei costi di ricerca dei consumatori, che li porta all'acquisto congiunto di più beni; conseguenza di questo è che semplici gruppi di negozi divengono più complesse destinazioni d'acquisto". Il darwinismo inteso, non solo come, "la competizione per il consumatore tra i negozi che vendono beni simili, ma anche la rivalità più generale per la risorsa scarsa comune ovvero il reddito disponibile del consumatore"

negozio” e che gli effetti che si ottengono sui differenti retailers non sono simmetrici; quest’ultima affermazione ammette l’esistenza di alcune categorie merceologiche che beneficiano maggiormente della vicinanza di altri negozi, di altre che preferiscono la presenza solo di determinate tipologie di retailers ed altre ancora che decidono di isolarsi. Pertanto, le scelte ottimali di posizionamento all’interno del tessuto urbano varieranno a seconda del tipo di negozio.

2.2.1 COME IMPATTA LE DECISIONI LA TIPOLOGIA DI PRODOTTO VENDUTO

Esattamente come il consumatore avrà degli atteggiamenti differenti a seconda del prodotto acquistato, anche il retailer effettuerà delle scelte di posizionamento diverse a seconda del prodotto venduto.

Una prima distinzione necessaria tra le tipologie di prodotti è stata proposta da (Nelson, 1970), il quale distingue i beni a seconda che essi “abbiano una qualità determinabile dal consumatore a priori, prima dell’acquisto”, i **search goods**, o “a posteriori, dopo l’acquisto”, gli **experience goods**. Da questa classificazione emerge che i negozianti trarranno beneficio dall’agglomerazione solo quando il consumatore riterrà necessaria la comparazione tra beni differenti, quindi nel caso di experience goods.

Un ulteriore classificazione dei beni è quella in funzione del livello di coinvolgimento tra consumatore e prodotto di (Holton, 1958). Esistono beni **convenience**, per i quali “il vantaggio derivante dalla comparazione di prezzo e qualità è piccolo rispetto alla stima dei costi di ricerca in termini di tempo, danari e fatica”; appartengono a questa categoria i beni non durevoli con domanda costante. Esistono, poi, beni **shopping**, per i quali “il vantaggio derivante dalla comparazione di prezzo e qualità tra le alternative proposte dal venditore è grande rispetto alla stima che il consumatore fa dei costi di ricerca in termini di tempo, soldi e fatica” (beni durevoli). Infine, i beni o marche **speciality**

caratterizzati da bassa offerta e alto numero di persone interessate all'acquisto; la ricerca di questa tipologia di prodotti richiede dei costi, in termini di tempo e denaro, molto superiori rispetto a quelli dei beni shopping.

Quindi si nota come le scelte agglomerative dei retailers variano parecchio a seconda del bene trattato. (G. A. Picone & D.B. Ridley, 2008), in *"Distance decreases with differentiation: Strategic agglomeration by retailers"*, e (Sevtzuk, 2014), in *"Location and Agglomeration: The Distribution of Retail and Food Businesses in Dense Urban Environments"* approfondiscono questo tema in due loro scritti. Nel primo gli autori propongono un modello che predica le scelte agglomerative dei negozianti in funzione del grado di differenziazione del prodotto. Esiste una relazione negativa tra "il grado di differenziazione del prodotto e la profittabilità della distanza geografica: se i prodotti sono meno differenziati, è più profittevole per i negozianti essere distanti così da ridurre la competizione, se al contrario i prodotti sono più differenziati i benefici dalla distanza diminuiscono fortemente".

Nel secondo paper, (Sevtzuk, 2014) propone un modello che studia i negozi raggiungibili a piedi in un raggio di 600 metri dal centro città. Egli studia diverse ipotesi, mediante un'analisi di regressione; la più importante delle quali è: "I retailers che vendono search goods, quali vestiti o ristoranti tenderanno ad agglomerare maggiormente di coloro che vendono convenience good, quali alimentari e liquori" (p. 389). Infatti, come detto precedentemente, i search goods sono dei beni più inclini alla comparazione, a differenza dei beni convenience, che non beneficiano dei vantaggi del "touch 'n feel it". Dalla sua analisi empirica emerge come le scelte di localizzazione siano significativamente correlate all'ubicazione dei concorrenti e al tipo di bene venduto: ad esempio i ristoranti e i bar sono una tipologia di esercizi commerciali che beneficiano in maggior misura della comparazione dei menù, e quindi tenderanno ad agglomerarsi;

viceversa i supermercati o i negozi che vendono merce di varia natura, di tipo convenience, saranno meno propensi alla concentrazione geografica.

2.2.2 *COME IMPATTA LE DECISIONI LA PRESENZA/ASSENZA DI AMENITÀ*

I servizi e le amenità hanno un ruolo nella scelta del posizionamento del retailer. (E. L. Gleaser, J. Kolko, & A. Saiz, 2001) in *“Consumer City”* osservano che si ha una crescita superiore in quelle città in cui è presente un maggior numero di servizi, viceversa, tutti i centri urbani con una minore concentrazione di amenità la crescita è inferiore. Considerata questa constatazione, le città del futuro sarebbero soggette alla capacità delle aree urbane di fornire sempre più servizi attrattivi a cittadini sempre più istruiti e abbienti. Esistono quattro principali amenità che guidano tale fenomeno:

- La presenza di una ricca varietà di servizi e beni di consumo: esistono dei particolari tipi di beni, definiti *“locali”*, in quanto parecchio correlati al territorio in cui si trovano (ristoranti, cinema, teatri...), tali per cui al crescere del numero si ha una maggiore crescita della città;
- Estetica e dotazioni materiali: con questa amenità ci si riferisce alla bellezza architettonica del luogo, alle condizioni climatiche della città e alle dotazioni che essa possiede;
- Buoni servizi pubblici: la crescita delle città è correlata anche alla qualità dei servizi pubblici offerti, come ad esempio la scuola o il controllo da parte delle forze dell'ordine;
- Velocità: un servizio è affidabile quando è facilmente fruibile e di semplice utilizzo da parte dei cittadini. Ci si riferisce in particolare alla facilità con cui i soggetti riescono a spostarsi all'interno della città, quindi alla velocità dei trasporti.

Gli autori confermano il paradigma della città moderna, intesa come luogo di consumo e non di produzione; il centro cittadino diventa un luogo in cui è

richiesta la presenza del maggior numero di servizi e beni di consumo, mentre le periferie diventano il vero fulcro produttivo.

All'affermarsi del nuovo paradigma, segue un numero sempre crescente di retailers che decidono di posizionarsi nei centri urbani. (C. A. Hidalgo & E. E. Castaner, 2015) in *"The Amenity Space and The Evolution of Neighborhoods"*, 2015 indagano "le motivazioni che guidano le scelte di posizionamento di amenità differenti e su quale sia la probabilità che due attrattività scelgano di posizionarsi vicine". Analizzando tutti i centri di agglomerazione presenti all'interno della città (sotto-centri), gli autori affermano che si ha la creazione di una rete di amenità, l'"Amenity Space", che permette di comprendere quale combinazione di servizi predica la presenza di altri. Così, è possibile prevedere e fornire raccomandazioni su quali nuove attività commerciali nasceranno in una zona in funzione di quelle già presenti.

2.2.3 UNO SGUARDO AI MODELLI DI AGGLOMERAZIONE

In questa sezione si volge lo sguardo a quelli che sono i principi che portano le imprese ad agglomerarsi e a come poter valutare questa concentrazione.

(Marshall, 1890) nel suo lavoro cercò di comprendere quelli che fossero i fondamenti economici alla base della concentrazione di imprese, arrivando a definirne tre: "interazione del mercato del lavoro, legami tra fornitori e distributori per beni intermedi e/o finali e, infine, spillover di conoscenza". Infatti, "La prossimità di imprese comporta la riduzione dei costi di trasporto nella movimentazione di persone, beni, e idee il che permette, fra le varie cose, di ridurre costi di spedizione."

Consideriamo adesso, due papers di (G. Ellison & E. Glaeser, 1994/1997), *"Geographic Concentration in U.S. Manufacturing Industries: a Dartboard Approach"*, nei quali, si cerca di carpire come i differenti settori si posizionano, in quali

circostanze le aziende presentano dei livelli di agglomerazione superiori a quelli che si verificherebbero se le scelte di collocazione geografica fossero del tutto casuali. A tal proposito affermano che “un buon indice di concentrazione geografica dovrebbe non solo misurare aspetti economicamente interessanti ma dovrebbe permettere il paragone su più settori e/o in lassi temporali differenti anche se questo è piuttosto complicato a causa della diversità dei dati a disposizione”. Secondo gli autori, le scelte di posizionamento sono guidate da due principali motivazioni:

- lo **spillover**, tecnologico e di conoscenza, che porta le imprese a concentrarsi
- il **vantaggio naturale**, che porta le imprese a localizzarsi ovunque vi sia un “*vantaggio*”, inteso ad esempio in termini di accesso a materie prime o clima favorevole.

Considerate queste due motivazioni, (G. Ellison & E. Glaeser, 1994/1997) propongono un indice “in grado di identificare la quota di agglomerazione derivante da queste due scelte e la quota derivante dal puro posizionamento casuale”. Occorre prestare attenzione, però, al fatto che l’agglomerazione “casuale” in realtà è impattata anch’essa da alcuni fattori. Infatti è dimostrato che “la concentrazione geografica delle imprese nelle diverse aree segue la concentrazione della categoria merceologica/settore stessa/o”; quest’ultima è analizzabile mediante l’indice di Herfindahl Hirschman. Inoltre, l’agglomerazione geografica di categoria è impattata anche dalla concentrazione della città. È indispensabile, pertanto, che gli indici di agglomerazioni nettino per l’effetto centro città, ovvero quell’esito tale per cui la maggiore concentrazione, e presenza, di negozi in centro città è semplicemente riconducibile al fatto che esse si trovano in centro (l’indice di G. Ellison ed E. Glaeser infatti elimina dalla concentrazione reale, la quota attribuibile al caso).

In tempi moderni, (Alcàcer, 2006) indagò nuovamente i temi affrontati da (Marchall, 1890), affermando che le scelte di localizzazione delle imprese sono guidate da due forze contrastanti: la competizione, che le porta a distanziarsi le une dalle altre e l'agglomerazione che le avvicina. "Queste forze possono avere delle entità differenti a seconda del tipo di attività svolta dall'impresa (ricerca e sviluppo, produzione o vendita) o delle competenze possedute dall'azienda, in termini di vantaggio assoluto di costo". Vi sono delle aziende, distanti dal mercato del prodotto⁷, che saranno meno impattate dai costi di competizione diretta e, quindi, tenderanno ad essere geograficamente più agglomerate; viceversa, altre attività beneficiano della concentrazione e, pertanto, creeranno dei clusters. Alla stessa maniera distinguendo in base alle capacità possedute, vi saranno aziende più capaci che beneficiano meno dei vantaggi dell'agglomerazione ed altre, meno capaci, che preferiranno posizionarsi vicino ad altri, in modo da poter usufruire il più possibile delle loro informazioni. (Alcàcer, 2006), quindi, constatò che i meccanismi esposti da Marshall, in particolar modo lo spillover, impattano in modo differente le scelte di collocazione geografica a seconda del contesto e dell'azienda di cui si sta parlando.

Il lavoro di (Marchall, 1890), fu ripreso ancora una volta da (E. Glaeser, G. Ellison, & W. R. Kerr, 2010) che in *"What Causes Industry Agglomeration? Evidence from*

⁷ Ad esempio esistono evidenze, tali per cui attività di R&S, basate su conoscenze tecnologiche e skills di alto livello, sono lontane dalla competizione diretta del mercato e, pertanto, sono meno soggette a quei comportamenti strategici, quali deterrenza all'entrata tramite eccesso di capacità (Spence, 1977) o minacce di uscita tramite prezzi predatori (Milgrom & Roberts, 1982) atti a ridurre la competizione su mercati caratterizzati da prodotti omogenei. Viceversa, le attività di vendita soffrono in maggior misura la competizione, e quindi tenderanno a disperdersi geograficamente.

Coagglomeration Patterns” confrontano le motivazioni classiche e moderne di agglomerazione di imprese. Secondo gli autori i vantaggi dell’agglomerazione e della co-agglomerazione derivano dalla riduzione dei costi di trasporto, in varie forme:

- **Prossimità a clienti, fornitori e beni:** le imprese di co-agglomerano in modo da poter risparmiare sull’approvvigionamento degli input e, quindi, sul costo di spedizione dei prodotti;
- **Condivisione del mercato del lavoro e delle persone:** per ottenere economie di scala è necessario l’utilizzo di una grande mole di lavoro, a cui si può avere accesso in modo più semplice se i lavoratori hanno la giusta mobilità per spostarsi tra settori differenti;
- **Spillover tecnologico o intellettuale, di idee:** è complicato comprendere come le idee emergano, ma il processo è più semplice se le attività di R&S possano fluire da un settore all’altro. le idee emergono, tuttavia si cerca di valutare come le attività di R&S fluttuino da un settore all’altro;
- **Vantaggio naturale:** due aziende derivanti da settori differenti si co-agglomerano perché ottengono un qualche vantaggio dalla stessa fonte naturale.

Conclusa la rassegna della letteratura relativa alla concentrazione urbana e su cosa impatta le scelte agglomerative di un retailer, si passa ad analizzare il mondo del Retail e come quest’ultimo è cambiato negli anni.

3. IL MONDO DEL RETAIL: TRASFORMAZIONI IN CORSO

3.1 COSA RAPPRESENTA L'ATTIVITÀ DI RETAILING

Per comprendere l'importanza del settore di cui si sta parlando, basti pensare al fatto che l'attività di retailing è alla base sia della nostra economia, che della nostra vita quotidiana in generale. I consumatori e le aziende acquistano quotidianamente prodotti e servizi, ma è la nostra familiarità con il mondo dello shopping che porta questo settore ad essere così rilevante.

L'attività del commercio consiste nello scambio di beni e/o servizi, solitamente in cambio di denaro. Questo concetto di scambio, fornisce le basi su cui la nostra economia è basata. I retailers, sono coloro che vendono beni al pubblico in relativamente piccole quantità, per uso di consumo. Nel "*Retail Market Monitoring Report*", la (Commissione Europea) si riferisce al settore in questione con il seguente approccio: "Il servizio del Retail comprende un'ampia varietà di forme (negozi, commercio elettronico, mercati, etc.), formati (dai piccoli negozietti agli ipermercati), prodotti, strutture legali (negozi indipendenti, gruppi di retailers indipendenti, franchises, gruppi integrati, etc.) e locations (urbane/rurali, centro città/suburbane etc.)".

Statisticamente, sempre secondo il *Retail Market Monitoring Report* della (Commissione Europea), i retailers sono classificati nelle seguenti categorie: quelli che vendono in dei negozi non specializzati (supermercati o negozi generici), quelli che vendono in negozi specializzati (come ad esempio tessile, cibo, audio&video, libri, etc.), quelli che vendono beni di seconda mano in negozio, quelli che vendono in stalli o mercati e quelli che vendono via mail o internet.

Le capacità competitive distintive di un retailers derivano dall'abilità superiore di produrre una gamma di beni che i consumatori vorranno acquistare e offrirli sul mercato nel giusto modo, nel giusto luogo e al giusto prezzo.

La rilevanza di questo settore può essere compresa facilmente dando uno sguardo ai numeri, in particolare si fa riferimento al Report 2018, pubblicato da (Deloitte, 2018), "*Global Power of Retailing, 2018*" nel quale vengono prese in considerazione le 250 più grandi aziende retailer al mondo (Top 250), analizzate le loro performance tra le varie geografie e settori di prodotto differenti. Nel report vengono mostrate delle rapide statistiche dei Top 250, che indicano l'ampiezza del fenomeno a livello mondiale:

- Crescita del fatturato retail in 5 anni, misurato attraverso il CAGR (Compound Annual Growth Rate) dal 2011 al 2016: **4,8%**;
- Fatturato Retail aggregato: **US\$4,4 trillion**;
- Fatturato medio, ad indicazione della grandezza media delle Top 250: **US\$17,6 billion**;
- Margine di profitto netto composto: **3,2%**;
- Numero medio di paesi in cui la singola azienda Retail opera: **10**.

3.2 I VARI FORMATI DEL RETAIL

Tutti i businesses che vendono beni o servizi ai consumatori stanno sotto l'ombrello del "retailing", ma, nella realtà, esistono differenti direzioni che possono essere intraprese. Per iniziare, c'è quella del Department Store, del Discount Store, dello Specialty Store e anche quella del retailer stagionale.

Senza entrare nello specifico della categoria di prodotto venduto, in generale, i segmenti possono essere suddivisi in:

- **Hard:** questi tipologie di beni includono elettrodomestici, elettronica, mobili, articoli sportivi, etc. Questa tipologia di rivenditori vengono definiti “hardline retailers”.
- **Soft:** queste categorie includono apparel, vestiti ed altre fabbriche.

Ogni retailer prova a differenziare sé stesso dalla competizione, ma è la strategia che le aziende utilizzano per vendere i propri prodotti a rappresentare il fattore di maggiore rilevanza. Esistono differenti tipologie di retailers:

- **Department Stores:** veramente grandi negozi con un enorme assortimento di beni e servizi.
- **Discounters:** vendono anch’essi un ampio insieme di prodotti e servizi, ma essi competono maggiormente sul prezzo.
- **Demographics:** sono retailers che non agiscono in un particolare segmento.

Ciascuna tipologia ha i suoi vantaggi competitivi distintivi, ma è fondamentale comprendere come si presentano quest’ultimi. Ad esempio, durante i periodi di crisi economica, i retailers di tipo Discount tendono a sovra-performare gli altri; l’opposto si verifica quando l’economia prospera. I retailers di maggior successo cercano di combinare le caratteristiche delle differenti tipologie per differenziarsi dalla competizione.

3.2.1 GIGANTI CHAIN STORES VS MOM-AND-POP STORES

Un’ulteriore distinzione che può essere effettuata, particolarmente utile ai fini dell’analisi di cui si tratterà successivamente, è quella tra i “Chain Stores” e i “Mom-and-Pop Stores”.

Con Chain Store si intende l’insieme di due o più negozi Retail che hanno la stessa proprietà e vendono la stessa tipologia di beni. I Chain Stores rappresentano un importante segmento di Retail in America, nell’Europa

dell'Ovest e in Giappone. Insieme ai Department Store e alle aziende mail-order, i Chain Store rappresentano la prima applicazione di successo di retailing dei metodi integrati di ampia scala. I regolari Chain Store devono essere differenziati dal Franchising e dalle catene di negozi volontarie o in cooperativa, in cui la singola unità di retail preserva la propria proprietà, ottiene i propri profitti e supporta finanziariamente le proprie perdite.

Con Mom-and-Pop Store si intende, invece, un negozio piccolo, indipendente, solitamente di proprietà familiare, che opera in un business impiegando il minimo numero di lavoratori e producendo piccole quantità di volumi, non opera in franchising e concentra il proprio business in una singola location. Questa tipologia di negozi ha molto da offrire: fedeltà e devozione da parte dei membri del team e scambi sinergici di nuove idee e conoscenze tra le diverse generazioni. Tuttavia, vi sono anche numerose sfide che le aziende familiari devono affrontare, come quella di chiarire i ruoli e le responsabilità, infatti la mancata delega dei compiti e la non corretta definizione dei ruoli dei dipendenti può spesso generare confusione in aziende di questo tipo, o quella di definire la ripartizione dell'equity.

La domanda a cui, al giorno d'oggi, occorre cercare di rispondere è: *Cosa accade a piccoli businesses in un'economia di libero mercato? Sopravvivono o vengono sopraffatti dalle grandi Corporations?*

Nel 1997, l'economista (LaFaive, s.d.), rispose affermando che le piccole imprese non soltanto sopravvivono, ma prosperano in un ambiente di libero mercato e, in particolare affermò che: "in un'economia di libero mercato, le grandi catene di negozi collocano i negozi più piccoli nella posizione di dover essere maggiormente competitivi; è vero che non tutte le piccole imprese riescono a sopravvivere a questa competizione, ma per effettuare un'analisi economica corretta occorre considerare l'aumento di produttività in termini netti".

Nell'ultimo ventennio, i così chiamati "Big Box Store" hanno trasformato l'ambiente del Retail nei mercati emergenti; secondo una nuova ricerca condotta dall'istituto Chazen, l'arrivo della vendita organizzata nelle economie in via di sviluppo ha contribuito a rafforzare la rete di negozi Mom-and-Pop, che ancora dominano il settore. "Dato che la vendita organizzata ha funzionato così bene in alcuni mercati emergenti, c'era la sensazione che si sarebbe verificata la morte per la vendita al dettaglio non organizzata. Ma l'impatto della vendita non organizzata non è uniforme nei diversi paesi ed economie, dipende dal contesto, dall'infrastruttura di vendita al dettaglio, dalle opzioni di finanziamento per i retailers non organizzati e da fattori dal lato della domanda, come le abitudini di consumo" spiega (Jerath, s.d.), professore associato alla Columbia Business School.

L'esistenza di una classe media in crescita, combinata con politiche di investimenti esteri diretti (FDI) sempre più liberali, ha aiutato la nascita di giganti internazionali, come Walmart e Carrefour, nei paesi in via di sviluppo. Come previsto, questo rastrellamento del settore, ha costretto alcuni retailers non organizzati e poco competitivi ad uscire dal mercato. Tuttavia, ciò che non era previsto, era che i rivenditori locali sopravvissuti avessero beneficiato delle efficienze di mercato implementate dalle grandi catene, come ad esempio l'accesso a più prodotti e a sistemi di distribuzione migliori. Con prezzi inferiori e minore concorrenza, i retailers locali possono proporre prezzi più alti e godere di margini di profitto più elevati.

I rivenditori organizzati e non, infatti, servono i consumatori in modi molto diversi. Una caratteristica importante dei negozi a conduzione familiare è che operano su una scala estremamente ridotta, servendo le famiglie che vivono a pochi passi dalla loro location; di conseguenza, i negozianti locali conoscono intimamente i loro clienti, sanno cosa, quando e quanto acquistano un determinato prodotto in una determinata settimana e, grazie a ciò, possono

mantenere uno stretto controllo sul livello di inventario, offrendo al tempo stesso un servizio altamente personalizzato al loro cliente.

D'altra parte, mentre i rivenditori organizzati possono offrire una più ampia varietà di prodotti a prezzi più bassi, arrivare al negozio comporta spesso costi aggiuntivi per il consumatore medio, come ad esempio le spese di trasporto e il tempo aggiuntivo per effettuare il viaggio; esistono poi, degli altri costi da tenere in considerazione, come ad esempio quelli relativi agli sprechi, infatti acquistando i prodotti alla rinfusa si finisce per acquistare più prodotti di quelli di cui effettivamente si necessita, aumentando il rischio che parte del prodotto venga sprecato, aumentandone il costo unitario. Questi fattori lato domanda, influenzano il modo in cui i consumatori decidono di acquistare e, di conseguenza, il modo in cui i retailers, organizzati e non, prendono le loro decisioni in materia di immissione sul mercato e pricing.

I policymakers possono mettere in atto delle azioni, al fine di bilanciare la crescita di entrambe le tipologie di retailers nei mercati emergenti. Potrebbero, ad esempio, elaborare politiche che consentano investimenti esteri diretti e gradualmente, in modo da offrire ai negozianti non organizzati il tempo per emulare le migliori pratiche di quelli organizzati e creare un'infrastruttura efficiente della catena di approvvigionamento. Un altro fattore che i policymakers dovrebbero tenere in considerazione è che chi decide di acquistare da retailers non organizzati incorre in prezzi maggiori, viceversa gli altri rischiano lo spreco. A causa di questi costi più alti per il consumatore, non è scontato che l'avanzata della vendita al dettaglio organizzata nei mercati emergenti sia necessariamente positiva per l'economia globale, pertanto le politiche future dovrebbero attentamente considerare questi impatti.

In un periodo difficile per la vendita al dettaglio, con tanti rivenditori che lottano per rimanere a galla, è facile pensare che siano i negozi più piccoli, Mom-and-Pop, ad avere la peggio. Tuttavia, la recente scomparsa di giganti retailer, come

ad esempio Abercrombie and Fitch, ha lasciato un vuoto nel mondo del bricks-and-mortar da riempire. In questa circostanza, molti esperti affermano che potrebbe essere meglio rimanere piccoli; essere un'azienda di dimensioni microscopiche non è certamente una protezione contro i "Big-Box" retailers o i rivenditori online, ma essere parte integrante di una comunità locale può aiutare a costruire una base clienti fedele.

Altre dinamiche stanno dando ai negozi "Mom-and-Pop" un vantaggio, secondo gli esperti del settore, infatti, i consumatori, nei tempi recenti, sono sempre più interessati agli articoli artigianali e la produzione locale è sempre più percepita come di qualità; cercano anche esperienze più autentiche e uniche, che i piccoli negozietti sono in grado di offrire. Inoltre, la tecnologia ha permesso a quest'ultimi di ottenere visibilità in tutto il mondo, senza espandere la propria impronta, sfruttando la potenza dei social network e del marketing mirato.

Il potentissimo impatto della tecnologia sarà argomento del prossimo paragrafo.

3.3 RETAIL TRENDS: PROFONDA TRASFORMAZIONE IN ATTO

Il settore del Retail è fortemente dinamico e sta subendo una profonda trasformazione. Ci troviamo in un mondo in cui l'acquirente è all'interno del posto di guida, grazie ad una tecnologia che gli consente di essere "costantemente connesso", e, mai così tanto prima d'ora, ha il potere di guidare i cambiamenti nel mondo dello shopping. L'"Everywhere Commerce" è un concetto ormai sedimentato, consentendo al consumatore di fare acquisti nella modalità che più desidera, ovunque, in qualsiasi momento, e come preferisce: in negozio, online, da dispositivi mobili, tramite comando vocale o mediante il cosiddetto "click-and-collect".

Nell'industria Retail, la distruzione del tradizionale modello di business ha portato a delle trasformazioni e a dei cambiamenti senza precedenti;

cambiamenti richiesti sia online che offline, al fine di servire al meglio la domanda degli acquirenti e di ridefinire la “customer experience”. Queste innovazioni e trasformazioni stanno avvenendo più velocemente e con maggiore intensità di sempre, rappresentando delle sfide per i retailers, abituati a bilanciare le convenzionali metriche di performance come crescita, profittabilità e produttività dello spazio. In tal senso, gli standard stanno cambiando, poiché alcuni dei più importanti retailers del mondo (riconosciuti come distruttori del settore, come Amazon e JD.com) stanno rinunciando alla profittabilità di breve termine, mirando invece all’acquisizione di nuovi clienti, all’espansione top-line e al dominio del settore. Questo ha portato retailers affermati da tempo al rischio di perdere i propri clienti e quote di mercato e cederle a questi “distruttori”, maggiormente in grado di gestire e sfruttare l’agilità organizzativa e operativa.

Mentre i negozi stanno chiudendo, la spesa del settore Retail si sposta verso l’online a ritmi vertiginosi. Nel 2017 gli Stati Uniti hanno registrato un numero record di chiusure, con 6885 negozi che hanno chiuso i battenti al primo dicembre. I negozi di tutto il mondo si trovano di fronte ad un simile destino, in quanto i rivenditori chiudono i negozi non profittevoli, per focalizzare le loro mosse verso locations più produttive e promettenti.

Le regole del commercio al dettaglio vengono riscritte in questo periodo di cambiamenti trasformativi. L’innovazione, la collaborazione, il consolidamento, l’integrazione e l’automazione sono elementi necessari per rinvigorire il commercio, influenzando profondamente il modo in cui i retailers fanno affari sia nel presente, che nel futuro.

Secondo il “Global Power of Retailing, 2018” di (Deloitte, 2018), vi sono delle competenze distintive che i retailers devono sviluppare per poter competere ed aver successo al giorno d’oggi:

- **Costruire capacità digitali a livello mondiale:** i retailers di tutto il mondo si stanno velocemente adattando all’idea che, dal punto di vista del consumatore, lo shopping non è solo “bricks versus clicks” o un canale piuttosto che un altro. Piuttosto, i consumatori vogliono essere channel-agnostic, ovvero indipendenti dal canale. L’esperienza di shopping e la ricerca pre-shopping sono un processo fluido su cui i consumatori rimbalzano, tra online ed offline, durante l’esperienza d’acquisto. Occorre prestare attenzione all’influenza che il digitale ha sulla spesa dei consumatori, infatti, nel report del 2016 “*The New Digital Divide*”, Deloitte ha scoperto che **l’interazione digitale influenza 56 centesimi di ogni dollaro speso nei negozi bricks-and-mortar**, a fronte dei 36 centesimi di solo tre anni fa. Quindi, le persone che acquistano utilizzando differenti metodi, incluso online, mobile e visita ai negozi, spendono più del doppio di quelli che acquistano solamente presso gli stores fisici. Questo significa che i retailers devono “*pianificare, strategizzare ed eseguire*” in modo adeguato ed olistico tutti i canali, indipendentemente dal fatto che la vendita avverrà in negozio o online (**Omni-canalità**). Infatti, un’esperienza di shopping senza soluzione di continuità non è più un piacere, ma un imperativo, ed è questa la ragione per cui i retailers di tutto il mondo stanno ampiamente investendo sull’online e sul digitale.
- **Combinare “bricks and clicks” per recuperare il tempo perduto:** il resto del mondo del Retail non sta pigramente seduto a guardare Amazon che irrompe nelle vendite e ruba quote di mercato; molti players, che potrebbero essere stati inizialmente in disparte, stanno cercando di recuperare il tempo perso in gran stile. Un recente studio ha rilevato che le vendite di prodotti alimentari attraverso canali online sono aumentate del 30 per cento nell’ultimo anno. Ne è un esempio, il più grande retailer al mondo, Wal-Mart, ha chiarito che l’e-commerce è uno dei pilastri strategici su cui l’azienda si basa; sta, difatti, spendendo miliardi di dollari

in investimenti di capitale per introdurre la Grocery Online, aumentare le sue capacità di “click-and-collect” e sfruttare la sua ampia rete di negozi per sposare online ed offline e guadagnare un vantaggio su Amazon.

- **Creare un’esperienza unica ed affascinante all’interno dei negozi:** i retail stores fisici non sono stati abbandonati, infatti il **90% delle vendite** Retail di tutto il mondo avvengono ancora presso il negozio fisico. Per competere con l’assortimento unico ed infinito offerto online, però, è cruciale la creazione di una significativa “customer experience” e del “brand engagement” (Apple Store e Nike retailer sono considerati in tal senso come gli standard d’oro da seguire). I retailers bricks -and-mortar stanno comprendendo l’importanza di creare merchandising di prodotti unici e curati, un’atmosfera emozionante e divertente e livelli di servizio simili a quelli che il consumatore può trovare online. Un buon esempio è rappresentato dal settore Grocery, infatti i negozi di generi alimentari stanno transitando da rivenditori di beni a fornitori di servizi e soluzioni, con cibo, salute e benessere che convergono in un ambiente unico. Allo stesso modo, è d’obbligo menzionare quanto velocemente i Fashion Retailers di tutto il mondo stanno continuando a distruggere il settore dell’abbigliamento. La spagnola Initex (Zara), la svedese H&M e la giapponese Fast Retailing (Uniqlo) hanno incrementato le vendite ad un ritmo annuale a due cifre, in media, negli ultimi cinque anni, hanno ridotto il ciclo della moda a circa cinque settimane, comparato con i sei/nove mesi dei rivenditori tradizionali e forniscono al cliente la gratificazione immediata di poter acquistare capi alla moda che li contraddistinguono.
- **Reinventare la vendita con le ultime tecnologie:** poche volte nella storia, i rapidi progressi della tecnologia e le innovazioni rivoluzionarie, avevano avuto la capacità di irrompere in modo così rapido e omnicomprensivo nel mondo del Retail distruggendone il modello di business. Se non sono già presenti, l’“Internet of Things”, l’intelligenza artificiale, la realtà

aumentata e virtuale (AR/VR) e i robots dovrebbero essere sul radar di ogni retailers. Questo tipo di tecnologie possono essere utilizzate come strumenti per elevare il business e migliorare la *customer relationships* sia da parte dei rivenditori fisici che online. Ad esempio, queste tecnologie all'avanguardia vengono utilizzate all'interno degli stores per migliorare e personalizzare l'esperienza d'acquisto e guidare il traffico all'interno del negozio. Forse uno degli usi più progressivi della tecnologia e dell'automazione che sta emergendo è nell'emergere di negozi non presidiati: il "Grab and Go" shopping è adesso una realtà, grazie alla tecnologia di pagamento "mobile". È il caso di AmazonGo, dove, nonostante l'esecuzione sia ancora nelle fasi iniziali, i consumatori possono già visitare il negozio, eseguire la scansione automatica degli oggetti mediante un'app per smartphone, quindi semplicemente pagare tramite telefono ed uscire!

Dopo aver descritto l'attività di Retailing e compreso quanto essa sia importante all'interno della nostra economia, sono stati analizzati i cambiamenti in atto all'interno di questo mondo e quali sono le mosse "paracadute" che i rivenditori dovrebbero affrontare per rimanere all'interno del business e non essere schiacciati dai new entrant.

4. DATA SET & TOOLS

Si passa adesso alla descrizione dei dati, che rappresentano la base empirica del lavoro, e degli strumenti (software) utilizzati per l'analisi, quali QGis e SPSS.

4.1 ORIGINE DEI DATI

Lo studio empirico effettuato, si focalizza sulla distribuzione spaziale delle attività commerciali, di svariate (21) categorie merceologiche, presenti nella città di Torino.

La popolazione totale di Torino, al 2016, ammontava a 900.000 abitanti circa, distribuiti con una densità media di 6.800 abitanti/km². L'area oggetto dell'esame si estende per circa 198.8 km² (14.100 m x 14.100 m), maggiore della reale estensione geografica della città di Torino (130 km²); infatti, l'area in analisi non è stata calcolata considerando i confini amministrativi, ma osservando le posizioni più estreme dei negozi presenti nel database, che ha permesso di definire un rettangolo di 198.8 km², rappresentato nell'immagine sottostante (Figura 1).

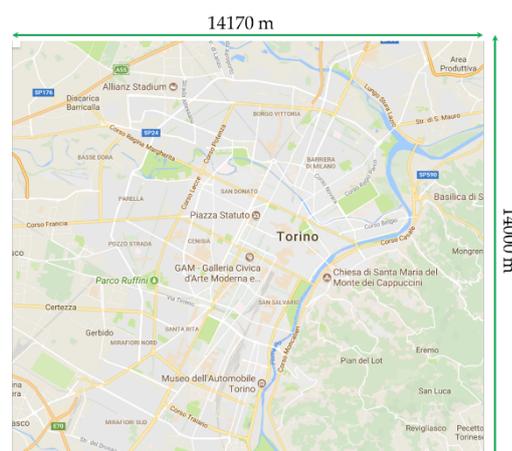


FIGURA 1- ESTENSIONE AREA DI TORINO

I dati su cui sono state effettuate le analisi sono stati ottenuti dal sito http://www.comune.torino.it/geoportale/ser_professionali_2.htm#tabs-%C2%AD%E2%80%90.

Delle categorie merceologiche presenti, è stata applicata una aggregazione di quelle particolarmente simili, come ad esempio Elettronica ed Informatica, e sono state eliminate quelle che presentavano una numerosità troppo bassa, nello specifico minore di 30 negozi, perché avrebbero potuto distorcere le analisi.

Il dataset ottenuto dal Comune di Torino, contiene 6.681 negozi distribuiti su 27 categorie merceologiche. Ciascuna entità commerciale, oltre ad essere georeferenziata⁸, è anche caratterizzata da una dimensione e da un anno apertura e un anno chiusura, ove questa si sia verificata; per ogni negozio si ha, pertanto, la sua estensione in m². Le categorie presenti, loro numerosità e la media dei metri quadri dei negozi sono riportati nella Figura 2.

Un'ulteriore modifica, parecchio rilevante ai fini dell'analisi, è stata quella di effettuare una distinzione tra i "Mom-and-Pops Stores" e i Department Stores. I primi sono stati ottenuti accorpando tutti quelle categorie che vendono generi alimentari, come Macellai, Panettieri, Botteghe e simili, che non appartengono ad uno specifico brand e, pertanto vivono del proprio fatturato e seguono determinate logiche di sopravvivenza. Differentemente da ciò si sono messi in evidenza i Minimercati, Supermercati e Ipermercati, che invece appartengono ad uno specifico network e, pertanto, godono di un'inerzia maggiore, in quanto sorretti dal proprio marchio, che permette di resistere in delle situazioni nelle quali un comune negozio morirebbe.

⁸ Tutti i dati della città di Torino sono stati acquisiti nel sistema di riferimento Gauss Boaga ed il codice internazionale di riferimento è EPSG 3003 – Monte Mario Italy Zone 1.

Categoria Merceologica	Numerosità	Media dei mq
Abbigliamento	1278	101.9
Alimenti biologici	43	81.5
Articoli funerari	75	52.1
Articoli per animali	116	108.4
Articoli sanitari e ortopedici	72	123.0
Autoveicoli e motoveicoli	579	321.8
Bigiotteria	175	59.6
Calzature	141	68.1
Carburanti	91	60.1
Cartolerie	255	53.7
Complementi d'arredo	300	119.5
Elettronica e informatica	550	70.8
Enoteca	51	76.6
Farmacia	152	79.6
Fioraio	113	48.7
Frutta e verdura	156	57.1
Gastronomia	250	52.6
Giornalaio	220	41.0
Intimo	46	75.4
Librerie	112	67.4
Macelleria	271	43.0
Oggetti preziosi	199	44.0
Opere d'arte	34	58.3
Ottica	90	69.6
Panetteria	482	62.5
Profumeria	631	66.9
Tabacchi	199	39.8

FIGURA 2- CATEGORIE MERCEOLOGICHE

Il **dataset complessivo**, utilizzato per l'analisi empirica, si pone l'obiettivo di andare a studiare gli **spazi** adibiti al Retail in Torino negli ultimi 13 anni, **dal 2004 al 2017**. Occorre prestare attenzione al fatto che, ogni riga del database rappresenta l'**ID** di uno spazio univoco in Torino, ottenuto mediante una concatenazione di Indirizzo, categoria merceologica e metratura. Per lo studio degli spazi, si sono tenuti in considerazione differenti eventi che possono esser avvenuti nel tempo:

- *uno spazio commerciale ha cambiato categoria merceologica*: considerato come un nuovo negozio;
- *uno spazio commerciale ha cambiato metratura*: considerato come un nuovo negozio;
- *uno spazio commerciale ha cambiato categoria merceologica e metratura*: considerato come un nuovo negozio;

- *uno spazio è rimasto invariato nel tempo*: se ne studia la sua storia, considerando il minimo Anno Apertura con cui si presenta nel database e il massimo Anno Chiusura.

In definitiva, esso contiene 17.865 istanze, distribuite su 21 categorie merceologiche. La Figura 3 mostra le categorie merceologiche e la loro numerosità, in termini di spazi.

Per studiare il comportamento delle singole categorie merceologiche sono state introdotte nel database delle **variabili dummy**⁹, utilizzate per stratificare il campione.

Categoria Merceologica	Numerosità
Abbigliamento	3598
Articoli Funerari	110
Supermercati	210
Librerie	334
Calzature	318
Elettronica	1194
Alimentari	6093
FarmaciaePara	455
Articoli per la Casa	767
Minimercati	143
Ipermercati	14
FioriePiante	240
Articoli per Animali	349
Cartolerie	562
Telefonia	534
Giocattoli	137
Mobili	573
Ottica	144
Profumeria e Cosmetica	1091
Quotidiani	456
Tabacchi	542

FIGURA 3- CATEGORIE MERCEOLOGICHE SELEZIONATE

⁹ variabile binaria, che può assumere solo due valori: 0 o 1.

4.2 QGIS

QGIS (Sistema Informativo Geografico) è un software opensource multiplatforma che permette di effettuare delle analisi territoriali. I dati, vengono suddivisi in *Layers* ed analizzati graficamente, mediante la creazione di un'immagine mappa. A seconda del tipo di analisi sui dati, sulla mappa possono essere visualizzate le istanze con diverse gradazioni di colore, sfumature o valori unici. Le mappe possono, inoltre, essere arricchite da etichette e icone, che dipendono dagli attributi degli elementi cartografici.

Il software permette l'esecuzione di svariate analisi spaziali, mediante l'utilizzo di appositi strumenti che mette a disposizione, e la gestione di dati geografici in maniera automatica; quest'ultima funzionalità risulta essere di particolare interesse ai fini dello studio in questione, in quanto, ad esempio, possono essere visualizzati facilmente su una mappa tutti i negozi presenti nella Città di Torino, e quindi nel dataset. È facile quindi tradurre i risultati ottenuti da analisi empiriche in modo visivo e facilmente comprensibile e intuitivo.

4.3 SPSS

IBM SPSS STATISTICS è un sistema completo di analisi dei dati e rappresenta uno dei programmi più utilizzati per le analisi statistiche nelle scienze psico-sociali.

SPSS permette di effettuare computazione statistica e di raffigurare graficamente i risultati; fornisce un'infinità di strumenti e tecniche statistiche, come ad esempio Analisi delle Frequenze, creazione di Descrittive dei dati, Test statistici (come il test T), Analisi ANOVA, Analisi delle Correlazioni, Analisi tramite modelli Lineari e Non Lineari, etc.

Nello specifico, per l'analisi effettuata, l'utilizzo del software è risultato fondamentale principalmente per due motivi:

- Studio della Correlazione Bivariata: SPSS fornisce per ogni variabile il numero di casi con valori non mancanti e le statistiche media e deviazione standard e per ogni coppia di variabili consente di calcolare il coefficiente di correlazione di Pearson, rho di Spearman e tau-b di Kendall, con i rispettivi livelli di significatività. Le correlazioni consentono la misurazione della relazione sussistente tra le variabili o i punteggi di rango (comandi utilizzati: Analyze> Correlation> Bivariate);
- Studio dell'Analisi di Sopravvivenza, mediante l'applicazione di una Regressione non lineare, che verrà descritta nel capitolo successivo (comandi utilizzati: Analyze> Survival> Life Table e Analyze>Survival>Cox Regression).

5. METODO UTILIZZATO: LA SURVIVAL ANALYSIS

5.1 COS'È UNA DURATION ANALYSIS?

L'analisi di sopravvivenza riguarda tutti quegli studi in cui si vuole analizzare l'incidenza di un determinato evento in un certo arco temporale, detti anche **studi di coorte**.

Quando si parla di *Duration Analysis* ci si riferisce all'analisi di Sopravvivenza o all'analisi di eventi storici. I dati sulla Duration possono essere generati da ciò che viene definito un "failure time process", ovvero lo studio del tempo di fallimento. Il termine tempo di *fallimento*, o tempo di sopravvivenza, deve essere utilizzato in modo estensivo, poiché si applica anche ad eventi differenti dalla morte ed assume significati diversi a seconda dell'evento a cui il ricercatore è interessato. Un *failure time process* consiste di unità –individuali, governative, nazionali e così via- che possono essere osservate a partire da un certo istante di tempo. Queste unità si presentano in un determinato stato – l'individuo è in salute, il governo è al potere, la nazione è in pace, eccetera- e vengono osservate nel tempo. In un dato istante di tempo, queste unità si presentano "a rischio" di subire un determinato evento, dove con "evento" ci si riferisce essenzialmente ad un cambiamento o transazione ad un altro stato – l'individuo muore, il governo cade, una nazione è in guerra. Dopo che l'evento si verifica, le unità non vengono più osservate o diventano a rischio di sperimentare un altro tipo di evento. In alcune circostanze, nelle unità non viene osservato l'evento, perché nessuna transizione da uno stato all'altro si verifica durante il tempo di osservazione – l'individuo rimane sano, il governo rimane al potere, la nazione rimane in pace. Questi casi vengono definiti "**censurati**", poiché non si può osservare la storia successiva dell'unità dopo l'ultimo istante di osservazione (nel caso dello studio in esame, ad esempio, tutti quegli spazi commerciali che al 2017, ovvero al

momento di conclusione dell'osservazione, erano ancora vivi, sono considerati come unità censurate). Questo processo è definito *failure time process* perché:

- a. Le unità vengono osservate a partire da un istante di tempo iniziale
- b. Le unità sopravvivono fino ad un certo istante di tempo
- c. Ad un certo punto le unità "falliscono" o vengono "censurate".

Esempio: **Studio empirico effettuato**

- Durata di uno Spazio Commerciale in Torino
 - Evento: Morte dello Spazio (l'attività chiude)
 - Timing: Anni di vita
 - Rischio: "Dato che il negozio in quello spazio è sopravvissuto fino all'anno X, qual è il rischio che morirà l'anno successivo?"

Quindi, **il timing implica il rischio**, ovvero ciò che ci si va a chiedere è: **"Dato che un evento non è ancora accaduto, qual è la probabilità che avvenga successivamente?"**

Ma cosa si intende esattamente per rischio e come si misura? Il rischio è un rapporto che indica una relazione tra la probabilità che qualcosa accada relativamente al fatto che non è ancora accaduta. Si tratta quindi di una probabilità condizionata:

$$RISCHIO = \frac{P(FALLIMENTO)}{P(SOPRAVVIVENZA)}$$

Dove con Fallimento si intende la probabilità incondizionata che un evento si verificherà; con Sopravvivenza si intende la probabilità che l'evento non si è verificato "fino ad ora" e con Rischio si intende il tasso di fallimento condizionato, che l'evento si verificherà, considerando il fatto che non si è verificato fino ad ora.

Convenzionalmente, nelle analisi di Duration, il rischio è chiamato **Hazard Ratio** e lo studio di quest'ultimo rappresenta il fulcro dell'intera analisi.

Nel modello di Sopravvivenza, la **variabile dipendente Y** è la durata di tempo speso in un determinato stato prima del verificarsi dell'**evento**; la variabile evento rappresenta proprio la transizione da uno stato all'altro, come ad esempio il verificarsi del fallimento o della morte. Viene modellata la relazione tra la *duration* e le **covariate X_i** , ovvero si studia come un incremento nella covariata X_i impatta la duration Y.

5.1.1 PERCHÉ NON SI UTILIZZA UN MODELLO OLS?

Sarebbe intuitivo pensare di utilizzare un modello OLS - *Ordinary Least Squares* per testare le ipotesi di interesse ed utilizzare la variabile Duration come variabile dipendente; in realtà, però, non è possibile utilizzare un modello di questo tipo con i dati di duration, pertanto si ha la necessità di muoversi verso dei modelli specificatamente disegnati per lo studio di dati di questa tipologia. Il modello OLS, infatti, ha alla base un'**assunzione di normalità**, ovvero assume che i tempi di durata (variabile dipendente Y) siano normalmente distribuiti; questa assunzione in questo contesto è irrealistica, in quanto le variabili dipendenti di duration assumono sempre valori positivi (numero di anni, numero di giorni, eccetera) e, inoltre, i dati possono mostrare spesso delle asimmetrie, specie se alcune osservazioni hanno una durata veramente lunga. Una soluzione che spesso si utilizza per ovviare a questo problema è quella di applicare una trasformazione logaritmica alla variabile dipendente prima di applicare il modello. Questo non è però l'unico problema che ne impedisce l'applicazione, infatti l'OLS Model non riesce facilmente a distinguere tra osservazioni "censurate" ed "non censurate"; eliminare queste osservazioni porterebbe portare ad un problema di selezione del campione. Un ulteriore ostacolo, parecchio rilevante, consiste nel non poter utilizzare delle covariate che cambiano valore nel tempo (TVC- Time Varying Covariates).

5.2 BACKGROUND – LA REGRESSIONE LOGISTICA

A differenza della regressione lineare (vedi OLS), **la regressione non-lineare** (in particolare per il caso in esame vedremo quella logistica) non assume la linearità della relazione tra la variabile dipendente e le variabili indipendenti, non richiede la normalità della distribuzione delle variabili e ha dei vincoli meno stringenti.

I modelli di **regressione logistica** nascono nell'ottica di spiegare la probabilità di possesso di un particolare attributo o di accadimento di uno specifico evento in relazione ad una serie di possibili determinanti, o variabili esplicative. Si tratta di un particolare caso di analisi di regressione¹⁰ in cui **la variabile dipendente è dicotomica**, piuttosto che quantitativa: convenzionalmente assume i valori 0 o 1 e ha una distribuzione, ovviamente, binomiale. La stima di Y, nel continuo, dovrà variare non nei reali, ma all'interno dell'intervallo [0,1] e assume il significato di probabilità, o rischio, che Y valga 1.

5.2.1 ESTENSIONE ALL'EVENT HISTORY ANALYSIS – IL RUOLO CRUCIALE DEL TEMPO

I modelli di **Survival Analysis** nascono negli anni '70 nel contesto della costruzione di modelli stocastici che permettano di esplicitare la dipendenza del processo da una serie di variabili esplicative, secondo la logica della regressione multipla. In tali modelli la variabile dipendente è la *probabilità, o meglio il rischio*, che l'evento di interesse accada nel tempo t , nel caso in cui il tempo venga misurato in unità discrete, o nell'intervallo di tempo che intercorre tra t e $t+\Delta t$, se invece la misurazione del tempo avviene mediante variabile continua. Le variabili indipendenti, o esplicative, invece possono essere rappresentate da tutti

¹⁰ Sono tutti modelli riconducibili alla classe dei modelli lineari generalizzati (McCullagh e Nelder, 1989)

quei fattori che si pensa possano, in qualche modo, condizionare lo studio in esame.

L'analisi di Sopravvivenza, o *event history analysis*, a tempi discreti rappresenta un'estensione del modello di regressione logistica nella quale si considera la probabilità non indipendentemente dal tempo di attesa di accadimento dell'evento; viene utilizzata, quindi, per analizzare i dati in cui è di interesse *il tempo all'evento*. La dimensione temporale viene introdotta tra le variabili esplicative e la variabile dipendente rimane dicotomica. È un modello che imposta il fenomeno studiato come un "*processo*" caratterizzato dal verificarsi di un particolare evento.

5.2.2 LA CURVA DI SOPRAVVIVENZA E LA "LIFE TABLE"

La tipologia di osservazioni, di cui si è parlato al paragrafo precedente, viene rappresentata graficamente mediante le **curve di sopravvivenza**, che consistono nella rappresentazione grafica della probabilità di sopravvivere all'evento in funzione del tempo. La curva di sopravvivenza è un output grafico della **Life Table**, che rappresenta una procedura descrittiva per esaminare la distribuzione della variabile *tempo all'evento*. L'idea alla base della Life Table è quella di suddividere il periodo di osservazione in intervalli di tempo più piccoli ed effettuare poi una stima della probabilità per ogni intervallo.

Nella figura sottostante (Figura 4) viene visualizzato il modo in cui la Life Table si presenta; l'output presentato è stato ottenuto **utilizzando il software SPSS**. Per la creazione della Life Table occorre inserire all'interno del software le seguenti **variabili**:

- **Time Variable**, o duration variable: è la variabile continua che indica la durata;

- **Status Variable:** è la variabile binaria o categorica, che rappresenta l'evento di interesse;
- **Factor Variable:** si può effettuare una comparazione della distribuzione della curva di sopravvivenza per ogni livello di un determinato fattore inserendo una factor variable.

Life Table								
		Interval Start Time	Number Entering Interval	Number Withdrawing during Interval	Number Exposed to Risk	Number of Terminal Events	Proportion Terminating	Proportion Surviving
First-order Controls								
custcat Customer category	1 Basic service	0	266	3	264.500	10	.04	.96
		3	253	10	248.000	17	.07	.93
		6	226	12	220.000	10	.05	.95
		9	204	11	198.500	10	.05	.95
		12	183	13	176.500	6	.03	.97
		15	164	10	159.000	5	.03	.97

Cumulative Proportion Surviving at End of Interval	Std. Error of Cumulative Proportion Surviving at End of Interval	Probability Density	Std. Error of Probability Density	Hazard Rate	Std. Error of Hazard Rate
.96	.01	.013	.004	.01	.00
.90	.02	.022	.005	.02	.01
.86	.02	.014	.004	.02	.00
.81	.02	.014	.004	.02	.01
.78	.03	.009	.004	.01	.00
.76	.03	.008	.004	.01	.00

FIGURA 4- ESEMPIO LIFE TABLE

L'output SPSS mostrato sopra, fornisce i seguenti risultati:

- *Interval Start Time:* il valore iniziale di ogni intervallo. Ogni intervallo si estende dal suo tempo di inizio al tempo di inizio dell'intervallo successivo;
- *Number Withdrawing during Interval:* il numero di casi censurati nell'intervallo;
- *Number Exposed to Risk:* il numero dei casi sopravvissuti meno metà dei casi cesurati. Si applica questa modifica per tenere in considerazione i possibili effetti dei casi censurati;
- *Number of Terminal Events:* il numero di casi nei quali si verifica l'evento nell'intervallo di tempo in considerazione. Questi saranno i casi nei quali il valore dell'evento è uguale ad 1;
- *Proportion Terminating:* il rapporto tra i *Terminal Events* e i *Exposed to Risk*;
- *Proportion Surviving:* complementare del *Proportion Terminating*;

- *Cumulative Proportion Surviving at the end of interval*: la proporzione di casi che sopravvivono dall’inizio della tabella alla fine dell’intervallo;
- *Probability Density*: una stima della probabilità di sopravvivenza all’evento terminale durante l’intervallo di tempo;
- *Hazard Rate*: una stima della probabilità del verificarsi dell’evento terminale durante l’intervallo di tempo, condizionatamente al fatto che si è sopravvissuti fino all’inizio dell’intervallo.

5.3 TERMINOLOGIA E NOTAZIONI

Sia T una variabile casuale continua positiva che denota il tempo di sopravvivenza, o la durata. Vediamo adesso alcune caratteristiche della variabile casuale T :

- **funzione di densità di probabilità (PDF) $f(t)$** , che rappresenta la *probabilità che l’evento avvenga al tempo t* . Dato che, $f(t)$ è una funzione continua (vedi esempio in Figura 5), per calcolare la probabilità bisogna guardare all’area sottesa dalla curva di densità per un infinitesimamente piccolo intervallo attorno a t .

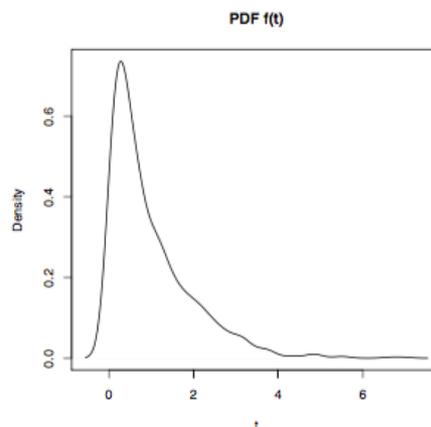


FIGURA 5 - ESEMPIO FUNZIONE DI DENSITÀ DI PROBABILITÀ

- $f(t)$ ha la sua corrispondente **funzione cumulata di probabilità (CDF) $F(t)$** , che rappresenta la probabilità che un evento si verifichi prima del tempo t (vedi esempio in Figura 6): $F(t) = \int_0^t f(u)du = P(T \leq t)$

Quindi, la **funzione di sopravvivenza** o **Survivor Function** (vedi esempio in Figura 7), che rappresenta la *probabilità che nessun evento si verifichi fino, almeno, al tempo t* , si calcola come:

$$S(t) = 1 - F(t)$$

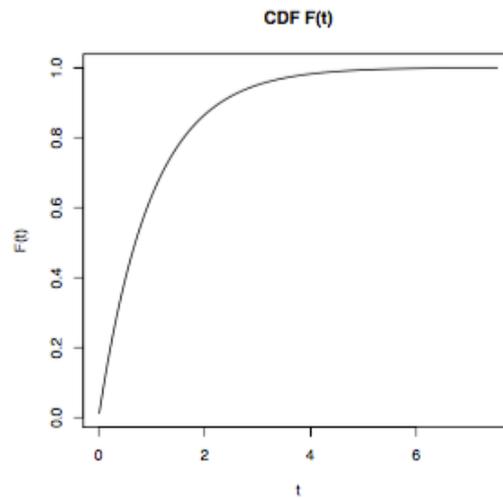


FIGURA 6 – ESEMPIO FUNZIONE CUMULATA DI PROBABILITÀ

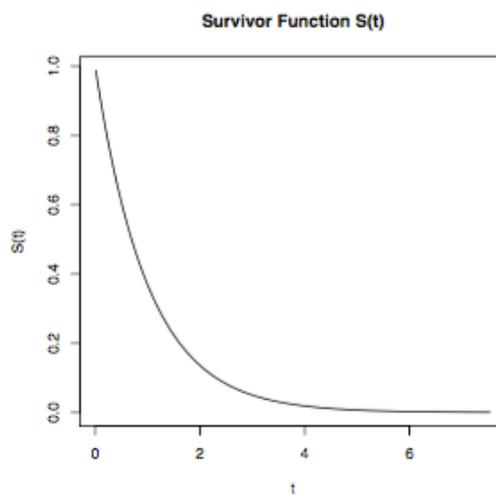


FIGURA 7 – ESEMPIO FUNZIONE DI SOPRAVVIVENZA

L'**Hazard Rate**, o **Hazard Function**, $h(t)$ è la *probabilità che l'evento avvenga al tempo t dato che si è sopravvissuti fino al tempo t* .

Se chiamiamo X l'evento – si ha sopravvivenza fino a t - e Y l'evento – l'evento si verifica in t - si ha che:

$$h(t) = P(Y|X) = \frac{P(X|Y)*P(Y)}{P(X)} = \frac{P(Y)}{P(X)} = \frac{f(t)}{S(t)}$$

$$f(t) = h(t)*S(t)$$

Solitamente, l'hazard rate viene modellato in funzione delle covariate inserite nel modello: $h_i(t)=g(X_i\beta)$; si possono poi interpretare sia i cambiamenti nell'hazard rate, che i cambiamenti nella duration media, dato che si sono verificati dei cambiamenti nelle covariate X . Quando tutte le covariate sono pari a zero, si parla di **baseline hazard**: $h_i(t)=g(\beta_0)$.

5.3.1 L'IPOTESI DI RISCHI PROPORZIONALI

Consideriamo un modello di Survival Analysis per dati individuali, ovvero un modello in cui i dati vengono strutturati in modo che ogni individuo sia distinto da tutti gli altri; ogni record rappresenterà, quindi, un'osservazione di una particolare unità temporale di uno specifico soggetto, il quale potrà, o meno, aver sperimentato l'evento. Ciò equivale alla realizzazione di un **processo Bernoulliano**, in cui ogni record è composto dalla variabile t , da tutte le variabili esplicative e dalla variabile che definisce il verificarsi o meno dell'evento di interesse $[Y_{t,i}]$. Il modello è, quindi, il seguente:

$$(Y_{t,i} | X = x_i) \sim \text{Bernoulli}(\lambda_{t,i})$$

$$\text{logit}(\lambda_{t,i}) = \alpha_t + x_t\beta$$

Il generico record i contribuirà all'analisi con t_i osservazioni:

- $t_i - 1$ insuccessi ed 1 successo, nel caso in cui in t_i sperimenti l'evento $[Y_{t,i} = 1]$;
- t_i insuccessi, nel caso in cui risulti un record censurato.

Nel caso in cui le variabili esplicative siano costanti nel tempo, riferendosi al rischio in termini di **odds**, si avrà che:

$$\text{logit}(\lambda_{t,i}) = \ln\left(\frac{\lambda_{t,i}}{1 - \lambda_{t,i}}\right) = \ln[\text{odds}(\lambda_{t,i})] = \ln\left[\frac{P(Y_{t,i} = 1|X = x_i)}{P(Y_{t,i} = 0|X = x_i)}\right] = \alpha_t + x_i\beta$$

Il parametro α_t non dipende dalle caratteristiche introdotte nel modello ed è pertanto considerato una baseline del rischio a cui è soggetta ogni osservazione al tempo t : $\alpha_t = \text{logit}(\lambda_{t,0})$, pertanto si ottiene che:

$$\text{logit}(\lambda_{t,i}) = \text{logit}(\lambda_{t,0}) + x_i\beta$$

$$\ln[\text{odds}(\lambda_{t,i})] = \ln[\text{odds}(\lambda_{t,0})] + x_i\beta$$

$$\text{odds}(\lambda_{t,i}) = \text{odds}(\lambda_{t,0}) * \exp(x_i\beta)$$

Il rischio del record i , al tempo t , è composto, come indica la formula sovrastante, da due fattori: uno relativo alla dimensione temporale, ovvero all'unità di tempo t , che rappresenta il **rischio base** [$\text{odds}(\lambda_{t,0})$]; il secondo si riferisce invece esclusivamente alle caratteristiche individuali [$\exp(x_i\beta)$].

Il modello ha alla base un'ipotesi fondamentale sottostante, che si evince facilmente dalla rappresentazione sopra enunciata, ovvero il fatto che le osservazioni sono soggette allo stesso rischio base, a prescindere dall'unità di tempo considerata; in altre parole, i rischi a cui le diverse unità sono soggette, risultano essere **proporzionali ad un rischio base, e pertanto proporzionali tra loro, secondo un fattore indipendente dal tempo.**

5.4 LA STIMA DEI PARAMETRI

Per la stima dei parametri non noti del modello si utilizza il metodo della Massima-Verosimiglianza – **Maximum Likelihood Estimation**.

L'obiettivo è quello di trovare un set di parametri θ , nel quale è incluso il parametro β e possibili altri, dato che si hanno a disposizione i dati y_i .

$$P(\theta|y) = \frac{P(y|\theta)P(\theta)}{P(y)} = k(y)P(y|\theta) \propto P(y|\theta)$$

$$\mathcal{L}(\theta|y) = \prod_{i=1}^n P(y_i|\theta)$$

Il massimo valore di $\mathcal{L}(\theta|y)$ in uno spazio multi-dimensionale è lo stimatore MLE del parametro in esame. Nel caso della Survival Analysis, si ha che $P(y|\theta) = f(t)$.

Le osservazioni "censurate" non forniscono alcuna informazione riguardo l'istante in cui l'evento si verifica, ma forniscono informazioni riguardo quanto a lungo essi sopravvivono. Per questo tipo di osservazioni si sa che sopravvivono fino ad un determinato istante t^c e che la loro effettiva duration (t^*) è superiore a t^c : $t^* \geq t^c$.

Per ogni osservazione viene quindi creato un "indicatore di censura" c_i che assume i valori $[0,1]$ (0 se l'osservazione è censurata, 1 nel caso opposto), e viene introdotta questa informazione all'interno della funzione di massima verosimiglianza: le osservazioni incensurate contribuiscono alla verosimiglianza con la loro funzione di densità, mentre quelle censurate con la loro funzione di sopravvivenza.

$$\begin{aligned} \mathcal{L} &= \prod_{i=1}^n [f(t_i)]^{c_i} [P(T_i^* \geq t_i^c)]^{1-c_i} = \prod_{i=1}^n [f(t_i)]^{c_i} [1 - F(t_i)]^{1-c_i} \\ &= \prod_{i=1}^n [f(t_i)]^{c_i} [S(t_i)]^{1-c_i} \end{aligned}$$

5.4.1 VERIFICA DELLA BONTÀ DEL MODELLO IN TERMINI DI ADATTAMENTO

Per verificare la bontà del modello si parte descrivendo il caso ottimale, ovvero quello in cui si riescono ad ottenere delle stime perfettamente coincidenti con i dati osservati. Se ad esempio, associassimo un singolo parametro ad una osservazione e annullassimo di volta in volta tutti gli altri avremmo che la stima

di questa, vincolata ad un solo dato, coinciderebbe esattamente con il valore del dato osservato. Si avrebbe quindi il **perfetto adattamento del modello ai dati osservati**.

Il modello nel quale il numero di parametri coincide con il numero n di osservazioni a disposizione è quello a cui corrisponde il massimo valore di verosimiglianza ottenibile (*modello saturato*). Si tratta di una soluzione, però, eccessivamente dispendiosa e per nulla flessibile ad un'estensione ad un insieme di altri dati.

Per comprendere se il modello in considerazione (p parametri esclusa l'intercetta) possa essere ritenuto o meno una buona approssimazione, il valore della massima verosimiglianza ad esso associato non dovrebbe subire eccessive diminuzioni rispetto a quello del modello saturato. Effettuando un test sul rapporto di verosimiglianza può essere impostato un confronto tra la situazione reale e quella ottimale:

$$D = -2 \ln \left[\frac{\max(\mathcal{L}_{\text{modello considerato}})}{\max(\mathcal{L}_{\text{modello saturato}})} \right] \sim \chi_{n-(p+1)}^2 = \chi_{n-p-1}^2$$

Il suddetto rapporto prende il nome di **devianza**. Per comprendere se il modello considerato sia una buona approssimazione di quello saturato, la devianza deve assumere valori piccoli; questo indica che esiste una pressoché equivalenza, in termini di verosimiglianza, tra i due modelli. Sotto questa condizione si può accettare l'*ipotesi nulla* di non significatività delle variabili non inserite nel modello.

5.4.2 ANALISI DEI RESIDUI

Per concludere il processo di analisi del modello, dopo aver effettuato il processo di stima, occorre verificare l'accuratezza di quest'ultime e delle ipotesi effettuate. Il procedimento più frequentemente utilizzato, per comprendere l'accuratezza,

consiste nel fornire una valutazione qualitativa, derivante dall'analisi grafica dei *residui* restituiti dal modello, ovvero di misure della distanza esistente tra i valori stimati e quelli osservati:

$$r_{ti} = y_{ti} - \widehat{y}_{ti}$$

La prima analisi da effettuare, per la verifica della bontà del modello, consiste nel verificare la distribuzione dei residui: essi dovrebbero presentare una distribuzione normale o quanto più possibile simmetrica. Dato che si sta effettuando un'analisi di sopravvivenza, però, la variabile dipendente ha natura bernoulliana e, pertanto, le stime di quest'ultima risultano aver valori compresi tra 0 e 1; la normalità, o simmetria, della funzione, in queste situazioni risulta, spesso, strutturalmente impossibile da ottenere.

L'alternativa, quindi, è quella di procedere con ulteriori analisi; quelle di maggiore importanza vengono riportate di seguito (nella descrizione viene indicato indistintamente con \widehat{y}_{ti} sia il valore stimato dal modello per la singola osservazione che quello della media riferita ad un gruppo di queste).

- Si può considerare la *k-esima* variabile esplicativa ed effettuare una rappresentazione grafica dei punti (x_{ki}, \widehat{y}_i) ; si utilizza questa rappresentazione per comprendere se risulta verificata o meno l'ipotesi di dipendenza lineare tra $\text{logit}[P(Y = 1|X)] = \text{logit}(\lambda)$ e X_k . Se quest'ultima dovesse venir rigettata, l'andamento grafico potrebbe comunque quale relazione sia in grado di esprimere la dipendenza in modo migliore;
- Si può creare un grafico dei residui relativamente a "tutte le unità statistiche elencate in ascissa", nel caso in cui la numerosità campionaria non dovesse risultare eccessivamente elevata. In un modello che si possa definire buono, il valore dei residui non dovrebbe discostarsi troppo dallo 0; pertanto, effettuare un'analisi di questo tipo permetterebbe di verificare

l'esistenza di alcuni valori anomali, detti *outliers*, considerando che quest'ultimi potrebbero essere dei valori che il modello preso in considerazione non è in grado di spiegare;

- Si può procedere, in alternativa, con la creazione di una rappresentazione grafica dei punti (r_i, \hat{y}_i) ; questi punti, in un modello accettabile dovrebbero essere distribuiti in modo quanto più possibile casuale attorno all'asse delle ascisse. Se questi dovessero presentare un andamento specifico (ad esempio un andamento sistematicamente crescente al crescere dei valori stimati), la scelta di un modello logistico potrebbe risultare inappropriata;
- Infine, si potrebbe effettuare una rappresentazione normale, definita anche "**normal plot**"; in un grafico di questo tipo vengono inseriti in ascissa i valori dei residui del modello e in ordinata un ugual numero di valori casuali di una normale standardizzata. I valori corrispondenti ai punti localizzati non eccessivamente distanti dalla diagonale principale sono quelli dei residui che possono essere considerati normalmente distribuiti; i restanti, invece, possono essere considerati come corrispondenti a valori anomali.

5.5 TECNICA DI REGRESSIONE SEMI-PARAMETRICA: COX REGRESSION

La procedura di **Regressione Cox** è utile per modellare *il tempo ad uno specifico evento*, basandosi sui valori di determinate covariate. L'output statistico al centro dell'analisi è l'hazard ratio; ciò che si stima, però, con questo tipo di analisi è il *rischio relativo* e non quello assoluto.

Si tratta di un modello **semi-parametrico**¹¹, che non fa alcuna assunzione riguardo la forma della funzione di hazard o sulla distribuzione della variabile duration (T_i).

La **variabile dipendente** del modello è la probabilità di sopravvivenza, che viene misurata mediante due variabili: la prima *numerica* – il tempo all'evento - e la seconda *dicotomica* – presenza o assenza dell'evento. La **variabile indipendente**, invece, può essere sia di tipo numerico, che qualitativo (binaria o categorica).

Il modello di *Cox Regression* ha delle **assunzioni di base**, che vengono di seguito elencate:

- I t_i vengono concettualizzati come eventi di tempo **discreti** piuttosto che come durations o tempi di sopravvivenza (ad esempio se $t=5$: un evento accade al tempo 5, piuttosto che l'osservazione i sopravvive fino al tempo 5);

¹¹ Esistono essenzialmente tre principali approcci: modello parametrico, semi-parametrico e non-parametrico.

- Modello parametrico: il superamento dell'assunzione stringente del modello OLS sulla distribuzione normale dei disturbi, porta all'adozione di modelli parametrici, come esponenziale, Weibull, log-logistico, e così via. Il problema di questi modelli è che fanno, ancora, delle assunzioni sulla distribuzione dei tempi di sopravvivenza.
- Modello semi-parametrico: la chiave per rimuovere l'assunzione distributiva sui tempi di sopravvivenza, è quella di ordinarli, dato che essi si verificano in determinati istanti di tempo. Si può così condurre l'analisi utilizzando semplicemente l'ordinamento dei tempi. Il tempo in quest'analisi non ha alcun ruolo, se non quello di consentire l'ordinamento delle osservazioni. Viene definito semi-parametrico in quanto, ha un approccio non-parametrico riguardo il tempo, ma è ancora parametrico perché parametrizza l'effetto di x (c'è ancora una componente parametrica – logit, probit– nell'analisi).
- Modello non-parametrico: questo approccio elimina qualsiasi assunzione riguardo come ogni valore osservato di x determini la probabilità di fallire. Sono metodi utilizzati nelle scienze sociali ed hanno spesso problemi con i dati di durata, perché non riescono a gestire i censurati. Quando non ci sono variabili indipendenti, o se quest'ultime sono hanno natura qualitativa, si possono utilizzare determinati metodi non-parametrici per stimare cose come la probabilità di sopravvivere nel passato ad un certo punto di tempo. Questi metodi non fanno assunzioni sulla distribuzione dei tempi di fallimento o su come le variabili indipendenti cambino la sopravvivenza.

- Gli hazard, o rischi, devono essere **proporzionali** tra loro, ovvero il rischio di ogni individuo è una proporzione fissa del rischio di ogni altro individuo ed è costante nel tempo;
- I rischi sono *moltiplicativi*;
- Gli eventi non possono verificarsi nello stesso istante: quindi esattamente un evento avviene in ogni istante di tempo t_i , per tutti i dati non-censurati.

Il modello è composto da due costitutive:

- Una **funzione hazard di base** $h_0(t)$, che viene lasciata non specificata, ma che deve essere positiva: rappresenta il valore dell'hazard quando tutte le covariate hanno valore nullo;
- Una funzione lineare di un insieme k di covariate, che rappresenta il rischio relativo $h_i(t)$.

La *Cox Regression*, esattamente come per i modelli parametrici, analizza il rischio, $h(t)$, mediante le covariate X_i ; viene presentata di seguito l'equazione di quest'ultimo:

$$\ln[h_i(t)] = \ln[h_0(t)] + \beta_1 X_{i,1} + \dots + \beta_k X_{i,k}$$

Applicando una trasformazione esponenziale essa diventa:

$$h_i(t) = h_0(t)e^{\beta_1 X_{i,1} + \dots + \beta_k X_{i,k}}$$

Per misurare il rischio relativo non è necessario conoscere l'hazard di base, infatti considerando, ad esempio, un modello semplice si ha che:

$$h(t, x) = h_0 e^{(\alpha + \beta x)}$$

per $x=0$ si ha la **baseline**: $h(t, 0) = h_0 e^{(\alpha)} = Ah_0$

per $x=k$ si ha: $h(t, k) = h_0 e^{(\alpha + \beta k)} = Ah_0 e^{(\beta k)}$

Se si vuole calcolare l'effetto della variabile esplicativa X , occorre misurare il rapporto tra gli hazard:

$$\frac{Ah_0e^{(\beta k)}}{Ah_0} = e^{(\beta k)}$$

Si ottiene quindi che la misura dell'effetto della X sulla baseline non contiene la baseline (A). Il modello viene stimato senza intercetta, rinunciando quindi alla *forma base del rischio*.

Per calcolare l'effetto dell'aumento di un'unità di una covariata rispetto al valore precedente, ad esempio, si calcola il rapporto tra gli hazard:

$$\frac{h_0(t)e^{\beta_1 X_1 + \dots + \beta_k X_k}}{h_0(t)e^{\beta_1 X_1 + \dots + \beta_k (X_k + 1)}} = e^{\beta_k (k - k + 1)} = e^{\beta_k}$$

Andando ad analizzare il significato del coefficiente β , si vede che questo rappresenta il logaritmo dell'incremento del rischio che si verifica in corrispondenza di un incremento unitario della rispettiva covariata.

In particolare, dando uno sguardo alla funzione di rischio, si nota subito che un **β positivo** suggerisce che ad un incremento della covariata x corrisponde un *incremento* del rischio di morte e, quindi, una diminuzione del tempo di sopravvivenza; viceversa un **β negativo** indica che un incremento della covariata x conduce da una *diminuzione* del rischio di morte; nel caso di **β uguale a uno**, ci si trova dinnanzi al valore limite, che mostra un *indipendenza* della covariata x dal rischio di morte.

Utilizzare un modello di questo tipo conduce essenzialmente a due vantaggi: il primo è che non occorre effettuare nessuna assunzione restrittiva riguardo la forma della funzione di hazard, mentre il secondo è che la Cox Regression risulta essere la scelta migliore nel caso in cui si voglia stimare l'effetto delle covariate e la natura della dipendenza dal tempo non è rilevante.

Alla stessa maniera, però, esistono dei contro a questa scelta, come ad esempio il fatto che l'unica quantità di interesse può essere l'hazard, che il modello può essere soggetto a sovra-adattamento o che la forma della funzione di rischio è sconosciuta, nonostante ci siano dei metodi semi-parametrici che consentano di derivare l'hazard e la funzione di sopravvivenza.

6. ANALISI EMPIRICA: IL CASO DELLA CITTÀ DI TORINO

Si procede, adesso, con una descrizione dell'analisi empirica effettuata; l'obiettivo di quest'ultima consiste nella comprensione di quelli che sono i fattori che impattano sulla sopravvivenza dei retailers nella città di Torino.

6.1 I RETAILERS A TORINO, STUDIO DEI FATTORI IMPATTANTI

Per lo studio sopraenunciato è stato utilizzato il dataset fornito dal Comune di Torino, descritto al punto 5.1. Come annunciato precedentemente, vengono analizzati 17.865 spazi commerciali geolocalizzati.

Grazie alla localizzazione geografica degli spazi commerciali, è stato possibile associare ad ogni spazio la rispettiva circoscrizione di appartenenza, la zona statistica ed il proprio indirizzo. La città di Torino, infatti, come previsto dal (Città di Torino, 2016) è suddivisa in 8 circoscrizioni¹² (mostrate in Figura 8) e in 92 zone statistiche (mostrate in Figura 9). Questa associazione, ha reso possibile uno studio sui quartieri di Torino e sugli effetti della distanza dal centro, come verrà mostrato in seguito.



FIGURA 8 - CIRCOSCRIZIONI TORINO

¹² Il territorio cittadino dal 01/01/2016 è, infatti, suddiviso in 8 circoscrizioni. L'ex circoscrizione 10 rientra adesso nella circoscrizione 2 e l'ex circoscrizione 9 rientra nella 8.

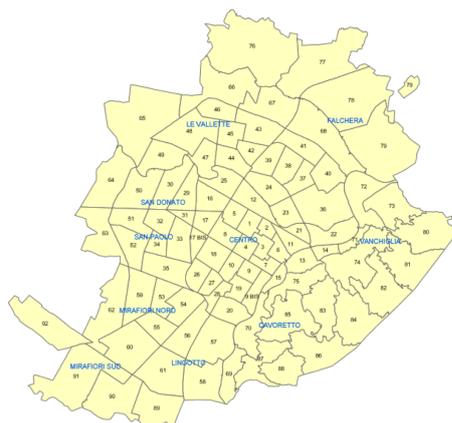


FIGURA 9- ZONE STATISTICHE TORINO

Differenti possono essere i fattori che impattano sulle scelte di comportamento di un retailer; ad esempio, la densità degli esercizi commerciali, intesa come il rapporto tra gli abitanti e il numero di punti vendita, oppure la tipologia di esercizio, distinto in base alla sua dimensione, grande o piccolo; la localizzazione del negozio o la sua categoria merceologica, che permette di comprendere la capacità e il potenziale di un luogo o di una categoria di attrarre un maggior numero di consumatori rispetto ad un altro luogo o categoria; infine, una maggiore concentrazione di offerta commerciale, che risulta maggiormente attrattiva per il consumatore, perché gli consente, in uno spazio fisico più contenuto e con un impiego di tempo inferiore, di valutare una maggiore quantità di articoli ed usufruire di una maggiore varietà di scelta.

Di seguito verranno analizzati i singoli fattori discriminanti e le procedure utilizzate per la loro analisi.

6.1.1 FATTORI IMPATTANTI - LA POPOLAZIONE

In prima analisi è possibile verificare, sempre grazie alla fonte (Città di Torino, 2016), che le circoscrizioni sono differenti tra loro, in termini di espansione geografica e quindi di popolazione presente in esse. Infatti, vi sono dei quartieri più popolosi ed altri che lo sono meno; questo potrebbe essere un primo fattore impattante nel mondo dei retailers. In prima istanza si potrebbe pensare, infatti,

che un retailer abbia una probabilità di sopravvivere maggiore qualora si trovi in una zona abitata da un maggior numero di persone, che possono usufruire dei servizi offerti, o dei prodotti venduti.

La popolazione di Torino al 31 dicembre 2017 risultava essere di 884.733 abitanti; effettuando un'analisi più dettagliata delle circoscrizioni, e della popolazione in esse presente, si vede che:

1. La *circoscrizione 1*, che comprende il **Centro** e il quartiere della **Crocetta**, è quella che presenta il minor numero di abitanti: 79.314. L'affermazione precedente, pertanto, non risulta essere veritiera in tutti i casi; infatti il Centro di Torino gode di un'alta densità di esercizi commerciali presenti, a fronte di una bassa densità di popolazione residente nella zona. Il 23% degli spazi commerciali presenti nel dataset analizzato si trova proprio in questa circoscrizione.
2. La *circoscrizione 2*, viceversa, è la più densamente popolata: vi si trovano infatti 136.842 abitanti. Essa comprende i quartieri **Santa Rita**, **Mirafiori Nord** e **Mirafiori Sud**. Ulteriore controprova del fatto che non sempre i retailers decidono di localizzarsi ove vi è maggiore presenza di cittadini è fornita dal fatto che in questa circoscrizione sono presenti solo il 10,4% degli spazi commerciali forniti dal dataset oggetto dello studio.
3. La *circoscrizione 3*, anch'essa parecchio popolata, ospita 125.443 abitanti. In essa vengono raggruppati i quartieri **San Paolo**, **Cenisia**, **Pozzo Strada**, **Cit Turin** e **Borgata Lesna**. Il 13,3% degli spazi commerciali si trovano in questa circoscrizione.
4. La *circoscrizione 4*, comprendente **San Donato**, **Campidoglio** e **Perella**, ha 97.339 abitanti. In essa si trovano solamente il 9% dei retailers analizzati.

5. Nella *circostrizione 5*, **Borgo Vittoria, Madonna di Campagna, Lucento e Vallette**, vivono 123.995 cittadini. Anch'essa, come la precedente, presenta un piccolo numero di spazi commerciali, il 9,5%.
6. La *circostrizione 6* presenta una popolazione di 106.265 abitanti e comprende le zone di **Barriera di Milano, Regio Parco, Bertolla, Falchera, Rebaudengo e Villaretto**. Essa esibisce una presenza di spazi commerciali, secondo il dataset, del 9,4%.
7. La *circostrizione 7*, meno popolata delle ultime presentate, ospita 86.500 cittadini. I quartieri **Aurora, Vanchilia, Sassi e Madonna del Pilone** appartengono ad essa; il 10% degli esercizi commerciali sono ubicati qui.
8. Infine, la *circostrizione 8*, è la seconda per popolazione, infatti risiedono in essa 120.035 abitanti. Il 15,7 % degli spazi commerciali si trovano nei quartieri di **San Salvario, Cavoretto, Borgo Po, Nizza Millefonti, Lingotto, Filadelfia**.

Per confutare o meno l'assunzione di maggiore sopravvivenza del retailer in presenza di una maggiore residenza di popolazione nella zona in cui esso è ubicato, si può inserire la **densità di popolazione** come variabile indipendente di un'analisi statistica, che verrà mostrata in seguito. Si tratta di una variabile che può assumere i valori *discreti* associati alle otto circoscrizioni.

6.1.2 FATTORI IMPATTANTI – IL FORMATO DEL RETAILER

In seconda analisi, è sembrato opportuno verificare l'esistenza o meno di un impatto sul comportamento dei retailers della loro **dimensione**. Il dataset a disposizione fornisce, infatti, per ciascun negozio la corrispondente superficie, espressa in metri quadri.

Sulla base della superficie, si è effettuata una distinzione tra i *formati* di appartenenza del negozio. È stata utilizzata la classificazione proposta da

(Osservatorio Regione Piemonte del Commercio, 2015), secondo la quale esistono:

- **Esercizi di Vicinato**, ovvero quegli esercizi commerciali la cui superficie di vendita non supera i 150 metri quadri;
- **Medie Strutture di Vendita**, ovvero quegli esercizi commerciali la cui superficie di vendita è compresa tra i 151 e i 1500 metri quadri;
- **Grandi Strutture di Vendita**, ovvero quegli esercizi commerciali la cui superficie di vendita è superiore ai 1500 metri quadri.

Ad ogni istanza del dataset è stato, pertanto, assegnato un formato e ci si è chiesto poi se i tre differenti formati portino o meno ad un comportamento differente. Lo studio è stato condotto analizzando i **flussi di entrata e uscita dal mercato** da parte dei retailers appartenenti al singolo formato; ovvero, ci si è chiesto se i negozi presentano delle dinamiche di nascita e morte differente a seconda della loro dimensione.

Prima di presentare i risultati a cui ha condotto questo studio, risulta indispensabile descrivere cosa si intende per tasso di entrata e di uscita in un determinato periodo di tempo¹³ e a cosa può condurre un'analisi di questo tipo:

- Il **tasso di entrata** (o tasso di nascita) è stato calcolato effettuando una media del rapporto tra il numero di retailers nati nel periodo e il numero di retailers attivi nel periodo;
- Il **tasso di uscita** (o tasso di morte) è stato calcolato effettuando una media del rapporto tra il numero di retailers morti nel periodo e il numero di retailers attivi nel periodo.

¹³ Il periodo di tempo a cui ci si riferisce da ora in avanti è di 13 anni, ovvero gli anni intercorsi tra il 2004 e il 2017, dei quali si sta effettuando l'analisi.

È possibile confrontare questi valori con quelli medi generici, ovvero il tasso di entrata medio e il tasso di uscita medio, ottenuti dalla media dei valori delle singole osservazioni; quello che si ottiene sono, quindi, gli **scostamenti rispetto alla media**. Dall'analisi degli scostamenti si possono ottenere essenzialmente quattro situazioni distinte:

1. Retailers che *nascono e non muoiono*, i quali hanno un tasso di nascita superiore alla media e un tasso di morte inferiore;
2. Retailers che *nascono e muoiono*, i quali presentano i tassi di nascita e di morte entrambi superiori alla media;
3. Retailers che *non nascono e non muoiono*, i quali presentano i tassi di nascita e di morte entrambi inferiori alla media;
4. Retailers che *non nascono e muoiono*; questo rappresenta il caso peggiore in quanto il tasso di nascita è inferiore alla media, mentre quello di morte superiore.

Volendo rappresentare graficamente queste quattro situazioni è possibile creare una **mappa percettiva**¹⁴ (o **mappa di posizionamento**), inserendo nell'asse delle ascisse il tasso di uscita e nell'asse delle ordinate il tasso di entrata; si ottiene il grafico mostrato in Figura 10, dove in verde viene evidenziata la situazione ideale, mentre in rosso la meno desiderabile.

¹⁴ La mappa di posizionamento è uno strumento solitamente utilizzato nel Marketing per riassumere le posizioni di determinati prodotti o di determinati brand sul mercato. E' una rappresentazione grafica su due dimensioni, che permette in maniera piuttosto semplice di analizzare le distanze e le differenze sussistenti tra gli oggetti che si stanno osservando, in funzione delle variabili di interesse



FIGURA 10 - MAPPA DI POSIZIONAMENTO

In riferimento all'analisi dei formati, sono stati calcolati, per ogni categoria (il formato), i tassi di entrata e di uscita dei retailers ad essa appartenenti, i rispettivi tassi medi e gli scostamenti rispetto alla media. Si ottengono tre differenti situazioni, mostrate in Figura 11.

I tassi di entrata ed uscita medi, rispetto ai quali si calcolano gli scostamenti hanno valori pari a **13,50%** e **3,78%** rispettivamente.

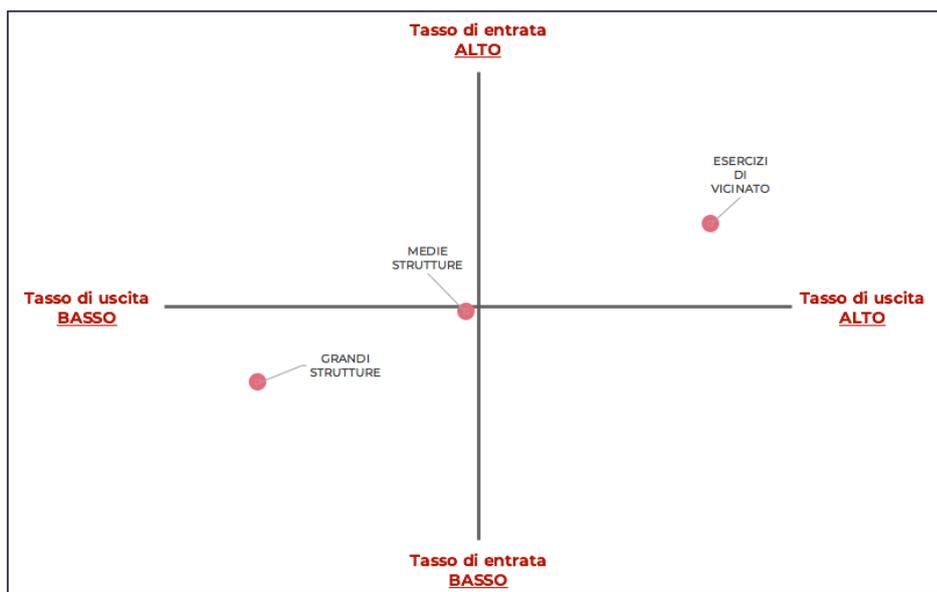


FIGURA 11- MAPPA DI POSIZIONAMENTO FORMATI

Dalla figura si vede come gli esercizi di vicinato si trovino in una situazione diametralmente opposta rispetto a quella delle grandi strutture: infatti i primi hanno entrambi i tassi, di entrata e di uscita, superiori alla media (con valori pari a 14,63% e 6,13% rispettivamente) e, pertanto, sono soggetti a fenomeni di nascita e morte dinamici e in rapida evoluzione; i secondi, invece, hanno entrambi i tassi inferiori alla media (12,59% e 1,80% rispettivamente), ovvero si trovano in una situazione stazionaria di non nascita e non morte. Diversa ancora è la situazione delle medie strutture, che rappresentano la realtà media rispetto alle due situazioni opposte presentate.

È evidente, quindi, come le dinamiche dei negozi appartenenti ai tre formati siano differenti e, pertanto, come la dimensione sia un fattore impattante. Per questo motivo, i metri quadri sono stati inseriti nel modello statistico, che verrà presentato successivamente, come variabile indipendente, in modo da poterne valutare l'impatto sulla sopravvivenza. Si tratta di una variabile *continua* che assume valori compresi tra 1 e 9995 metri quadri, la cui distribuzione è mostrata in Figura 12.

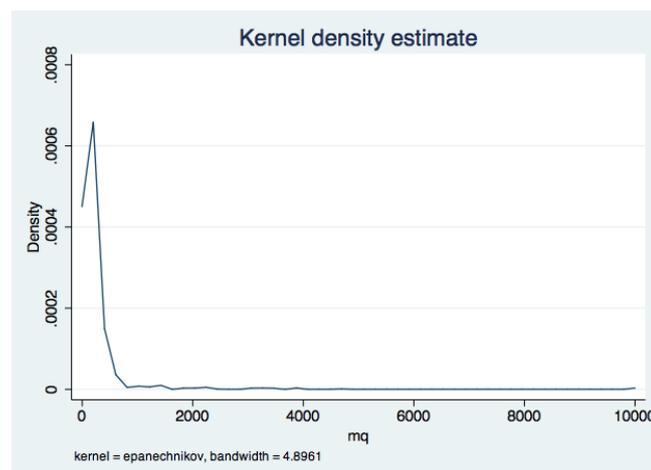


FIGURA 12 - DISTRIBUZIONE METRI QUADRI

6.1.3 FATTORI IMPATTANTI – IL POSIZIONAMENTO GEOGRAFICO DEL RETAILER

La differenza esistente tra le circoscrizioni (presentata al punto 7.1.1), può condurre ad un'ulteriore tipo di analisi: può essere, infatti, particolarmente interessante studiare i flussi degli spazi commerciali presenti in esse. Si tratta di verificare con quali tassi i retailers entrano ed escono dalle circoscrizioni e se questi presentano delle differenze sostanziali a seconda dei casi analizzati.

Per un'analisi di questo tipo, si è utilizzato lo stesso procedimento descritto al punto precedente (flussi di entrata e uscita dal mercato): si sono calcolati, per ogni circoscrizione, i tassi di entrata e di uscita dei retailers che in esse sono ubicati, i rispettivi tassi medi (che hanno dei valori pari a **14,56%** nel caso del tasso di entrata medio e **6,21%** nel caso di quello di uscita) e gli scostamenti rispetto alla media.

Quello che si ottiene, come si può vedere in Figura 12, è che i retailers presenti nella circoscrizione 1 ed 8 hanno entrambi i tassi, di entrata e di uscita, inferiori alla media (con valori pari a 14,40% e 5,40% nel primo caso e 14,10% e 5,30% nel secondo): essi quindi non nascono e non muoiono. Situazione differente e peggiore è, invece, quella che si presenta per i negozi che si trovano nella circoscrizione 2, i quali nascono e muoiono, con un tasso di entrata inferiore alla media (14,20%) e uno di uscita superiore (7,41%). La circoscrizione 3 si comporta esattamente nella media; mentre emerge un agglomerato, composto dalle circoscrizioni 4, 5, 6 e 7, che presentano un comportamento simile: tasso di uscita nella media e tasso di entrata leggermente superiore alla media (il primo ha un valore di circa 14,70% e il secondo di circa 6,10%).

Esistono, quindi, differenti comportamenti dei retailers all'interno delle diverse circoscrizioni; risulta di notevole interesse, però, capire quale sia la causa di quest'ultimi. Per poter effettuare questa analisi si è pensato di indagare l'impatto

del posizionamento geografico della singola circoscrizione all'interno della città, e quindi dei negozi in essa presenti.

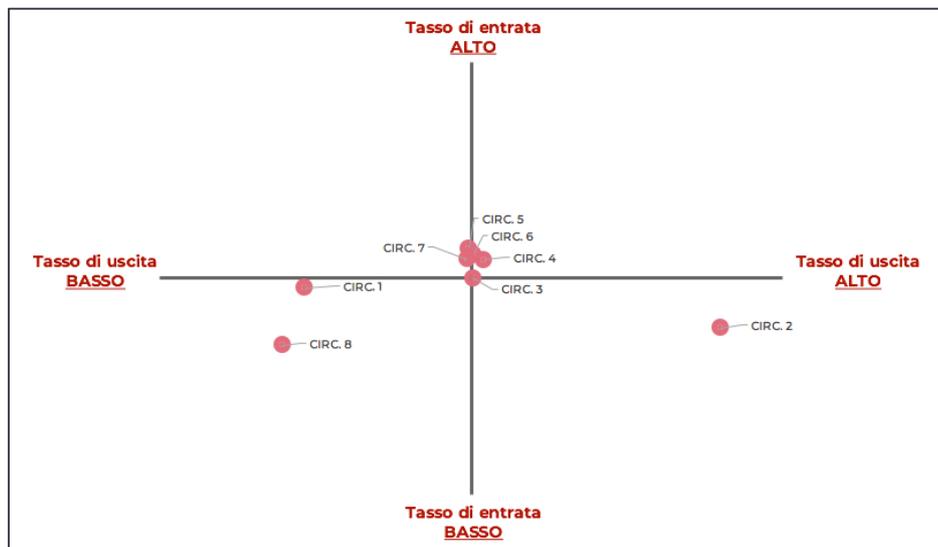


FIGURA 13- MAPPA DI POSIZIONAMENTO QUARTIERI

Si è approssimato l'effetto dell'appartenenza ad una determinata "circoscrizione" con quello della "distanza dal centro", anche vista l'importanza che quest'ultima assume nel mondo del Retail, come mostrato nell'Analisi della letteratura al capitolo 2.

Con il termine *centro* ci si riferisce alle coordinate geografiche del punto centrale della Città di Torino, che risulta corrispondere all'indirizzo di Piazza San Carlo, 201. Dopo aver fissato il centro di riferimento, è semplice il calcolo della distanza dal centro di ogni negozio, mediante l'utilizzo della formula della *distanza euclidea tra due punti in uno piano*¹⁵ riportata sotto.

¹⁵ Con distanza tra due punti in un piano cartesiano, o distanza euclidea, si fa riferimento alla formula che permette il calcolo della distanza tra due punti partendo dalle loro coordinate cartesiane. Per definizione sarà sempre positiva o nulla (nel caso di coincidenza dei due punti), ma mai negativa. La distanza tra due punti risulta essere pari alla radice quadrata della somma del quadrato della differenza delle ascisse e il quadrato della differenza delle ordinate.

Se considerassimo due punti (*Retailer_i* e *Centro*), ognuno individuato dalle proprie coordinate cartesiane [(*x_i*, *y_i*) e (*x_{centro}*, *y_{centro}*)], la loro distanza euclidea (misurata in metri) è pari a:

$$Distanza(Retailer_i, Centro) = \sqrt{(x_{centro} - x_i)^2 + (y_{centro} - y_i)^2}$$

La distanza dal centro è quindi uno dei fattori impattanti che è stato calcolato per ogni retailer presente nel dataset ed inserito all'interno del modello di analisi, come variabile indipendente, per verificare o meno se sia rilevante ai fini della sopravvivenza. Si tratta di una variabile continua che può assumere valori compresi tra 0 e 8,46 chilometri, la cui distribuzione è riportata in Figura 14.

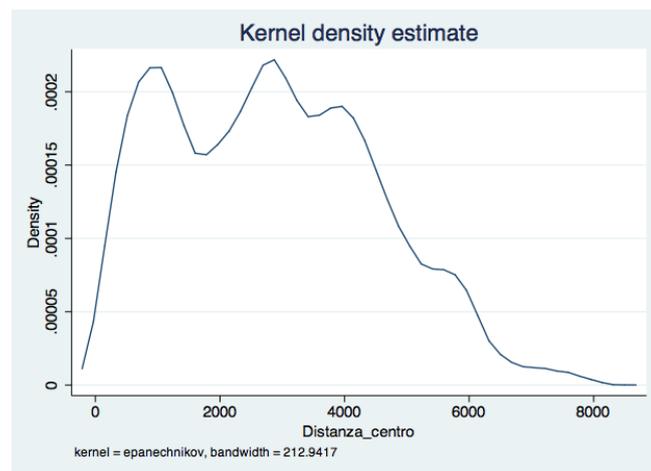


FIGURA 14 - DISTRIBUZIONE DISTANZA DAL CENTRO

6.1.4 FATTORI IMPATTANTI – LA CATEGORIA MERCEOLOGICA DEL RETAILER

Una caratteristica fondamentale che il database fornisce è la *categoria merceologica* a cui il retailer appartiene. Uno studio analogo a quello effettuato per il formato e per il posizionamento è stato eseguito per verificare se l'appartenenza ad una specifica categoria merceologica sia un fattore discriminante o meno nella determinazione dei differenti comportamenti dei retailers.

Per le 21 categorie merceologiche, presentate in Figura 3, utilizzando la metodologia di analisi dei flussi di entrata ed uscita dal mercato (vedi 7.1.2), sono stati calcolati i tassi di entrata e di uscita dei singoli retailers, i rispettivi tassi medi e gli scostamenti rispetto alla media; successivamente si è effettuata un'analisi grafica di questi, presentata nella figura sottostante (Figura 13).

I tassi di entrata ed uscita medi, rispetto ai quali si calcolano gli scostamenti hanno valore pari a **14,84%** e **4,72%** rispettivamente.

Appare chiara l'appartenenza dei retailers a quattro possibili situazioni distinte, corrispondenti ai quattro quadranti della mappa; si verificano infatti delle dinamiche differenti a seconda che il negozio appartenga ad una categoria piuttosto che ad un'altra:

- Retailers che *nascono e muoiono*: si tratta di realtà commerciali soggette ad un alto turnover e che agiscono con dinamismo all'interno del mercato; è il caso dei negozi di **Calzature**, che presentano il più alto tasso di entrata (19,75%) ed uscita (8,63%) in assoluto nel territorio cittadino. Vi sono poi due agglomerati di categorie merceologiche che presentano dei comportamenti parecchio simili tra loro: il primo è quello composto dai negozi di **Articoli per Animali, Telefonia e Fiori e Piante**, che nonostante siano realtà con caratteristiche e prodotti venduti molto differenti, mostrano un tasso di entrata ed uscita dal mercato simile; in particolare, benché si tratti di valori alti in entrambi i casi, il loro scostamento rispetto al valore medio del tasso di entrata è superiore a quello di uscita (il valore dei due tassi è pari a circa 17,80% e 6,8%). Il secondo agglomerato è composto dalle categorie **Articoli per la Casa e Abbigliamento**, le quali hanno, invece, uno scostamento rispetto alla media del tasso di uscita superiore a quello di entrata, con valore dei tassi rispettivamente pari a circa il 16,70% e 8,5%. Infine, vi sono i negozi di **Giocattoli e Mobili**, per i

quali i fenomeni di uscita dal mercato si presentano maggiori di quelli di ingresso, anche se con valori molto vicini alla media.

- Retailers che *non nascono e muoiono*: si tratta di settori in sofferenza e crisi, per i quali non si verificano nuovi ingressi nel mercato, nonostante si verifichi la morte dei negozi esistenti; è il caso dei negozi di **Elettronica**, che manifestano un alto tasso di uscita rispetto alla media (valore pari a 7,62%), a fronte di un tasso di entrata con valore esattamente pari a quello medio. Rientrano in questo quadrante anche i negozi di **Alimentari**, che ricordando quello che è stato enunciato nel capitolo 5, sono tutti quegli esercizi di vicinato, definiti “Mom-and-Pops stores”, che non godono dell’appartenenza ad un network o brand e, pertanto, sono sorretti solo dalle proprie vendite e dal proprio fatturato; questa tipologia di esercizi commerciali presenta un tasso di uscita dal mercato superiore al valore medio (5,92%) e un tasso di ingresso in esso inferiore alla media (14,44%).
- Retailers che *non nascono e non muoiono*: si tratta di situazioni più stazionarie o di mercati saturi, come ad esempio quello dei negozi di **Quotidiani** e **Tabacchi**; queste due categorie hanno una dinamica piuttosto simile e presentano il più basso tasso di ingresso sul mercato (pari a circa il 9%). Vi sono poi, gli **Ipermercati**, che a fronte di un basso tasso di ingresso sul mercato, presentano anche il più basso tasso di uscita da esso (esattamente pari a 0, il che implica che nessun Ipermercato nato in Torino è mai morto). Intermedia è la situazione dei **Supermercati** e delle **Farmacie**, che mostrano un comportamento simile, con degli scostamenti rispetto alla media assimilabili e con valore negativo (pari a circa il -4,5% nel caso del tasso di ingresso e -3,5% nel caso di quello di uscita). Le **Librerie** si trovano anch’esse in questo quadrante presentando un tasso di entrata sul mercato di 12,38% e uno di uscita pari al 4,15%. Hanno scostamenti negativi rispetto alla media anche i **Minimercati**, con valori pari al 13,16% nel caso dei fenomeni di ingresso e 4,33% nel caso di quelli

di uscita. Situazione esattamente pari alla media nel caso del tasso di uscita dal mercato è quella delle **Cartolerie**, nonostante presentino uno scostamento negativo rispetto alla media, -1,30%, del tasso di ingresso.

- Retailers che *nascono e non muoiono*: rientrano in questo quadrante i negozi di **Articoli Funerari**, i quali mostrano uno scostamento positivo rispetto alla media riguardo il tasso di entrata (quest'ultimo è pari a 16,40%) e uno negativo relativamente al tasso di uscita (quest'ultimo è pari a 3,13%). Vi sono poi i negozi di **Ottica**, i quali hanno un tasso di uscita dal mercato nella media e uno scostamento rispetto alla media del tasso di entrata nel mercato pari al 2,22%. I negozi di **Profumeria e Cosmetica** si trovano sulla mappa esattamente nel punto di intersezione tra gli assi, ovvero hanno degli scostamenti nulli rispetto alla media.

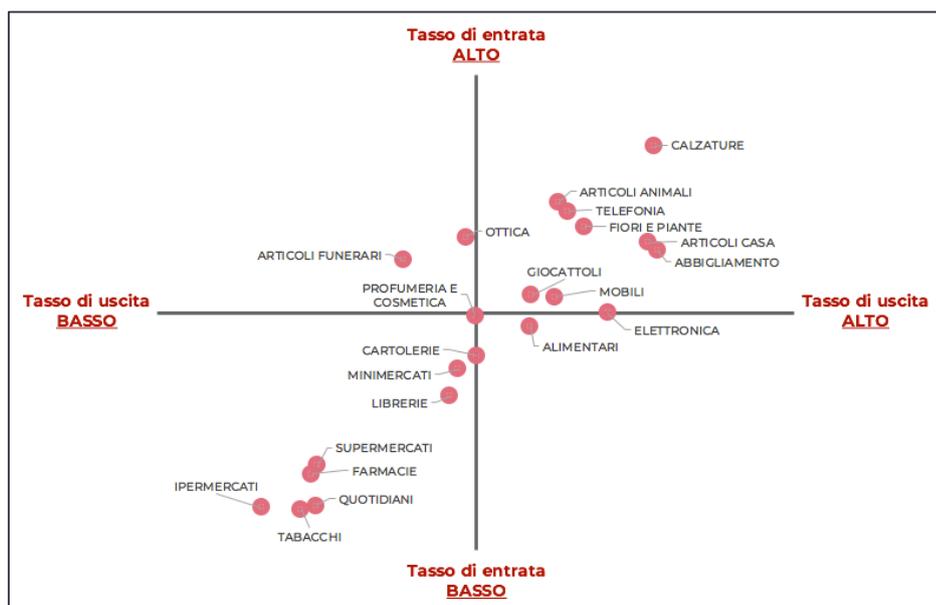


FIGURA 15- MAPPA DI POSIZIONAMENTO CATEGORIE MERCEOLOGICHE

Dall'analisi sopra enunciata risulta evidente come la categoria merceologica sia un fattore fondamentale da tenere in considerazione per motivare le scelte e le dinamiche dei singoli retailers. Dei negozi possono mostrare dei fenomeni completamente differenti relativamente alla sopravvivenza, semplicemente per il tipo di prodotto o servizio che forniscono. È questa la motivazione per cui si è

deciso strutturare l'analisi statistica tenendo in considerazione, come fattore discriminante fondamentale, la categoria merceologica, che rientra nel modello come variabile indipendente. Si tratta di inserire all'interno del modello statistico 21 variabili binarie (0,1), che assumono valore 1 nel caso in cui il negozio in questione è appartenente alla categoria merceologica considerata e 0 in tutti gli altri casi.

6.1.5 FATTORI IMPATTANTI – L'AGGLOMERAZIONE TRA CATEGORIE MERCEOLOGICHE

Dopo aver compreso l'importanza che la categoria merceologica assume ai fini dell'analisi condotta, si è deciso di indagare il posizionamento delle stesse all'interno della città di Torino. Nello specifico, mediante l'utilizzo del software QGis (per la cui descrizione e funzionamento si rimanda al punto 5.2) si è effettuata un'analisi grafica, al fine di mostrare una fotografia al 2017 della distribuzione delle singole categorie merceologiche all'interno della città di Torino; i pallini colorati sulla mappa rappresentano i retailers ad esse appartenenti. L'istantanea al 2017 è stata ottenuta applicando un doppio filtro, sia sulla categoria d'interesse che sui retailers che al 2017 risultano non "morti", ovvero tutti quei casi per cui non si sia verificata una chiusura.

Si riportano a titolo esemplificativo due categorie appartenenti a settori differenti e che presentano comportamenti altrettanto differenti: **Abbigliamento** e **Supermercati**, mostrati in Figura 14 e 15. Si rimanda all'Appendice A per la visualizzazione del posizionamento delle restanti categorie merceologiche.

Osservando le due mappe appare evidente come, a prescindere dalla numerosità dissimile del campione, i due casi presentati presentino delle dinamiche di localizzazione non analoghe. Il motivo può risiedere nel fatto che si stanno osservando delle categorie merceologiche differenti, dal punto di vista della tipologia di bene venduto.

Come già descritto nell'analisi della letteratura, al paragrafo 2.2.1, esistono differenti classificazioni di beni: i negozi di Abbigliamento vendono beni, definiti da (Holton, 1958) "*shopping*", per i quali il consumatore trae un vantaggio tanto superiore quanto maggiore è la possibilità di effettuare una comparazione tra prodotti. Pertanto, una maggiore concentrazione di offerta commerciale risulta maggiormente attrattiva agli occhi del cliente finale, perché gli consente di effettuare una scelta più oculata, valutando un maggior numero di articoli e godendo di una maggiore varietà a disposizione, in uno spazio fisico più contenuto e con un impiego di tempo inferiore. Questa è la motivazione che, in generale, sta alla base della scelta, da parte dei negozi di Abbigliamento, di agglomerarsi. Nello specifico, anche semplicemente da una prima osservazione della disposizione dei negozi sulla città (in Figura 14) si può vedere, come questa scelta venga rispettata anche nel caso della città di Torino.

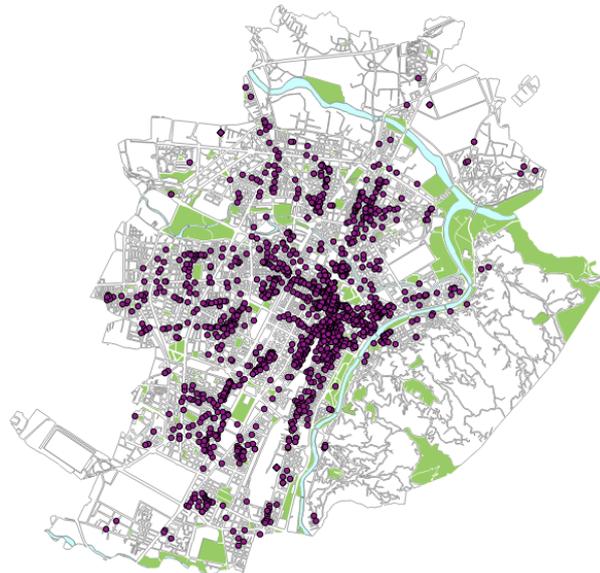


FIGURA 16 - Istantanea al 2017 Abbigliamento

Contrariamente a ciò, i Supermercati, mostrati in Figura 15, forniscono prodotti, definiti da (Holton, 1958), "*convenience*", ovvero beni di largo consumo, che i consumatori comprano frequentemente ed immediatamente, quindi senza effettuare confronti. Pertanto, negozi che vendono questa tipologia di prodotti,

non traggono beneficio nel concentrarsi geograficamente in punti vicini dello spazio.



FIGURA 17 - Istantanea al 2017 SUPERMERCATI

L'analisi grafica suddetta ha reso visibile come le logiche di posizionamento dei retailers siano differenti tra loro. Si tratta di un'ulteriore differenza riscontrata ed è sembrato opportuno approfondirla meglio, al fine di comprendere se diverse scelte portino a risultati differenti, in termini di sopravvivenza.

Dall'analisi della letteratura, si è visto come l'agglomerazione possa assumere due forme differenti, come affermato da (Fox, et al., 2007):

- » **Agglomerazione "intra-type", o intra-categoria:** quando si considera il fenomeno di clusterizzazione geografica di negozi appartenenti alla stessa categoria merceologica.
 - » Per lo studio di questo tipo di fenomeno è stato introdotto all'interno dell'analisi statistica un indice, definito A_i , capace di calcolare quanto maggiore, o minore, sia la *probabilità* di trovare attorno ad un generico negozio i , appartenente ad una specifica categoria X , un negozio della sua stessa categoria, piuttosto che uno appartenente ad una qualsiasi categoria generica.

Per comprendere come questo indice sia stato ottenuto, occorre tenere in considerazione un'approssimazione effettuata sulla forma della città di Torino, assunta **circolare** e composta da cerchi concentrici.

Si procede adesso con una descrizione dei passi eseguiti per ricavare l'indice suddetto.

Sia i il *generico negozio della categoria X*; attorno ad ogni i si traccia un cerchio immaginario di raggio **600 metri** (la motivazione di quest'ultima scelta deriva dall'analisi dello studio operato da (Sevtzuk, 2014), il quale ritiene che dieci minuti di passeggiata, corrispondenti a 600 metri, siano il massimo tempo/distanza che il consumatore è disposto a percorrere per effettuare dei confronti tra beni simili).

Come detto più volte, ogni istanza del dataset è geolocalizzata sulla città di Torino, pertanto si conosce la relativa latitudine e longitudine. È possibile calcolare, quindi, il numero di esercizi commerciali che il generico negozio i ha attorno, nell'arco di 600 metri, nel suo anno di apertura, ovvero, nel momento in cui la scelta di localizzazione è stata effettuata. Questa operazione è stata effettuata sia per calcolare il numero di negozi totali attorno ad i , sia per ricavare quelli appartenenti alla stessa categoria di i , ovvero X .

In particolare, è stata calcolata, mediante l'applicativo Excel, la distanza tra i e qualsiasi altro negozio attivo nell'anno della nascita di i (o solamente della categoria X nel secondo caso) tramite la formula della **distanza tra due punti individuati da coordinate geografiche (D)**, calcolata in metri:

$$D = \arccos\{\cos [\text{rad}(90 - A_1)] * \cos (\text{rad}(90 - B_1)) \\ + \sin (\text{rad}(90 - A_1)) \\ * \sin(\text{rad} - B_1) * \cos (\text{rad}(A_2 - B_2))\} * 6371 * 1000$$

Dove: A_1 e A_2 sono rispettivamente la latitudine e la longitudine del primo punto (il negozio i), mentre B_1 e B_2 del secondo.

Considerando solo le distanze inferiori o uguali a 600 si ottiene proprio il numero di negozi presenti nell'area di interesse.

Per ogni i si può, quindi, calcolare la probabilità di trovare attorno ad esso un negozio della categoria X ($P_{i,X}$) mediante la formula:

$$P_{i,X} = \frac{N^\circ \text{ negozi } X \text{ attorno a } i}{N^\circ \text{ negozi } X \text{ in Torino}}$$

Allo stesso modo si può ottenere la probabilità di trovare attorno ad i un negozio di una generica categoria merceologica (P_i), come segue:

$$P_i = \frac{N^\circ \text{ negozi attorno a } i}{N^\circ \text{ negozi in Torino}}$$

L'indice A_1 sarà dato dal rapporto delle due probabilità:

$$A_1 = \frac{P_{i,X}}{P_i}$$

Possono verificarsi tre situazioni distinte:

- $A_1 < 1$, che indica la maggiore probabilità di trovare attorno al negozio i un negozio generico, piuttosto che uno della categoria X ;
 - $A_1 = 1$, nel caso in cui le due probabilità siano equivalenti;
 - $A_1 > 1$, se attorno al negozio i è più probabile che si trovi un retailer della categoria X , piuttosto che uno generico negozio; questo avviene quando esiste un'agglomerazione intra-categoria.
- » In alternativa, il fenomeno può essere studiato inserendo nel modello il **Numero di Vicini della categoria** ($N_{\text{vicini, cat } X}$), per comprendere se sia semplicemente il fatto di avere vicino retailers simili che aiuta a sopravvivere.

- **Agglomerazione "inter-type", o inter-categoria:** quando si considera il fenomeno di concentrazione geografica di differenti tipologie di negozi, appartenenti a qualsiasi categoria merceologica.

» Per lo studio di questo tipo di fenomeno è stato introdotto all'interno del modello statistico un indice, chiamato A_2 , che calcola quanto maggiore, o minore, sia la *densità* di negozi attorno ad i , rispetto alla generica densità presente all'interno della città. A differenza dell'indice presentato prima non si effettua più un rapporto di probabilità, bensì di densità, che viene calcolata come il rapporto tra numero di negozi attorno ad i , o il numero di negozi in Torino, e l'area su cui questi sono localizzati. Nel primo caso l'area non sarà altro che la superficie di un cerchio di raggio 600, ovvero πr^2 (1.130.973,36 m²); nel secondo si fa riferimento all'area della città di Torino, calcolata per ogni anno, come la superficie di un cerchio di raggio r , dove r è ottenuto individuando il retailer più distante dal centro nell'anno di interesse.

La densità di negozi attorno ad i (ρ_i), e la densità di negozi generica in Torino (ρ_{Torino}) si ottengono come segue:

$$\rho_i = \frac{N^\circ \text{ negozi attorno a } i}{\text{AREA attorno ad } i}$$

$$\rho_{\text{Torino}} = \frac{N^\circ \text{ negozi Torino}}{\text{AREA Torino}}$$

L'indice A_2 sarà dato dal rapporto delle due densità:

$$A_2 = \frac{\rho_i}{\rho_{\text{Torino}}}$$

Anche in questo caso, possono verificarsi tre situazioni distinte:

- $A_2 < 1$, che indica la minore densità di negozi attorno ad i , rispetto alla generica densità presente in Torino;
- $A_2 = 1$, nel caso in cui le due densità siano equivalenti;

- $A_2 > 1$, se attorno al negozio i vi è una maggiore densità commerciale, rispetto a quella che si trova in città; questo è quello che avviene quando esiste un'agglomerazione inter-categoria.
- » In alternativa, il fenomeno può essere studiato inserendo nel modello il **Numero di Vicini** (N_{vicini}), per comprendere se sia semplicemente il fatto di avere traffico attorno che aiuta a sopravvivere.

Gli indici A_1 e A_2 , e i valori $N_{vicini,catX}$ e N_{vicini} , sono stati calcolati per ogni retailer presente all'interno del database ed inseriti nel modello di regressione, che verrà di seguito descritto.

A_1 è una variabile indipendente continua che assume valori compresi tra 0, nel caso in cui il negozio i non abbia alcun negozio della sua stessa categoria attorno, e 137,03 (quest'ultimo valore corrisponde, ad esempio, ad un negozio di Articoli Funerari in Via Agostino Bertani 103, attorno al quale si trovano solamente tre negozi, due dei quali appartengono alla medesima categoria merceologica).

A_2 è una variabile indipendente continua che assume valori compresi tra 0, nel caso in cui il negozio i non abbia alcun negozio attorno, e 29,47 (quest'ultimo è il caso di un negozio di Abbigliamento in Via G.L. Lagrange 7, che ha una densità di circa 955 negozi per chilometro quadrato attorno a sé, a fronte dei circa 32 negozi per chilometro quadrato presenti genericamente in Torino).

6.2 ANALISI DI SOPRAVVIVENZA

Sono stati, finora, presentati i singoli step che hanno permesso di comprendere l'esistenza di fattori fondamentali che un retailer deve tenere in considerazione per affrontare le proprie decisioni. Per ogni fattore è stata derivata una variabile, che consente, una volta inserita in un modello di regressione, di misurarne gli effetti.

Viene descritta adesso la vera e propria analisi statistica effettuata, la Survival Analysis, per la cui descrizione si rimanda al capitolo 5.

In prima istanza, si è effettuato studio generico sull'andamento della sopravvivenza dei retailers in Torino, interrogando il database in modo da ottenere il primo output che la Survival Analysis mette a disposizione: la **Life Table** e il corrispondente grafico della **Curva di Sopravvivenza**; i rispettivi output vengono mostrati nelle figure sottostanti (Figura 18 e 19).

Tabella vita ^a												
Ora di inizio intervallo	Numero di entrati nell'intervallo	Numero di rimossi nell'intervallo	Numero di esposti al rischio	Numero di eventi terminali	Proporzione di terminazione	Proporzione di sopravvivenza	Proporzione di sopravvivenza cumulativa al termine dell'intervallo	Errore standard della proporzione di sopravvivenza cumulativa al termine dell'intervallo	Densità di probabilità	Errore standard di densità di probabilità	Tasso di rischio	Errore standard di tasso di rischio
0	18935	0	18935,0	641	0,03	0,97	0,97	0,00	0,034	0,001	0,03	0,00
1	18294	810	17889,0	1585	0,09	0,91	0,88	0,00	0,086	0,002	0,09	0,00
2	15899	750	15524,0	1215	0,08	0,92	0,81	0,00	0,069	0,002	0,08	0,00
3	13934	800	13534,0	987	0,07	0,93	0,75	0,00	0,059	0,002	0,08	0,00
4	12147	742	11776,0	653	0,06	0,94	0,71	0,00	0,042	0,002	0,06	0,00
5	10752	642	10431,0	611	0,06	0,94	0,67	0,00	0,042	0,002	0,06	0,00
6	9499	625	9186,5	461	0,05	0,95	0,64	0,00	0,034	0,002	0,05	0,00
7	8413	548	8139,0	367	0,05	0,95	0,61	0,00	0,029	0,001	0,05	0,00
8	7498	483	7256,5	297	0,04	0,96	0,58	0,00	0,025	0,001	0,04	0,00
9	6718	562	6437,0	233	0,04	0,96	0,56	0,00	0,021	0,001	0,04	0,00
10	5923	692	5577,0	167	0,03	0,97	0,54	0,00	0,017	0,001	0,03	0,00
11	5064	581	4773,5	146	0,03	0,97	0,53	0,00	0,017	0,001	0,03	0,00
12	4337	537	4068,5	121	0,03	0,97	0,51	0,00	0,016	0,001	0,03	0,00
13	3679	3679	1839,5	0	0,00	1,00	0,51	0,00	0,000	0,000	0,00	0,00

a. Il tempo di sopravvivenza mediano è 13,00

FIGURA 18 - LIFE TABLE

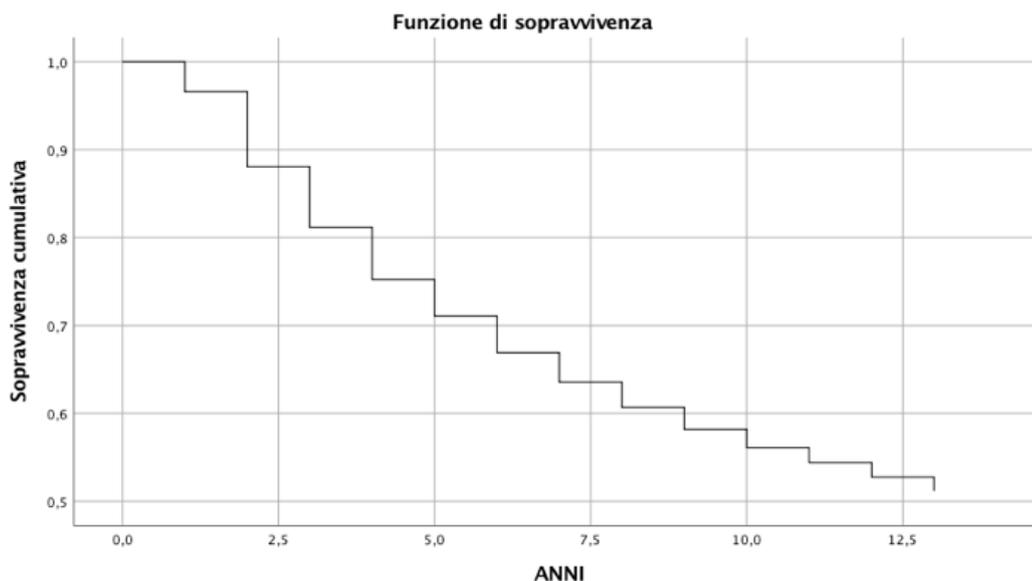


FIGURA 19 - CURVA DI SOPRAVVIVENZA

Effettuando un focus sull'andamento della curva di sopravvivenza, si può comprendere meglio quale sia la probabilità che un retailer, localizzato in Torino e avente X anni di età, sopravviva all'anno $X+1$. Nello specifico, il grafico in Figura 19 mostra come, a titolo esemplificativo, un negozio vivo da 2 anni e mezzo abbia una probabilità di sopravvivere superiore all'85%; questa probabilità si riduce a circa il 55% nel caso in cui il negozio sia vivo da 10 anni.

Per le motivazioni enunciate al paragrafo 6.1.4, si è deciso, però, di strutturare l'analisi utilizzando come fattore principale di discriminazione la categoria merceologica, che identifica in una qualche maniera l'identità del retailer stesso. È proprio dallo studio dell'impatto della categoria merceologica sulla sopravvivenza che ha inizio l'analisi.

6.2.1 LIFE TABLE E COX REGRESSION – CATEGORIA MERCEOLOGICA

Per ciascuna delle ventuno categorie merceologiche ci si è chiesti quale sia la probabilità che un retailer sopravviva per il semplice fatto di appartenere ad essa.

La *prima Research Question* che ci si è posti può essere enunciata come segue:

RQ1, “Quanto sopravvive un negozio della categoria X, rispetto a tutti i negozi che non appartengono alla categoria X”?

Per rispondere a questa domanda, si può pensare di utilizzare lo strumento della **Life Table**, che consente, inserendo un *fattore* all’interno dell’analisi (rappresentato proprio dalla categoria merceologica), di visualizzare graficamente la sovrapposizione di due curve di sopravvivenza, derivanti da due differenti casi: *fattore* = 1 e *fattore* = 0. La curva corrispondente al *fattore* = 1, solitamente colorata in **Rosa**, è quella che mostra la sopravvivenza dei negozi che appartengono alla specifica categoria merceologica di interesse; viceversa, quella corrispondente al *fattore* = 0, solitamente colorata in **Azzurro**, fa riferimento alla sopravvivenza di tutti i negozi che non appartengono alla categoria in analisi.

Da un’analisi di questo tipo possono emergere **tre** situazioni distinte:

1. La curva rosa (categoria=1) si trova sopra la curva azzurra (categoria=0): questa casistica mostra che tutti i negozi appartenenti alla categoria d’interesse hanno una curva di sopravvivenza, e quindi una probabilità di sopravvivere, maggiore di tutti quelli appartenenti alle restanti venti categorie; la differenza tra le due curve è, inoltre, statisticamente significativa¹⁶;
2. Le due curve sono sovrapposte: questa è la circostanza che si verifica quando l’appartenenza ad una specifica categoria merceologica porta ad una probabilità di sopravvivere non differente dalla media; si tratta di casi in cui la differenza tra le due curve non è statisticamente significativa;

¹⁶ Il metodo utilizzato per appurare la differenza significativa tra le due curve è il **Test di Wilcoxon**: è il test utilizzato per comparare le distribuzioni di sopravvivenza tra gruppi differenti; il test statistico è basato sulla differenza tra i valori medi dei due gruppi. Si è deciso di utilizzare questa tipologia di test, piuttosto che il *LogRank Test*, perché il primo è sensibile a piccole variazioni tra dati di sopravvivenza, mentre il secondo a grandi.

3. La curva rosa si trova sotto quella azzurra: questa casistica mostra che tutti i negozi appartenenti alla categoria d'interesse hanno una curva di sopravvivenza, e quindi una probabilità di sopravvivere, minore di tutti i negozi appartenenti alle restanti venti categorie; la differenza tra le due curve è, inoltre, statisticamente significativa.

A titolo esemplificativo, si riportano le Curve di Sopravvivenza dei due casi descritti precedentemente, Abbigliamento e Supermercati, e si rimanda all'Appendice B per la visualizzazione di tutte le restanti categorie.

La Figura 20 mostra la curva di sopravvivenza dei retailers di Abbigliamento; appare evidente come questo esempio rientri nella terza casistica: la curva di sopravvivenza rosa si trova, infatti, sotto quella azzurra; pertanto, tutti i negozi di abbigliamento hanno una probabilità di sopravvivere inferiore, per il fatto stesso di appartenere a questa categoria merceologica, rispetto a tutti quelli che invece non vi appartengono.

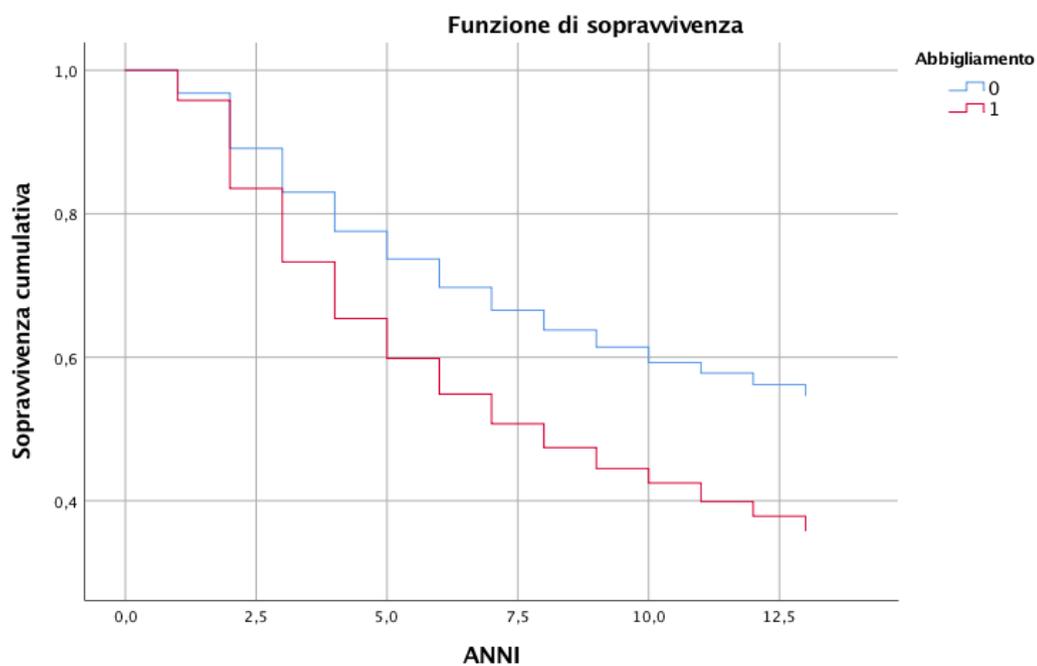


FIGURA 20 – CURVA DI SOPRAVVIVENZA ABBIGLIAMENTO

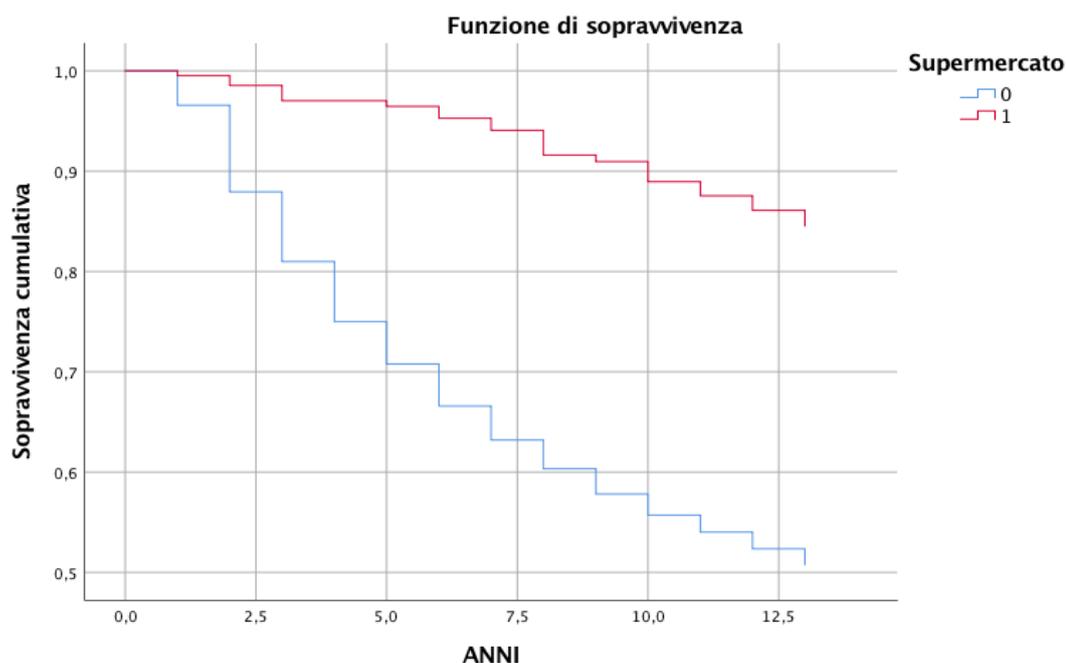


FIGURA 21 - CURVA DI SOPRAVVIVENZA SUPERMERCATI

Differente è la situazione mostrata in Figura 21, dove è semplice verificare che il fatto di essere un Supermercato porta il retailer ad avere una curva di sopravvivenza “molto” maggiore del non esserlo.

Rientrano nel primo caso categorie come Articoli Funerari, Cartolerie, Farmacie e Parafarmacie, Ipermercati (che mostrano una curva di sopravvivenza costante e pari ad 1, in quanto tutti vivi!), Librerie, Minimercati, Profumeria e Cosmetica, Quotidiani e Periodici e Tabacchi.

Le due curve non sono significativamente differenti a livello statistico nel caso di categorie come Alimentari, Articoli per Animali, Fiori e Piante, Giocattoli, Mobili, Ottica e Telefonia.

Il terzo caso (curva rosa sotto la curva azzurra) è rappresentativo di categorie come Articoli per la Casa, Calzature ed Elettronica.

Durante la citazione si è evidenziata la parola “molto”, per mettere in risalto un inconveniente di questo tipo di analisi: le informazioni che la Life Table, e la

relativa curva di sopravvivenza, forniscono, sono di tipo qualitativo e, inoltre, permettono di effettuare delle valutazioni esclusivamente sulla singola categoria merceologica; non è possibile, quindi, mediante l'utilizzo di questo strumento misurare la *distanza* tra le due curve, per fornire dei parametri quantitativi, né effettuare *confronti* tra categorie merceologiche differenti. Si tratta, comunque, di un ottimo strumento da utilizzare in prima analisi, per fornire una valutazione qualitativa del comportamento del fattore studiato; occorre, però, approfondire meglio lo studio, mediante l'utilizzo di ulteriori tecniche.

Per effettuare dei confronti tra categorie ed ottenere gli effetti quantitativi sulla probabilità di sopravvivenza dei retailers è stato utilizzato lo strumento della **Cox Regression**. Occorre prestare attenzione al procedimento con cui si è condotta l'analisi: le singole categorie (*variabili indipendenti binarie*) non sono state inserite in blocco all'interno del modello, bensì si è optato per effettuare delle analisi a sé stanti, immettendo una categoria per volta; solo così, infatti, è possibile effettuare dei paragoni, in quanto i valori, di volta in volta ottenuti, indicano quanto maggiore o minore sia la probabilità di sopravvivere del negozio per il fatto di appartenere alla specifica categoria merceologica, rispetto al valore medio (ottenuto da tutte le categorie meno quella di interesse).

L'analisi di regressione suddetta ha l'obiettivo di stimare la probabilità di sopravvivenza del retailer in funzione di determinati parametri. Il modello Cox riceve, pertanto, in input le informazioni riguardanti il numero di **Anni di vita** del retailer stesso (al momento in cui l'analisi è stata condotta: anno 2017) e la *variabile dipendente binaria* **Morte**, che presenta valore 0 nel caso in cui il retailer al 2017 sia ancora vivo e 1 nel caso contrario; vengono inserite inoltre tutte le *variabili indipendenti* che si vogliono utilizzare per il calcolo della stima.

Per una valutazione quantitativa degli effetti occorre guardare all'*hazard ratio* (**Exp(B)**), ovvero al rischio di morte, che ci permette di rispondere alla domanda:

“Essere un negozio della categoria X di quanto permette di ridurre (o aumentare) il rischio di morte?”

Ricordiamo che, come descritto nel capitolo 5, dall’analisi dell’hazard ratio, possono emergere tre situazioni distinte:

1. $\text{Exp}(B) > 1$: ci si trova in questo caso ogni qual volta una variazione della variabile in questione porta ad un *aumento* del rischio di morte;
2. $\text{Exp}(B) = 1$: in questo caso una variazione della variabile in questione non ha impatto sul rischio di morte; i due fenomeni sono pertanto indipendenti;
3. $\text{Exp}(B) < 1$: ci si trova in questo caso ogni qual volta una variazione della variabile in questione porta ad una *riduzione* del rischio di morte.

Nel caso dello studio in esame l’hazard ratio indica la variazione del rischio di morte che si verifica nel passaggio da Categoria=0 a Categoria=1.

Vengono presentati adesso gli output ottenuti (mediante il software SPSS) e l’interpretazione di quest’ultimi.

Relativamente alle categorie che si è deciso di utilizzare per scopi esemplificativi (Abbigliamento e Supermercati) vengono presentate le tabelle sottostanti (Tabella 1 e 2). Si rimanda all’Appendice C per la visualizzazione degli output delle restanti categorie merceologiche¹⁷.

¹⁷ Occorre prestare attenzione al fatto che la categoria Ipermercati non è stata inserita in questo tipo di analisi, in quanto non esistendo istanze con valore di morte pari a uno, sarebbe impossibile una computazione dei valori.

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Abbigliamento	0,481	0,027	319,880	1	0,000	1,618

TABELLA 1 - SOPRAVVIVENZA ABBIGLIAMENTO

La Tabella 1 mostra un valore dell'hazard ratio maggiore di uno; questo implica che l'essere un negozio di Abbigliamento conduce ad un aumento del rischio di morte per il retailer. In particolare l'output ci dice che nel passaggio da 0 a 1 della variabile indipendente il rischio di morte aumenta del 61,8%.

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Supermercati	-1,486	0,200	55,016	1	0,000	0,226

TABELLA 2 - SOPRAVVIVENZA SUPERMERCATI

Viceversa, la Tabella 2 mostra un valore dell'hazard ratio minore di uno; questo implica che l'essere un Supermercato conduce ad una riduzione del rischio di morte per il retailer. Nel passaggio da 0 a 1 della variabile indipendente il rischio di morte si riduce del 77,4% ($1 - 0,226 = 0,774$).

Come detto, in Appendice C si possono trovare gli output delle restanti categorie merceologiche; per effettuare dei confronti, però, si è deciso di riportare una tabella riassuntiva dove vengono presentate solo le categorie che mostrano valori significativi. Nella tabella sottostante (Tabella 3) le categorie sono ordinate per hazard ratio crescente. Si è effettuata una separazione tra quelle categorie, che per il fatto di essere tali, **riducono il rischio di morte** rispetto al non esserlo (evidenziate in verde) e quelle che lo **aumentano** (evidenziate in rosso).

Occorre far attenzione al fatto che, nonostante categorie come Calzature, Articoli per la Casa e Abbigliamento portano ad un aumento del rischio di morte, si è visto anche, dall'analisi delle categorie in Torino (al paragrafo 6.1.4), come queste facciano parte del quadrante dei retailers che *nascono e muoiono*. Si tratta,

pertanto, di categorie soggette ad un alto turnover; non bisogna attribuire, quindi, un'accezione negativa a quest'ultimo risultato ottenuto.

Differente, e peggiore, è la situazione dei negozi di Elettronica, che oltre ad aumentare il rischio di morte, hanno anche un tasso di entrata sul mercato esattamente pari al valore medio (sempre dall'analisi al 6.1.4).

	Sign.	Exp(B)	Aumento (+) o diminuzione (-) del rischio di morte
Tabacchi	0,000	0,158	- 84,2%
Farmacia e Parafarmacia	0,000	0,200	- 80,0%
Quotidiani	0,000	0,221	- 77,9%
Supermercati	0,000	0,226	- 77,4%
Articoli Funerari	0,001	0,490	- 51,0%
Librerie	0,001	0,734	- 26,6%
Profumeria e Cosmetica	0,000	0,767	- 23,3%
Cartolerie	0,002	0,796	- 20,4%
Calzature	0,001	1,331	+ 33,1%
Elettronica	0,000	1,346	+ 34,6%
Articoli per la Casa	0,000	1,440	+ 44,0%
Abbigliamento	0,000	1,618	+ 61,8%

TABELLA 3 - RIASSUNTO SOPRAVVIVENZA CATEGORIE

La categoria merceologica è il *primo driver* della sopravvivenza individuato, infatti determina l'identità del retailer stesso e ne può ridurre o aumentare il rischio di morte.

6.2.2 COX REGRESSION – DRIVERS DEL GENERICO NEGOZIO IN TORINO

I fattori impattanti e le rispettive variabili indipendenti individuate, emersi nel capitolo 6.1, sono stati inseriti all'interno del modello:

- **Densità di popolazione**, che misura il numero di persone per metro quadro;
- **Dimensione del retailer**, misurata in metri quadri;

- **Distanza dal centro**, misurata in metri;
- **Indice di agglomerazione A₁ e Numero vicini della categoria X**;
- **Indice di agglomerazione A₂ e Numero vicini generico**.

Prima di condurre l'analisi per le singole categorie merceologiche, si è studiato l'impatto delle variabili indipendenti sopra enunciate sul *generico negozio* in Torino. Si riportano in Tabella 4 e 5 gli output relativi, dove in rosso vengono evidenziate le variabili non significative. La differenza tra le due tabelle consiste solamente nel modo in cui l'agglomerazione è stata calcolata.

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Densità di Popolazione	0,010	0,005	3,361	1	0,067	1,010
Metri Quadri	-0,001	0,000	108,998	1	0,000	0,999
Distanza dal Centro	0,012	0,011	1,303	1	0,254	1,012
A1	-0,010	0,009	1,165	1	0,280	0,990
A2	0,006	0,005	1,131	1	0,288	1,006

TABELLA 4 - OUTPUT GENERICO 1

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Densità di Popolazione	0,009	0,005	2,697	1	0,101	1,009
Metri Quadri	-0,001	0,000	108,031	1	0,000	0,999
Distanza dal centro	0,024	0,010	5,945	1	0,015	1,024
Nvicini	0,000	0,000	0,465	1	0,495	1,000
Nvicini,categoria	0,002	0,000	40,209	1	0,000	1,002

TABELLA 5 - OUTPUT GENERICO 2

La *seconda Research Question* che ci si è posti può essere enunciata come segue:

RQ2, "la densità di popolazione presente nella zona in cui il Retailer è localizzato è un fattore predittivo della probabilità di sopravvivenza di quest'ultimo?"

In entrambi i casi, la densità di popolazione non risulta essere un parametro significativamente predittivo della probabilità di sopravvivenza del generico

negozio in Torino. Una motivazione di questo risultato può risiedere nel fatto che si ha carenza di dati a disposizione; infatti, i valori di popolazione sono associati alle circoscrizioni e non alla singola zona statistica di appartenenza del negozio. Il fatto di avere solo otto valori differenti, porta ad una sistematica non significatività del valore della variabile dipendente, che per questo motivo si è deciso, quindi, di escludere dall'analisi condotta; non si può, quindi, dare una risposta alla RQ2.

La terza *Research Question* che ci si è posti può essere enunciata come segue:

RQ3, “la dimensione del Retailer è un fattore predittivo della probabilità di sopravvivenza di quest’ultimo?”

La dimensione, come mostrato in Tabella 5, è un fattore predittivo statisticamente significativo (al 99%) della probabilità di sopravvivenza di un retailer. Alla variabile “Metri Quadri” corrisponde un beta di -0,001 e un relativo hazard ratio, o $\text{Exp}(B)$, di 0,999. Questo implica che: se il retailer aumentasse la sua dimensione di un'unità, pertanto la variazione della variabile indipendente sarebbe di 1, il rischio di morte sarebbe ridotto del 0,1%. Vediamo altri due esempi, corrispondenti ad una variazione rispettivamente di 50 e di 100:

- Se il retailer variasse la sua dimensione di 50 metri quadri (passando ad esempio da 10 a 60 mq), il hazard ratio sarebbe di: $HR = e^{(-0,001*50)} = 0,951$, a cui corrisponde una riduzione del rischio di morte di $1-HR$, ovvero del 4,9%;
- Se il retailer variasse la sua dimensione di 100 metri quadri (passando ad esempio da 10 a 110 mq), il hazard ratio sarebbe di: $HR = e^{(-0,001*100)} = 0,904$, a cui corrisponde una riduzione del rischio di morte di $1-HR$, ovvero del 9,6%.

La risposta alla **RQ3** è, quindi, affermativa; la dimensione è il *secondo driver* della sopravvivenza individuato e, in particolare, può essere definito un driver con **impatto positivo**, in quanto *all'aumentare di questo, il rischio di morte viene ridotto*.

Si passa adesso alla *quarta Research Question*, **RQ4**: *“la distanza dal centro del Retailer è un fattore predittivo della probabilità di sopravvivenza di quest'ultimo?”*

Come mostrato dalle tabelle 4 e 5, la dimensione non è un fattore statisticamente significativo della probabilità di sopravvivere del generico retailer in Torino.

In riferimento all'agglomerazione, invece, ci si è posti due differenti domande:

- **RQ5**: *“il fatto di avere attorno al negozio i (della categoria X) tanti Retailer della sua stessa categoria ha impatto sulla probabilità di sopravvivere dello stesso?”*

A questa domanda, come descritto al paragrafo 6.1.5, si può rispondere guardando all'impatto dell'indice A_1 e del parametro $N_{vicini,categoria}$:

- » L'indice A_1 non è un fattore statisticamente significativo della probabilità di sopravvivere del generico retailer in Torino.
- » Il parametro $N_{vicini,categoria}$ ha invece un effetto significativo e, in particolare, ha un impatto **negativo** sul rischio di morte; all'aumentare del numero di negozi della stessa categoria del retailer i attorno ad esso, il rischio di morte di quest'ultimo aumenta. Nello specifico, dato che il beta ha un valore pari a 0,002 e l' $\exp(\text{beta})$ è pari a 1,002, se il numero di vicini aumentasse di uno, il rischio aumenterebbe dello 0,20%. Alla stessa maniera, se aumentasse di dieci, ad esempio, il rischio aumenterebbe del 2% ($HR = e^{(0,002*10)} = 1,020$).

Guardando al secondo risultato si può quindi affermare che, per il generico negozio in Torino, il numero di vicini della stessa categoria rappresenta il *terzo driver negativo* della probabilità di sopravvivenza.

- **RQ6:** *“il fatto di avere attorno al negozio i (della categoria X) tanti Retailer della qualsiasi categoria ha impatto sulla probabilità di sopravvivere dello stesso?”*

A questa domanda, come descritto al paragrafo 6.1.5, si può rispondere guardando all’impatto dell’indice A_2 e del parametro N_{vicini} : entrambi i fattori, come mostrato dalle tabelle, non hanno valori statisticamente significativi, pertanto non risultano predittivi della probabilità di sopravvivere del generico retailer in Torino.

6.2.3 COX REGRESSION – DRIVERS PER LE DIVERSE CATEGORIE MERCEOLOGICHE

Data l’importanza della categoria merceologica all’interno dell’analisi, si è deciso di studiare l’impatto dei fattori individuati per ognuna di esse.

Lo scopo di questa tipologia di studio è quello di comprendere se il fatto di essere un retailer di una specifica categoria merceologica porta, in base al fattore in analisi, a dei fenomeni di sopravvivenza diversi da quelli del generico negozio in Torino.

Si riportano a titolo esemplificativo delle categorie merceologiche specifiche, ma si rimanda all’Appendice D per la visualizzazione degli output delle restanti.

6.2.3.1 Metri Quadri

Relativamente alla **RQ3 – impatto del Formato** – si è scelto di mostrare, oltre alle categorie a cui si è fatto riferimento per l’intera analisi, Abbigliamento e Supermercati, anche l’output delle Farmacie. Le tre categorie mostrano infatti dei fenomeni differenti, visibili nelle tabelle sottostanti, rispettivamente 6, 7 e 8.

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	78,009	1	0,000	0,999
Abbigliamento	0,992	0,143	48,335	1	0,000	2,697
Abbigliamento*Mq	-0,001	0,000	3,344	1	0,000	0,999

TABELLA 6 - METRI QUADRI ABBIGLIAMENTO

L'output in Tabella 6 mostra l'impatto del formato sul generico negozio, alla riga Mq, e quello relativo all'interazione di due variabili: Abbigliamento e Mq. La variabile di interazione, data dal prodotto di una variabile binaria e di una continua, verrà "attivata" ogni qual volta il negozio in questione è proprio uno di Abbigliamento. L'inserimento di una variabile di interazione all'interno dell'analisi, permette di comprendere se "il fatto di essere un negozio di abbigliamento", porti ad un impatto differente da quello del negozio generico; in particolare possono verificarsi situazioni in cui:

- **il beta della variabile di interazione ha lo stesso segno del beta calcolato per il negozio medio ed è significativo:** in questa circostanza l'interpretazione da attribuire è che l'impatto viene amplificato, ovvero è più forte di quello che si verificherebbe se il negozio non fosse di quella categoria, ma con lo stesso segno;
- **il beta della variabile di interazione ha un valore non significativo:** in questo caso l'impatto della variabile di interazione non è da tenere in considerazione; si tratta infatti di un valore non statisticamente significativo. Il valore associato alla categoria in questione sarà, pertanto, quello del negozio medio.
- **il beta della variabile di interazione ha segno opposto a quello del beta calcolato per il negozio medio ed è significativo:** in questo caso l'interpretazione da attribuire è che la categoria porta ad un impatto opposto rispetto a quello che avviene per il generico negozio; può anche succedere che il segno opposto porti ad un effetto nullo, nel caso in cui i

due valori siano identici. L'effetto nullo è da leggere come "la variabile in questione non ha un impatto sulla probabilità di sopravvivenza del negozio".

Come si può vedere dalla Tabella 6, la categoria Abbigliamento rientra nel primo caso: il beta della variabile di interazione ha, infatti, lo stesso segno di quello del generico negozio (negativo). Questo vuol dire che se per il generico negozio in Torino "aumentare la dimensione dello stesso porta ad una riduzione del rischio di morte", per il generico negozio di Abbigliamento questo effetto è ancora maggiore. Osservando nello specifico gli effetti si ha che:

- per il generico negozio in Torino, ad una variazione di 1 della variabile metri quadri corrisponde una **riduzione** del rischio di morte dello 0,1%;
- per il generico negozio di Abbigliamento in Torino, ad una variazione di 1 della variabile metri quadri corrisponde una **riduzione** del rischio di morte dello 0,2%; calcolato come:

$$HR = e^{(-0,001*1-0,001*1)} = e^{(-0,002)} = 0,998$$

$$1 - HR = 0,002$$

È più importante, al fine di ridurre il rischio di morte, aumentare la dimensione in un negozio di Abbigliamento, che in un generico negozio non di Abbigliamento in Torino.

Nella seconda situazione rientrano, invece, i Supermercati, come viene mostrato in Tabella 7.

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	62,807	1	0,000	0,999
Supermercati	-1,174	0,381	9,465	1	0,002	0,309
Mq*Supermercati	0,001	0,000	1,421	1	0,233	1,001

TABELLA 7 - METRI QUADRI SUPERMERCATI

La variabile di interazione Mq*Supermercati ha un effetto non significativo a livello statistico, pertanto, non si può affermare che “il fatto di essere un Supermercato porti ad un impatto aggiuntivo rispetto a quello del generico negozio in Torino”.

Differente è la situazione che si verifica per le Farmacie e le Parafarmacie, la cui rappresentazione è visibile in Tabella 8.

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	105,802	1	0,000	0,999
Farmacia e Parafarmacia	-1,851	0,167	122,755	1	0,000	0,157
Farmacia e Parafarmacia*Mq	0,002	0,001	13,260	1	0,000	1,002

TABELLA 8 - METRI QUADRI FARMACIE E PARAFARMACIE

In questa circostanza, ci si trova nella terza casistica possibile; il beta della variabile di interazione ha, infatti segno opposto (positivo) a quello della variabile generica e la somma dei due valori è diversa da zero. L'effetto della dimensione, per le Farmacie e Parafarmacie assume un'interpretazione differente: “all'aumentare della dimensione del negozio il rischio di morte aumenta!”. A livello quantitativo, si avrà che:

- per il generico negozio in Torino, ad una variazione di 1 della variabile metri quadri corrisponde una **riduzione** del rischio di morte dello 0,1%;
- per la generica Farmacia o Parafarmacia in Torino, ad una variazione di 1 della variabile metri quadri, corrisponde un **aumento** del rischio di morte dello 0,1%; calcolato come:

$$HR = e^{(-0,001*1+0,002*1)} = e^{(0,001)} = 1,001$$

$$HR - 1 = 0,001$$

Come detto precedentemente, si rimanda all'Appendice D per la visualizzazione del comportamento di tutte le restanti categorie merceologiche; per visualizzare un quadro riassuntivo, è stata creata la tabella sottostante (Figura 22), in cui si riportano le venti¹⁸ categorie merceologiche e i loro rispettivi comportamenti, rispetto alla variabile dimensione.

MQ	
Tabacchi	
Farmacie	↑
Quotidiani	↑
Supermercati	
Art. Funerari	
Librerie	
Profumerie	
Cartolerie	↑
Calzature	
Elettronica	
Art. Casa	
Abbigliamento	↓
Alimentari	
Art. Animali	↓
Fiori e Piante	
Minimercati	↓
Mobili	
Ottica	
Telefonia	
Giocattoli	↓

FIGURA 22 - DRIVER DIMENSIONE PER TUTTE LE CATEGORIE

In Figura 22 vengono lasciate le caselle vuote per tutte le rispettive categorie merceologiche per cui il fattore dimensione non assume valore statisticamente significativo; le caselle che presentano una freccia verde verso il basso sono

¹⁸ Si ricorda, che la categoria Ipermercati è stata esclusa dall'analisi, per questioni di fattibilità (tutti vivi, impossibile studiarne la morte)

relative a quelle categorie merceologiche per cui la dimensione rappresenta un *driver positivo e statisticamente significativo* della sopravvivenza: è il caso di retailers di Abbigliamento, Articoli per Animali, Minimercati e Giocattoli, per i quali all'aumentare della loro superficie, il rischio di morte diminuisce. Contrariamente a ciò, le caselle che presentano una freccia rossa che punta verso l'alto si riferiscono a quelle categorie merceologiche per cui la dimensione rappresenta un *driver negativo e statisticamente significativo* della sopravvivenza: è il caso delle Farmacie, dei negozi di Quotidiani e delle Cartolerie, per i quali all'aumentare della loro superficie, il rischio di morte aumenta.

6.2.3.2 Distanza dal Centro

Per quanto riguarda la **RQ4 – impatto della distanza dal centro** – è stata utilizzata una metodologia d'analisi equivalente a quella presentata per il fattore dimensione. Per ogni categoria, quindi, sono state inserite delle variabili di interazione tra la categoria stessa e il fattore di interesse, ovvero la distanza; l'obiettivo di uno studio di questo tipo è quello di scoprire se la specifica categoria merceologica porta a dei fenomeni di sopravvivenza differenti da quelli individuati per il generico negozio in Torino. Nello specifico, come mostrato precedentemente nella Tabella 4 e 5, la distanza dal centro non ha un ruolo statisticamente significativo ai fini della sopravvivenza per il generico negozio.

Focalizzandosi, invece, sulle singole categorie merceologiche, si riportano tre categorie che presentano delle dinamiche, relativamente alla distanza dal centro, differenti: le prime due corrispondono a quelle presentate a titolo esemplificativo durante l'intera analisi, ovvero Abbigliamento e Supermercati (in Tabella 8 e 9); a queste si aggiunge la categoria Profumeria e Cosmetica (in Tabella 10).

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	77,906	1	0,000	0,999
Distanza	0,013	0,008	2,392	1	0,122	1,013
Abbigliamento	0,468	0,057	68,315	1	0,000	1,598
Abbigliamento*Mq	-0,001	0,000	5,959	1	0,005	0,999
Abbigliamento*Distanza	0,041	0,016	6,510	1	0,006	1,042

TABELLA 9 - DISTANZA DAL CENTRO ABBIGLIAMENTO

Dalla Tabella 9, si può notare come il fattore Distanza dal centro (abbreviato in Distanza) abbia un valore, coerentemente con quanto affermato in precedenza, non significativo in generale; la variabile di interazione, tra abbigliamento e distanza, assume però valore significativo. Una lettura che si può dare ad un fenomeno di questo tipo è che: “in generale, la distanza dal centro non è significativamente rilevante ai fini della sopravvivenza, ma se il negozio in questione è di Abbigliamento, essere vicino o meno al centro diventa importante a livello statistico”. Guardando agli effetti dal punto di vista quali e quantitativo, si ha che essendo l'esponentiale di beta maggiore di uno, il rischio di morte aumenta all'aumentare della variabile “distanza dal centro”¹⁹. In particolare l'exp(beta) assume valore pari a **1,042**; questo implica che:

- Se la variabile distanza subisse una variazione di 1 km, l'hazard ratio corrispondente sarebbe pari a 1,042 e, pertanto, il rischio di morte per il generico negozio di Abbigliamento aumenterebbe del **4,2%**;
- Se la variabile distanza subisse una variazione di 2 km, l'hazard ratio corrispondente sarebbe pari a: $HR = e^{(0,041*2)} = e^{(0,082)} = 1,085$; il rischio di morte per il generico negozio di Abbigliamento, pertanto, aumenterebbe del **8,5%**.

¹⁹ variabile misurata in chilometri, che assume valori compresi tra 0 e 8,5 km

In conclusione, quindi, per i negozi di Abbigliamento, essere localizzati o meno, in corrispondenza del centro cittadino Torinese, o vicini ad esso, è rilevante, infatti allontanandosi dal centro il rischio di morte aumenta.

Differente è la situazione che si verifica per i Supermercati, riportati in Tabella 10, per i quali la variabile distanza dal centro non risulta essere significativa a livello statistico. Questo implica che, essere vicini o meno al centro non è un fattore che impatta la sopravvivenza del generico supermercato in Torino.

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	62,519	1	0,000	0,999
Distanza	0,004	0,007	0,391	1	0,532	1,004
Supermercati	-1,132	0,548	4,265	1	0,039	0,323
Mq*Supermercati	0,001	0,000	1,402	1	0,236	1,001
Distanza*Supermercati	-0,014	0,129	0,012	1	0,914	0,986

TABELLA 10 - DISTANZA DAL CENTRO SUPERMERCATI

Situazione ancora differente, è quella rilevata per i negozi di Profumeria e Cosmetica, in Tabella 11: la distanza dal centro in questo caso risulta essere una variabile significativa ai fini della sopravvivenza, ma in direzione opposta rispetto a quella enunciata per i retailers di Abbigliamento. Nella fattispecie, infatti, l'esponentiale di beta presenta un valore minore di uno, **0,931**, il che implica che all'aumentare della distanza dal centro il rischio di morte per la generica Profumeria diminuisce. Dal punto di vista degli effetti quantitativi, si ha che:

- Se la variabile distanza subisse una variazione di **1 km**, l'hazard ratio corrispondente sarebbe pari a 0,931 e, pertanto, il rischio di morte per il generico negozio di Profumeria e Cosmetica diminuirebbe del **6,9%**;
- Se la variabile distanza subisse una variazione di **2 km**, l'hazard ratio corrispondente sarebbe pari a: $HR = e^{(-0,071*2)} = e^{(-0,142)} = 0,868$; il

rischio di morte per il generico negozio di Profumeria e Cosmetica, pertanto, sarebbe ridotto del **13,24%**.

Da quello che emerge dall'analisi, quindi, in generale per i negozi di Profumeria e Cosmetica è maggiormente conveniente, ai fini della sopravvivenza, localizzarsi distanti dal centro.

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	110,118	1	0,000	0,999
Distanza	0,007	0,007	1,014	1	0,314	1,007
Profumeria	-0,075	0,128	0,347	1	0,556	0,927
Mq*Profumeria	0,000	0,001	0,064	1	0,800	1,000
Distanza*Profumeria	-0,071	0,034	4,274	1	0,003	0,931

TABELLA 11 - DISTANZA DAL CENTRO PROFUMERIA E COSMETICA

Sono state analizzate tutte le categorie merceologiche e si rimanda all'appendice D per la visualizzazione dei corrispondenti output. Esattamente come il driver dimensione, si riporta una tabella riassuntiva (Figura 23) che indica il modo in cui varia il rischio di morte all'aumentare della variabile "distanza dal centro"; in particolare le caselle in bianco indicano un effetto non significativo, mentre le frecce verso l'alto o il basso corrispondono ai casi di aumento o diminuzione del rischio di morte.

La distanza dal centro rappresenta un *driver positivo e statisticamente significativo* solo nel caso dei negozi di Profumeria e Cosmetica; viceversa, come mostra la figura, nel caso dei retailers di Calzature, Abbigliamento, Alimentati, Minimercati e Mobili la distanza dal centro è un *driver negativo e statisticamente significativo* della probabilità di sopravvivere.

	MQ	CENTRO
Tabacchi		
Farmacie	↑	
Quotidiani	↑	
Supermercati		
Art. Funerari		
Librerie		
Profumerie		↓
Cartolerie	↑	
Calzature		↑
Elettronica		
Art. Casa		
Abbigliamento	↓	↑
Alimentari		↑
Art. Animali	↓	
Fiori e Piante		
Minimercati	↓	↑
Mobili		↑
Ottica		
Telefonia		
Giocattoli	↓	

FIGURA 23 - DRIVER DISTANZA DAL CENTRO PER TUTTE LE CATEGORIE

6.2.3.3 Agglomerazione Intra-Categoria e Inter-Categoria

Come enunciato al paragrafo 6.1.5, per procedere con l'analisi si è pensato di visualizzare sulla mappa di Torino una fotografia attuale del posizionamento dei retailers per le singole categorie merceologiche e ci si è resi conto che, incidentalmente, alcune categorie per cui la distanza dal centro risulta un driver significativo della probabilità di sopravvivenza, come Abbigliamento, risultano anche essere più agglomerate. A questo punto dello studio è, quindi, sorto un dubbio: il centro cittadino risulta essere un polo agglomerativo di negozi, infatti, come visto nella descrizione delle circoscrizioni, quella relativa al centro di Torino (circoscrizione 1), è la più piccola, in termini di dimensioni, ma ospita il maggior numero di negozi (il 23% degli spazi commerciali presenti nel database), che necessariamente saranno localizzati gli uni vicini agli altri. Può accadere che, per alcune categorie sia l'agglomerazione, quindi il fatto di stare vicino ad altri (della stessa categoria o meno), il driver significativo della sopravvivenza e che

questo fenomeno sia stato “inglobato” all’interno della variabile “distanza dal centro”, in quanto il centro è proprio il maggior sinonimo di agglomerazione. Occorre, quindi, chiedersi: è l’agglomerazione, e non il fatto di essere in centro o meno, il reale driver che guida i fenomeni di sopravvivenza?

Per rispondere alla **RQ5 – impatto dell’agglomerazione intra-categoria** – e alla **RQ6 – impatto dell’agglomerazione inter-categoria** - si presentano i risultati ottenuti per le categorie merceologiche Abbigliamento, Supermercati e Calzature (in Tabella 12, 13 e 14). Si rimanda all’appendice E per i risultati di tutte le altre categorie merceologiche.

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	78,009	1	0,000	0,999
Distanza	0,022	0,012	3,199	1	0,074	1,022
A1	-0,007	0,008	0,684	1	0,408	0,993
A2	0,007	0,006	1,057	1	0,304	1,007
Abbigliamento	0,992	0,143	48,335	1	0,000	2,697
Abbigliamento*Mq	-0,001	0,000	3,344	1	0,000	0,999
Abbigliamento*Distanza	-0,029	0,026	1,309	1	0,253	0,971
A1*Abbigliamento	-0,306	0,088	11,962	1	0,001	0,737
A2*Abbigliamento	-0,001	0,013	0,005	1	0,946	0,999

TABELLA 12 - AGGLOMERAZIONE ABBIGLIAMENTO

Quello che accade per la categoria Abbigliamento, presentata nella Tabella 12, è esattamente quello che si era previsto: introducendo all’interno del modello le variabili di interazione relative all’agglomerazione²⁰, la variabile “distanza dal centro” perde la significatività statistica, dimostrando di non essere il reale driver della sopravvivenza. Per quanto riguarda le variabili di interazione relative all’agglomerazione:

²⁰ si ricorda che l’indice A1 fa riferimento all’agglomerazione intra-categoria, mentre l’indice A2 a quella inter-categoria

- L'agglomerazione intra-categoria (A1*Abbigliamento) risulta significativa a livello statistico, presentando un esponenziale di beta minore di 1, che implica una diminuzione del rischio di morte all'aumentare della variabile A1. Nello specifico, il valore dell'hazard ratio è pari a **0,737**; pertanto, ad una variazione unitaria della variabile A1 corrisponde una riduzione del rischio di morte del **26,3%**. Per il generico negozio di Abbigliamento in Torino risulta fondamentale, ai fini della sopravvivenza, essere agglomerato, quindi avere attorno a sè altri negozi di Abbigliamento.
- L'agglomerazione inter-categoria (A2*Abbigliamento) non risulta essere una variabile significativa a livello statistico e, pertanto, il fatto di avere negozi vicini in generale non risulta essere un fattore fondamentale ai fini della sopravvivenza per il generico negozio in Torino.

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	62,747	1	0,000	0,999
A1	-0,010	0,009	1,167	1	0,280	0,990
A2	0,006	0,005	1,348	1	0,246	1,006
Distanza	0,013	0,010	1,591	1	0,207	1,013
Supermercati	-1,548	1,030	2,258	1	0,133	0,213
Mq*Supermercati	0,001	0,000	1,566	1	0,211	1,001
Distanza*Supermercati	0,063	0,175	0,128	1	0,720	1,065
A1*Supermercati	-0,074	0,234	0,100	1	0,752	0,929
A2*Supermercati	0,054	0,099	0,305	1	0,581	1,056

TABELLA 13 - AGGLOMERAZIONE SUPERMERCATI

Quello che avviene per i Supermercati è presentato nella Tabella 13. Inserendo le variabili di interazione relative all'agglomerazione (intra e inter categoria), che risultano essere non significative, la distanza continua ad essere anch'essa non significativa. Questo implica che, ai fini della sopravvivenza di un supermercato in Torino, non ha rilevanza né essere vicini al centro, né essere agglomerati ad altri supermercati, né avere attorno negozi in generale; è evidente come, per una

categoria di questo tipo, la sopravvivenza sia guidata da altri fattori che non vengono indagati in questo tipo di analisi.

Il caso dei retailers di Calzature, riportato in Tabella 14, è, in un certo senso, simile a quello descritto per i negozi di Abbigliamento. Come mostrato, sia in Figura 23, che in Appendice D, per il generico negozio di Calzature in Torino la “distanza dal centro” rappresenta un driver negativo della probabilità di sopravvivenza, ovvero all’aumentare della distanza il rischio di morte aumenta a sua volta. Introducendo all’interno del modello le variabili di interazione relative all’agglomerazione, la variabile corrispondente alla distanza perde di significatività statistica, come nel caso di Abbigliamento. Anche per la categoria Calzature si può affermare, quindi, che la distanza dal centro non sia il reale driver della sopravvivenza, ma che risultava essere significativo perché stava “inglobando” gli effetti reali, dovuti all’agglomerazione. Per quanto riguarda le variabili di interazione relative all’agglomerazione:

- L’agglomerazione intra-categoria ($A1*Calzature$) non risulta essere una variabile significativa a livello statistico e, pertanto, il fatto di avere degli altri negozi di Calzature attorno al generico negozio di Calzature non risulta essere un fattore fondamentale ai fini della sopravvivenza.
- L’agglomerazione inter-categoria ($A2*Calzature$) è invece una variabile statisticamente significativa. Essa presenta un esponenziale di beta minore di uno, il che implica una diminuzione del rischio di morte all’aumentare della variabile A2. In particolare, il valore dell’hazard ratio è pari a **0,918**, e quindi ad una variazione unitaria della variabile A2, corrisponde una riduzione del rischio di morte dell’**8,2%**. È fondamentale, quindi, per una categoria merceologica come Calzature, avere accanto dei vicini, che generano traffico, per sopravvivere.

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	107,049	1	0,000	0,999
Distanza	0,014	0,011	1,765	1	0,184	1,014
A1	-0,009	0,009	1,054	1	0,305	0,991
A2	0,008	0,005	2,589	1	0,108	1,008
Calzature	0,959	0,427	5,038	1	0,025	2,609
Calzature*Mq	-0,001	0,001	0,459	1	0,498	0,999
Calzature*Distanza	-0,014	0,087	0,028	1	0,868	0,986
A1*Calzature	-0,043	0,093	0,213	1	0,645	0,958
A2*Calzature	-0,086	0,039	4,834	1	0,005	0,918

TABELLA 14 - AGGLOMERAZIONE CALZATURE

Viene mostrata, adesso, un quadro riassuntivo degli effetti su tutte le categorie merceologiche dei quattro driver presentati finora: dimensione, distanza dal centro, agglomerazione intra-categoria e agglomerazione inter-categoria. Dando uno sguardo al quadro presentato in Figura 24 si riesce facilmente ad evincere quali siano per ogni categoria i fattori che ne guidano la sopravvivenza.

	MQ	CENTRO	Agglomerazione INTRA-CATEGORIA	Agglomerazione INTER-CATEGORIA
Tabacchi				
Farmacie	↑			
Quotidiani	↑			
Supermercati				
Art. Funerari				
Librerie				↓
Profumerie		↓		
Cartolerie	↑			
Calzature				↓
Elettronica				
Art. Casa				↓
Abbigliamento	↓		↓	
Alimentari		↑		
Art. Animali	↓			
Fiori e Piante				
Minimercati	↓	↑		
Mobili		↑		
Ottica				
Telefonia				↓
Giocattoli	↓			

FIGURA 24 - DRIVERS PER TUTTE LE CATEGORIE MERCEOLOGICHE

7. CONCLUSIONI

L'obiettivo di questa trattazione è stato quello di dare una risposta alla domanda di fondo che il lavoro di tesi si è posto: "Quali sono i drivers della sopravvivenza di un Retailer?". Come si evince dall'analisi empirica, esiste una sommaria risposta a questa domanda, ottenuta mediante lo studio del generico negozio presente in Torino; allo stesso modo, ne esistono tante specifiche, ottenute contestualizzando di volta in volta rispetto alla categoria merceologica di interesse.

Occorre notare che, l'indagine effettuata si concentra solo sul sottoinsieme di variabili che risultano osservabili e misurabili²¹; ovviamente la sopravvivenza di un negozio dipenderà da un insieme di altri fattori non riscontrabili a priori, come la gestione del negozio, il prezzo servito ed il ciclo di vita stesso dell'imprenditore.

Come già precedentemente annunciato, comprendere le logiche e le dinamiche dei moderni retailers può rappresentare la chiave di volta nell'interpretazione e regolazione dei processi di sviluppo urbano, sociale ed economico, (M. Saraiva, P. Pinho, 2015). Un elaborato di questo tipo può risultare di particolare interesse per un *policy maker* che si trovi a dover ridefinire delle politiche pubbliche o a sviluppare nuove strategie di progresso cittadino, riconfigurando gli spazi commerciali all'interno della città; allo stesso tempo, può rappresentare una linea guida per un retailer che abbia l'interesse di avviare una nuova attività commerciale e si trovi ad analizzare il mercato per intraprendere delle decisioni sulla propria localizzazione.

²¹ variabili presenti nel dataset ottenuto dal sito del Comune di Torino o ricavabili da esse

Dall'analisi emergono delle evidenze empiriche, relative alla città di Torino, di particolare rilievo:

- Analizzando i retailers dal punto di vista della **categoria merceologica** si evincono differenti effetti riguardo il rischio di morte, in particolare, delle venti categorie analizzate, otto²² di esse riducono il rischio di morte, quattro ²³ lo aumentano e le restanti producono un effetto non significativamente differente da quello del negozio medio. Da uno studio incrociato sui risultati ottenuti dall'analisi dei flussi di entrata ed uscita dal mercato e quelli derivanti dalla Survival Analysis, quello che emerge è l'esistenza in Torino di due macro-gruppi di categorie merceologiche: quelle con un alto turnover, che presentano quindi congiuntamente un alto tasso di nascita e di morte, e quelle con un basso turnover, che viceversa mostrano un basso tasso di nascita e di morte. Rare sono, invece, le situazioni di turnover con valori intermedi, derivanti da casi di non nascita e morte e nascita e non morte. È fondamentale quindi, al fine di non derivare informazioni distorte, focalizzarsi su entrambi i tipi di analisi. Una delle categorie, ad esempio, che in maggior misura riduce il rischio di morte in Torino è **Edicole**²⁴. Focalizzarsi solo sull'analisi di sopravvivenza, e quindi sulla durata media di vita, potrebbe indurre a pensare che il settore dell'editoria sia attrattivo nella realtà torinese. Le edicole però, nonostante incorrano in un basso rischio di morte, sono caratterizzate anche da un basso tasso di entrata sul mercato. È interessante notare che, nonostante la crisi che l'editoria cartacea sta subendo in quest'ultimi anni con l'avvento di internet, le edicole in Torino

²² Tabacchi, Farmacia e Parafarmacia, Edicole, Supermercati, Articoli Funerari, Librerie, Profumeria e Cosmetica e Cartolerie

²³ Calzature, Elettronica, Articoli per la Casa e Abbigliamento

²⁴ l'essere un negozio di quotidiani riduce il rischio di morte del 77,9% rispetto al non esserlo (essere quindi un negozio di qualsiasi categoria merceologica eccetto che un'edicola)

non chiudano. Una spiegazione di questo fenomeno può risiedere nell'ampliamento della tipologia di beni venduti dagli edicolanti, così come può essere insita nella struttura stessa della maggior parte di edicole torinesi; si tratta infatti di chioschi che difficilmente possono essere adibiti ad una differente attività.

Un altro caso esemplificativo della realtà torinese è quello delle **Librerie**²⁵, anch'esse con un basso turnover ed un basso rischio di morte. L'analisi effettuata ha fatto emergere l'esistenza di un ampio numero di librerie storiche ²⁶ in Torino; esse vivono grazie alla loro forte identità e specializzazione su campi distintivi. Nello specifico, delle 150 librerie già attive sul territorio al momento iniziale dell'analisi (anno 2004) ad oggi ne sono ancora attive i due terzi, nonostante la crescita dell'e-commerce e la diffusione dei libri digitali;

- Dall'analisi di un'altra caratteristica fondamentale nel mondo del Retail, il **formato**, è emerso che essa rappresenta un driver positivo della sopravvivenza. Il generico negozio in Torino ha un rischio di morte via via decrescente al crescere della sua dimensione. Questo risultato conferma ciò che già in letteratura viene espresso dai modelli gravitazionali, (W. J. Reilly, 1931) e (D. L. Huff, 1963), che guardano alla dimensione come proxy dell'attrattività del negozio stesso: un negozio di maggiori dimensioni, infatti, offrirà una maggiore gamma e varietà di prodotti e sarà quindi maggiormente allettante agli occhi del consumatore. Un'ulteriore spiegazione al risultato trovato risiede nel fatto che, i negozi di grandi dimensioni nella maggior parte dei casi sono dei *Chain Stores* ed appartengono quindi ad un network; l'attrattività di questi

²⁵ l'essere una libreria riduce il rischio di morte del 26,6% rispetto al non esserlo (essere quindi un negozio di qualsiasi categoria merceologica eccetto che una libreria)

²⁶ ai fini dell'analisi sono state definite storiche tutte quelle librerie vive da oltre 10 anni

tipi di negozi, spesso ancora del commercio, è maggiore e questo fa sì che le logiche di sopravvivenza a cui sono sottoposti siano differenti da quelle dei *Mom-and-Pop Stores* (si rimanda al capitolo 3.2 per la differenza tra le due tipologie di negozio);

- Ulteriore evidenza empirica è quella relativa alla **localizzazione** nella quale il retailers sceglie ex-ante di posizionarsi. In letteratura esiste in merito il concetto di “eterogeneità spaziale”, secondo il quale con l’evoluzione della struttura urbana si è affermata sempre più la tendenza da parte delle attività economiche ad agglomerarsi tra loro e a posizionarsi in dei centri che ne favoriscano l’interazione, (G. Giuliano & K. A. Small, 1991). Nel caso della città di Torino il centro cittadino rappresenta il luogo in cui si verifica la maggiore agglomerazione di retailers; per questo motivo, si è deciso di analizzare il fenomeno studiando dapprima l’effetto della distanza dal centro e, successivamente, quello dell’agglomerazione. Si è seguito questo procedimento perché l’analisi ha fatto emergere che il centro risulta essere un fattore significativamente predittivo della probabilità di sopravvivere per tutte quelle categorie che vendono beni di tipo “shopping”, (Holton, 1958), ovvero prodotti per cui il vantaggio derivante dalla comparazione di prezzo e qualità delle alternative proposte è particolarmente elevato. Queste categorie, come Abbigliamento, Calzature, Articoli per la Casa, sopravvivono maggiormente proprio nel centro della città di Torino, dove si misura la maggiore agglomerazione tra i retailers; è l’agglomerazione, quindi, uno dei driver della loro sopravvivenza. Differente è, invece, l’evidenza riguardante i Minimercati, per i quali è proprio il fatto di essere vicini al centro, inteso come luogo nello spazio e non come sinonimo di agglomerazione, che consente la sopravvivenza. Questo risultato può essere spiegato guardando alla competizione tra i formati: nel centro della città di Torino i Minimercati sopravvivono principalmente, perché vi è

assenza di competizione tra i formati (Minimercati, Supermercati, Ipermercati).

8. BIBLIOGRAFIA

- A. Anas, R. Arnett, & K. A. Small. (1998). *Urban Spatial Structure*.
- A. Wolinsky. (1983). *Retail Trade Concentration due to Consumers' Imperfect Information*.
- Alcàcer. (2006).
- C. A. Hidalgo, & E. E. Castaner. (2015). *The Amenity Space and The Evolution of Neighborhoods*.
- Città di Torino. (2016). *Servizio Telematico Pubblico*. Tratto da Regolamento del decentramento: <http://www.comune.torino.it/regolamenti/374/374.htm>
- Commissione Europea. (s.d.). *Retail Market Monitoring Report*.
- D. L. Huff. (1963). *A Probabilistic Analysis of Shopping Center Trade Areas*.
- Deloitte. (2018). *Global Power of Retailing, 2018*.
- Duranton & Puga. (2004).
- E. Glaeser, G. Ellison, & W. R. Kerr. (2010). *What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns*.
- E. L. Glaeser, & J. D. Gottlieb. (2006). *Urban Resurgence and the Consumer City*.
- E. L. Gleaser, J. Kolko, & A. Saiz. (2001). *Consumer City*.
- Fox, et al. (2007). *The Impact of Retail Location on Retailer Revenues: An Empirical Investigation*.
- G. A. Picone, & D.B. Ridley. (2008). *Distance decreases with differentiation: Strategic agglomeration by retailers*.

- G. Ellison, & E. Glaeser. (1994/1997). Geographic Concentration in U.S. Manufacturing Industries: a Dartboard Approach.
- G. Giuliano, & K. A. Small. (1991). Subcenters in the Los Angeles Region.
- Holton. (1958).
- Hotelling. (1929).
- Jerath, K. (s.d.).
- LaFaive, M. D. (s.d.).
- M. Dudey. (1990).
- Marchall. (1890).
- McDonald. (1987).
- Milgrom, P., & Roberts, J. (1982). Limit pricing and entry under incomplete information: An equilibrium analysis. *Econometrica: Journal of the Econometric Society*, 443-459.
- Nelson. (1970).
- Osservatorio Regione Piemonte del Commercio. (2015). *Il Commercio in Piemonte 2015*.
- S. Angel, J. Parent, D.L. Civico, & A. M. Blei. (2010). The Persistent Decline in Urban Densities: Global and Historical Evidence of “Sprawl”.
- Sevtzuk. (2014). Location and Agglomeration: The Distribution of Retail and Food Businesses in Dense Urban Environments.
- W. Alonso. (1960). A Theory of Urban Land Market.

W. J. Reilly. (1931). The Law of Retail Gravitation.

9. APPENDICE A

Si riportano di seguito le mappe, ottenute tramite l'utilizzo del software QGis, del posizionamento delle singole categorie merceologiche all'interno della città di Torino; i pallini colorati rappresentano i retailersss attivi al 2017, per i quali, quindi, non si è mai verificata una chiusura.

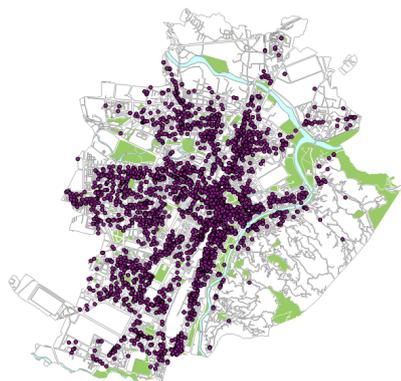


FIGURA 25 – ALIMENTARI



FIGURA 28 - ARTICOLI FUNERARI



FIGURA 26 - ARTICOLI ANIMALI



FIGURA 29 – CALZATURE

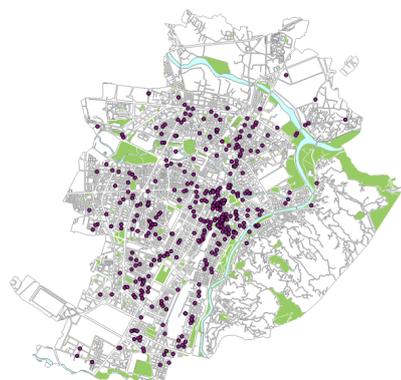


FIGURA 27 - ARTICOLI PER LA CASA



FIGURA 30 – CARTOLERIE



FIGURA 31 – ELETTRONICA

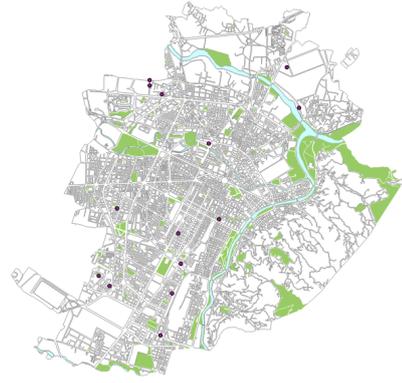


FIGURA 35 – IPERMERCATI



FIGURA 32 – FARMACIE



FIGURA 36 – LIBRERIE



FIGURA 33 - FIORI E PIANTE

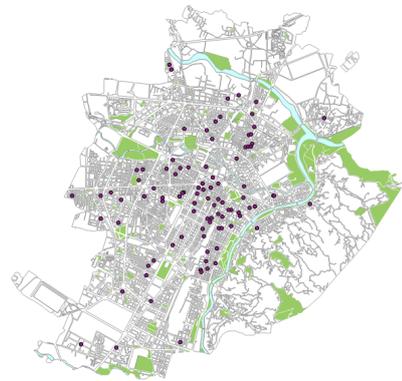


FIGURA 37 – MINIMERCATI



FIGURA 34 – GIOCATTOLI



FIGURA 38 – MOBILI



FIGURA 39 – OTTICA



FIGURA 42- TABACCHI

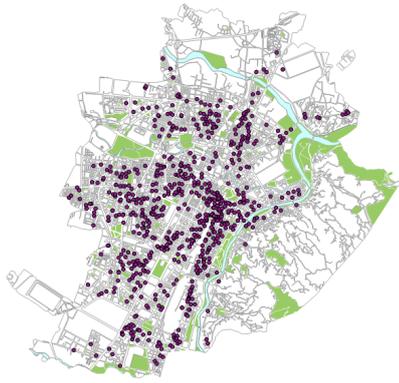


FIGURA 40 - PROFUMERIA E COSMETICA

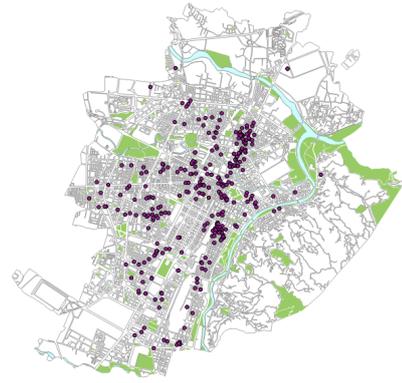


FIGURA 43 – TELEFONIA



FIGURA 41 – QUOTIDIANI

10. APPENDICE B

Si riportano di seguito le curve di sopravvivenza delle 19 categorie merceologiche non descritte durante l'analisi.

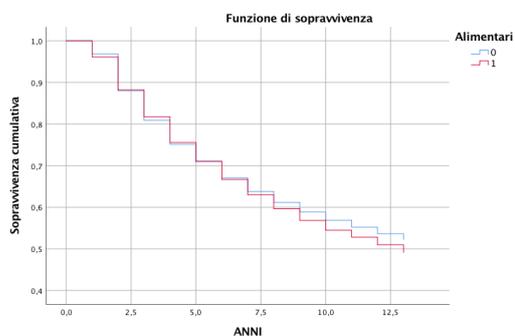


FIGURA 44 - ALIMENTARI

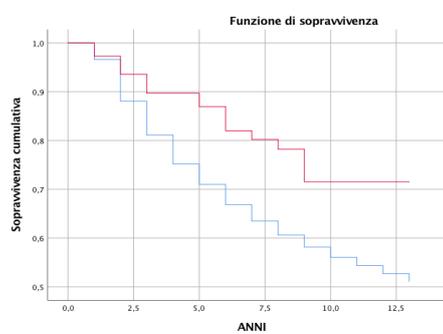


FIGURA 47 - ARTICOLI FUNERARI

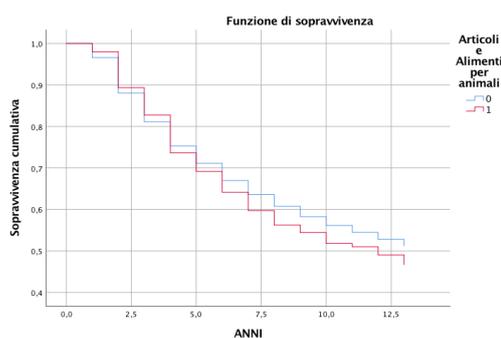


FIGURA 45 - ARTICOLI ANIMALI

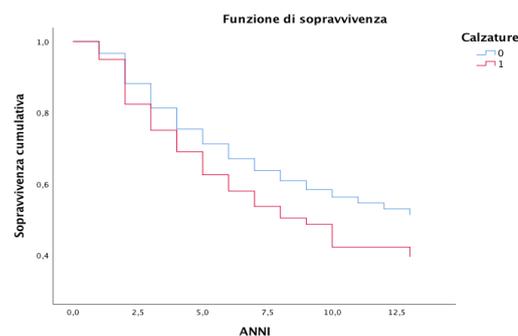


FIGURA 48 - CALZATURE

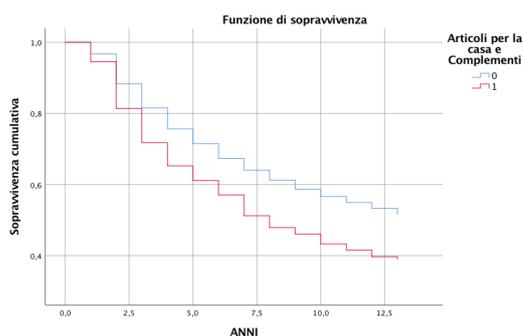


FIGURA 46 - ARTICOLI PER LA CASA

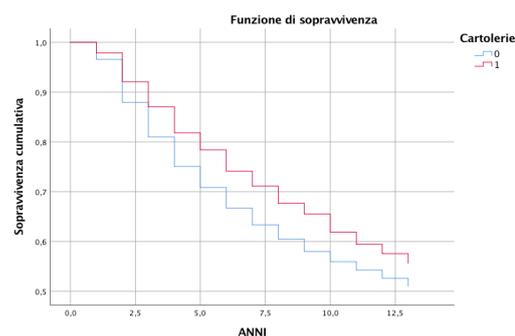


FIGURA 49 - CARTOLERIE

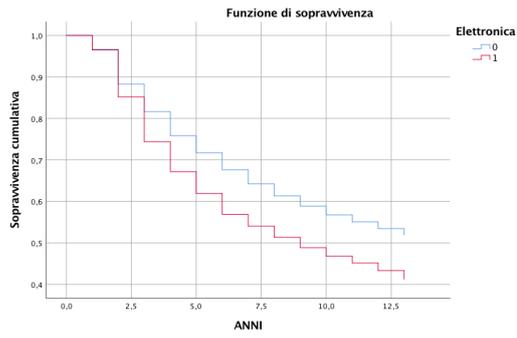


FIGURA 50 – ELETTRONICA

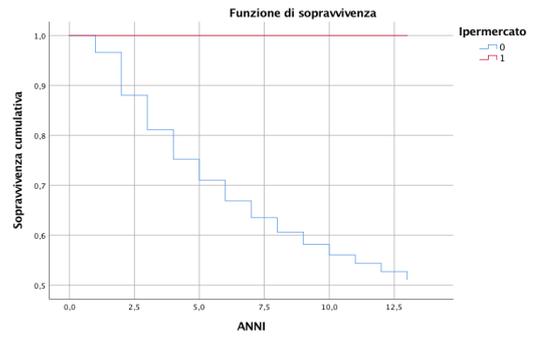


FIGURA 54 – IPERMERCATI

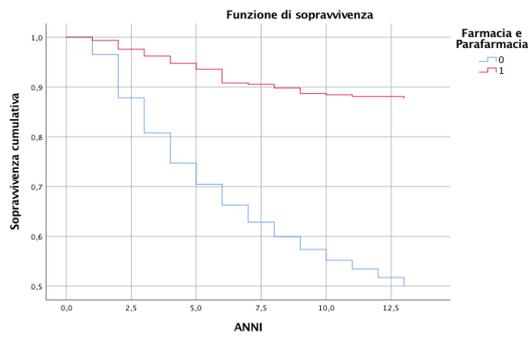


FIGURA 51 - FARMACIA E PARAFARMACIA

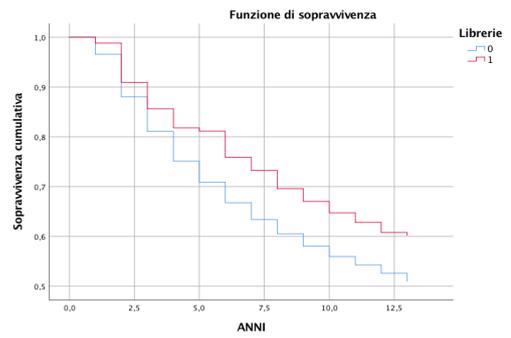


FIGURA 55 – LIBRERIE

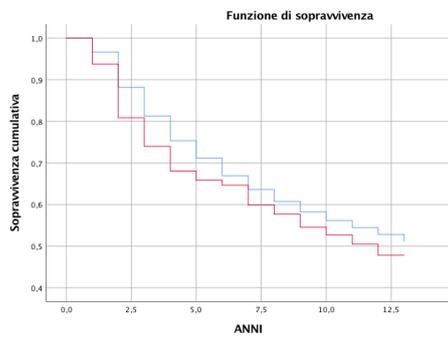


FIGURA 52 - FIORI E PIANTE

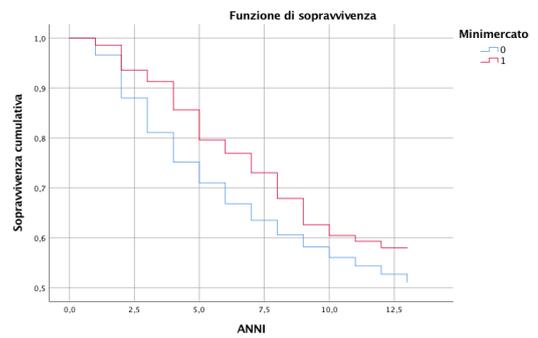


FIGURA 56 - MINIMERCATI

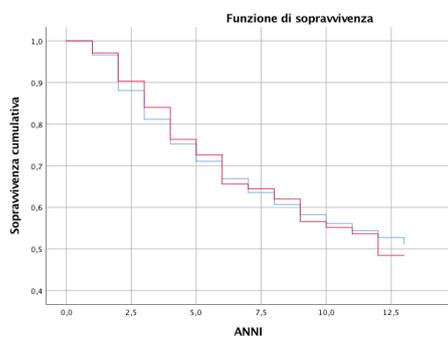


FIGURA 53 – GIOCATTOLI

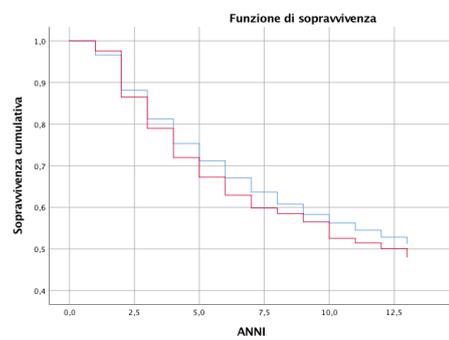


FIGURA 57 – MOBILI

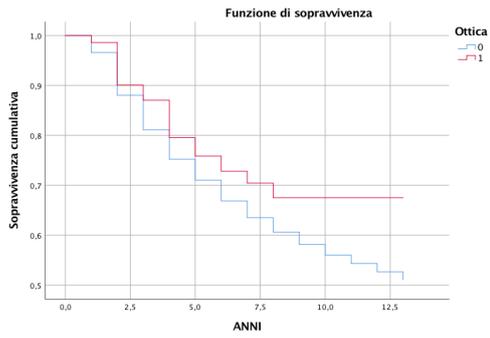


FIGURA 58 – OTTICA

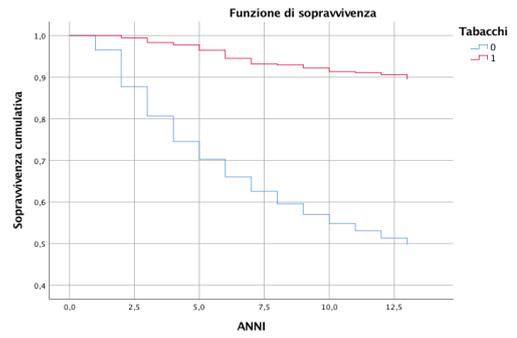


FIGURA 61 – TABACCHI

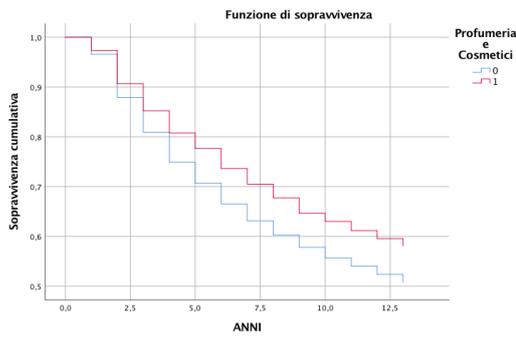


FIGURA 59 - PROFUMERIA E COSMETICA

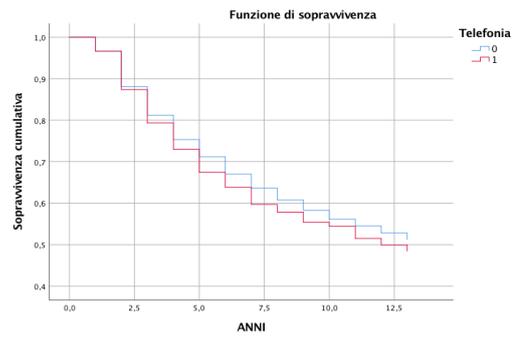


FIGURA 62 – TELEFONATA

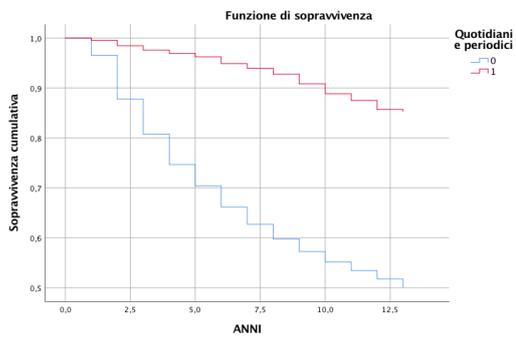


FIGURA 60 - QUOTIDIANI E PERIODICI

11. APPENDICE C

Di seguito sono presentati gli output della sopravvivenza delle categorie merceologiche; si riportano sia i risultati significativi che quelli che non lo sono.

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Alimentari	0,037	0,025	2,291	1	0,130	1,038

TABELLA 15 – ALIMENTARI

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Articoli Animali	0,038	0,090	0,175	1	0,676	1,038

TABELLA 16 - ARTICOLI ANIMALI

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Articoli per la Casa	0,364	0,052	48,461	1	0,000	1,440

TABELLA 17 - ARTICOLI CASA

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Articoli Funerari	-0,714	0,219	10,680	1	0,001	0,490

TABELLA 18 - ARTICOLI FUNERARI

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Calzature	0,286	0,086	11,101	1	0,001	1,331

TABELLA 19 – CALZATURE

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Cartolerie	-0,228	0,072	9,914	1	0,002	0,796

TABELLA 20 – CARTOLERIE

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Elettronica	0,297	0,042	49,613	1	0,000	1,346

TABELLA 21 – ELETTRONICA

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Farmacia	-1,609	0,142	128,452	1	0,000	0,200

TABELLA 22 – FARMACIA

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Fiori e Piante	0,151	0,102	2,210	1	0,137	1,163

TABELLA 23 - FIORI E PIANTE

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Giocattoli	-0,007	0,138	0,002	1	0,961	0,993

TABELLA 24 – GIOCATTOLI

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Librerie	-0,309	0,095	10,497	1	0,001	0,734

TABELLA 25 – LIBRERIE

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Minimercati	-0,284	0,146	3,769	1	0,052	0,753

TABELLA 26 – MINIMERCATI

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mobili	0,085	0,065	1,674	1	0,196	1,088

TABELLA 27 – MOBILI

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Ottica	-0,359	0,163	4,878	1	0,027	0,698

TABELLA 28 – OTTICA

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Profumeria	-0,265	0,055	23,188	1	0,000	0,767

TABELLA 29 – PROFUMERIA

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Quotidiani	-1,508	0,130	135,236	1	0,000	0,221

TABELLA 30 - QUOTIDIANI

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Tabacchi	-1,843	0,139	175,143	1	0,000	0,158

TABELLA 31 – TABACCHI

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Telefonia	0,061	0,071	0,720	1	0,396	1,062

TABELLA 32 – TELEFONIA

12. APPENDICE D

Vengono di seguito riportati gli output relativi all'impatto dei fattori **dimensione e distanza dal centro** sulla probabilità di sopravvivenza; si fa riferimento a tutte le categorie merceologiche, al fine di scoprire se "l'essere un negozio di una determinata categoria merceologica" porta a dei fenomeni di sopravvivenza differenti dalla media. Si badi al fatto che vengono riportate sia le categorie per cui l'effetto risulta significativo a livello statistico, che quelle in cui non vi è alcun impatto significativo (in questo caso la significatività verrà evidenziata in rosso).

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	92,977	1	0,000	0,999
Distanza	-0,009	0,009	0,954	1	0,329	0,992
Alimentari	-0,112	0,056	3,966	1	0,046	0,894
Alimentari*Mq	0,000	0,000	0,023	1	0,879	1,000
Alimentari*Distanza	0,035	0,015	5,606	1	0,010	1,036

TABELLA 33 - MQ E DISTANZA ALIMENTARI

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	102,639	1	0,000	0,999
Distanza	0,003	0,007	0,160	1	0,689	1,003
ArticoliAnimali	0,308	0,255	1,458	1	0,227	1,360
ArticoliAnimali*Mq	-0,008	0,002	13,346	1	0,000	0,992
ArticoliAnimali*Distanza	0,073	0,062	1,375	1	0,241	1,075

TABELLA 34 - MQ E DISTANZA ARTICOLI ANIMALI

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	103,096	1	0,000	0,999
Distanza	0,003	0,007	0,167	1	0,682	1,003
Articoli Casa	0,371	0,110	11,330	1	0,001	1,449
Articoli Casa*Mq	-0,001	0,001	0,836	1	0,361	0,999
Articoli Casa*Distanza	0,023	0,030	0,592	1	0,442	1,023

TABELLA 35 - MQ E DISTANZA ARTICOLI PER LA CASA

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	109,683	1	0,000	0,999
Distanza	0,005	0,007	0,429	1	0,512	1,005
ArticoliFunerari	-0,132	0,682	0,037	1	0,847	0,877
ArticoliFunerari*Mq	-0,011	0,010	1,097	1	0,295	0,989
ArticoliFunerari*Distanza	-0,044	0,145	0,094	1	0,759	0,957

TABELLA 36 - MQ E DISTANZA ARTICOLI FUNERARI

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	106,545	1	0,000	0,999
Distanza	0,002	0,007	0,051	1	0,821	1,002
Calzature	0,017	0,187	0,009	1	0,926	1,018
Calzature*Mq	-0,001	0,001	1,013	1	0,314	0,999
Calzature*Distanza	0,148	0,051	8,358	1	0,004	1,160

TABELLA 37 - MQ E DISTANZA CALZATURE

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	111,924	1	0,000	0,999
Distanza	0,003	0,007	0,141	1	0,708	1,003
Cartolerie	-0,485	0,164	8,792	1	0,003	0,616
Cartolerie*Mq	0,002	0,001	2,924	1	0,010	1,002
Cartolerie*Distanza	0,041	0,044	0,863	1	0,353	1,041

TABELLA 38 - MQ E DISTANZA CARTOLERIE

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	97,957	1	0,000	0,999
Distanza	0,003	0,007	0,120	1	0,729	1,003
Elettronica	0,331	0,097	11,738	1	0,001	1,393
Elettronica*Mq	0,000	0,000	0,029	1	0,865	1,000
Distanza*Elettronica	-0,013	0,027	0,245	1	0,620	0,987

TABELLA 39 - MQ E DISTANZA ELETTRONICA

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	105,767	1	0,000	0,999
Distanza	0,006	0,007	0,823	1	0,364	1,006
Farmacia	-1,884	0,321	34,537	1	0,000	0,152
Farmacia*Mq	0,002	0,001	13,148	1	0,000	1,002
Farmacia*Distanza	0,010	0,087	0,013	1	0,908	1,010

TABELLA 40 - MQ E DISTANZA FARMACIA E PARAFARMACIA

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	108,063	1	0,000	0,999
Distanza	0,003	0,007	0,200	1	0,655	1,003
FioriePiante	0,102	0,254	0,160	1	0,689	1,107
FioriePiante*Mq	0,000	0,003	0,009	1	0,925	1,000
Distanza*FioriePiante	0,006	0,065	0,010	1	0,921	1,006

TABELLA 41- MQ E DISTANZA FIORI E PIANTE

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	105,680	1	0,000	0,999
Distanza	0,004	0,007	0,307	1	0,580	1,004
Giocattoli	0,488	0,313	2,424	1	0,120	1,629
Giocattoli*Mq	-0,002	0,002	1,769	1	0,184	0,998
Distanza*Giocattoli	-0,073	0,084	0,748	1	0,387	0,930

TABELLA 42 - MQ E DISTANZA GIOCATTOLI

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	107,711	1	0,000	0,999
Distanza	0,002	0,007	0,093	1	0,760	1,002
Librerie	-0,238	0,158	2,268	1	0,132	0,788
Librerie*Mq	0,001	0,001	2,520	1	0,112	1,001
Distanza*Librerie	-0,071	0,069	1,074	1	0,300	0,931

TABELLA 43 - MQ E DISTANZA LIBRERIE

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	99,817	1	0,000	0,999
Distanza	0,002	0,007	0,069	1	0,792	1,002
Minimercati	-0,124	0,334	0,139	1	0,710	0,883
Minimercati*Mq	-0,003	0,001	7,903	1	0,005	0,997
Distanza*Minimercati	0,261	0,079	10,860	1	0,001	1,299

TABELLA 44 - MQ E DISTANZA MINIMERCATI

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	105,961	1	0,000	0,999
Distanza	0,001	0,007	0,017	1	0,895	1,001
Mobili	-0,082	0,139	0,352	1	0,553	0,921
Mobili*Mq	0,000	0,000	0,463	1	0,496	1,000
Distanza*Mobili	0,086	0,040	4,725	1	0,015	1,090

TABELLA 45- MQ E DISTANZA MOBILI

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	108,864	1	0,000	0,999
Distanza	0,004	0,007	0,263	1	0,608	1,004
Ottica	-0,187	0,439	0,181	1	0,670	0,830
Mq*Ottica	0,000	0,004	0,001	1	0,970	1,000
Distanza*Ottica	-0,070	0,107	0,431	1	0,512	0,932

TABELLA 46 - MQ E DISTANZA OTTICA

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	110,118	1	0,000	0,999
Distanza	0,007	0,007	1,014	1	0,314	1,007
Profumeria	-0,075	0,128	0,347	1	0,556	0,927
Mq*Profumeria	0,000	0,001	0,064	1	0,800	1,000
Distanza*Profumeria	-0,071	0,034	4,274	1	0,003	0,931

TABELLA 47- MQ E DISTANZA PROFUMERIA E COSMETICA

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	126,786	1	0,000	0,999
Distanza	0,006	0,007	0,809	1	0,368	1,006
Quotidiani	-1,409	0,257	30,079	1	0,000	0,244
Mq*Quotidiani	0,001	0,001	2,396	1	0,006	1,001
Distanza*Quotidiani	-0,072	0,076	0,904	1	0,342	0,930

TABELLA 48- MQ E DISTANZA QUOTIDIANI

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	124,055	1	0,000	0,999
Distanza	0,006	0,007	0,668	1	0,414	1,006
Tabacchi	-2,192	0,339	41,734	1	0,000	0,112
Mq*Tabacchi	0,002	0,005	0,208	1	0,649	1,002
Distanza*Tabacchi	0,062	0,079	0,620	1	0,431	1,064

TABELLA 49 - MQ E DISTANZA TABACCHI

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	107,208	1	0,000	0,999
Distanza	0,003	0,007	0,152	1	0,697	1,003
Telefonia	0,053	0,182	0,085	1	0,771	1,054
Mq*Telefonia	-0,003	0,002	1,264	1	0,261	0,997
Distanza*Telefonia	0,032	0,050	0,398	1	0,528	1,032

TABELLA 50 - MQ E DISTANZA TELEFONIA

13. APPENDICE E

Vengono di seguito riportati gli output relativi all'impatto dei quattro drivers individuati (dimensione, distanza dal centro, agglomerazione intra e inter-categoria) sulla probabilità di sopravvivenza; si fa riferimento a tutte le categorie merceologiche, al fine di scoprire se "l'essere un negozio di una determinata categoria merceologica" porta a dei fenomeni di sopravvivenza differenti dalla media. Si badi al fatto che vengono riportate sia le categorie per cui l'effetto risulta significativo a livello statistico, che quelle in cui non vi è alcun impatto significativo (in questo caso la significatività verrà evidenziata in grassetto).

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	92,789	1	0,000	0,999
Distanza	0,000	0,013	0,001	1	0,982	1,000
A1	-0,013	0,011	1,399	1	0,237	0,987
A2	0,006	0,006	0,825	1	0,364	1,006
Alimentari	-0,265	0,180	2,178	1	0,140	0,767
Alimentari*Mq	0,000	0,000	0,024	1	0,876	1,000
Alimentari*Distanza	0,030	0,022	1,887	1	0,010	1,031
A1*Alimentari	0,151	0,134	1,287	1	0,257	1,164
A2*Alimentari	0,003	0,012	0,063	1	0,803	1,003

TABELLA 51 - OUTPUT COMPLESSIVO ALIMENTARI

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	102,835	1	0,000	0,999
Distanza	0,013	0,011	1,595	1	0,207	1,013
A1	-0,011	0,010	1,243	1	0,265	0,989
A2	0,007	0,005	1,874	1	0,171	1,007
ArticoliAnimali	0,614	0,442	1,926	1	0,165	1,848
ArticoliAnimali*Mq	-0,008	0,002	13,326	1	0,000	0,992
ArticoliAnimali*Distanza	0,031	0,076	0,168	1	0,682	1,032
A1*ArticoliAnimali	-0,001	0,087	0,000	1	0,995	0,999
A2*ArticoliAnimali	-0,045	0,052	0,738	1	0,390	0,956

TABELLA 52 - OUTPUT COMPLESSIVO ARTICOLI ANIMALI

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	103,467	1	0,000	0,999
Distanza	0,014	0,011	1,804	1	0,179	1,015
A1	-0,009	0,009	1,019	1	0,313	0,991
A2	0,008	0,005	2,134	1	0,144	1,008
Articoli Casa	0,752	0,268	7,875	1	0,005	2,121
Articoli Casa*Mq	-0,001	0,001	0,782	1	0,376	0,999
Articoli Casa*Distanza	-0,041	0,049	0,717	1	0,397	0,960
A1*Articoli Casa	0,007	0,116	0,004	1	0,950	1,007
A2*Articoli Casa	-0,040	0,025	2,597	1	0,015	0,960

TABELLA 53 - OUTPUT COMPLESSIVO ARTICOLI CASA

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	110,515	1	0,000	0,999
Distanza	0,014	0,011	1,687	1	0,194	1,014
A1	0,009	0,016	0,284	1	0,594	1,009
A2	0,006	0,005	1,215	1	0,270	1,006
ArticoliFunerari	-1,519	1,342	1,280	1	0,258	0,219
ArticoliFunerari*Mq	-0,007	0,010	0,468	1	0,494	0,993
ArticoliFunerari*Distanz	0,127	0,197	0,416	1	0,519	1,135
A1*ArticoliFunerari	-0,026	0,035	0,540	1	0,462	0,974
A2*ArticoliFunerari	0,186	0,159	1,380	1	0,240	1,205

TABELLA 54 - OUTPUT COMPLESSIVO ARTICOLI FUNERARI

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	111,877	1	0,000	0,999
Distanza	0,011	0,011	0,984	1	0,321	1,011
A1	-0,010	0,009	1,234	1	0,267	0,990
A2	0,005	0,005	1,051	1	0,305	1,005
Cartolerie	-0,886	0,368	5,782	1	0,016	0,412
Cartolerie*Mq	0,002	0,001	3,007	1	0,008	1,002
Cartolerie*Distanza	0,090	0,063	2,023	1	0,155	1,094
A1*Cartolerie	0,068	0,102	0,447	1	0,504	1,070
A2*Cartolerie	0,041	0,036	1,311	1	0,252	1,042

TABELLA 55 - OUTPUT COMPLESSIVO CARTOLERIE

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	108,287	1	0,000	0,999
Distanza	0,014	0,011	1,651	1	0,199	1,014
A1	-0,010	0,009	1,220	1	0,269	0,990
A2	0,007	0,005	1,838	1	0,175	1,007
FioriePiante	0,141	0,520	0,074	1	0,786	1,152
FioriePiante*Mq	0,000	0,003	0,012	1	0,912	1,000
Distanza*FioriePiante	-0,026	0,092	0,079	1	0,779	0,974
A1*FioriePiante	0,119	0,120	0,991	1	0,320	1,126
A2*FioriePiante	-0,018	0,057	0,097	1	0,756	0,983

TABELLA 56 - OUTPUT COMPLESSIVO FIORI E PIANTE

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	98,234	1	0,000	0,999
Distanza	0,014	0,011	1,711	1	0,191	1,014
A1	-0,011	0,010	1,256	1	0,262	0,989
A2	0,008	0,005	2,190	1	0,139	1,008
Elettronica	0,201	0,233	0,740	1	0,390	1,222
Elettronica*Mq	0,000	0,000	0,030	1	0,862	1,000
Distanza*Elettronica	-0,014	0,039	0,126	1	0,723	0,986
A1*Elettronica	0,115	0,091	1,581	1	0,209	1,121
A2*Elettronica	0,002	0,022	0,008	1	0,928	1,002

TABELLA 57 - OUTPUT COMPLESSIVO ELETTRONICA

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	105,986	1	0,000	0,999
Distanza	0,015	0,011	1,950	1	0,163	1,015
A1	-0,009	0,009	1,091	1	0,296	0,991
A2	0,007	0,005	1,973	1	0,160	1,007
Giocattoli	1,669	0,761	4,811	1	0,028	5,305
Giocattoli*Mq	-0,003	0,002	1,795	1	0,180	0,997
Distanza*Giocattoli	-0,239	0,139	2,928	1	0,008	0,788
A1*Giocattoli	-0,178	0,170	1,086	1	0,297	0,837
A2*Giocattoli	-0,110	0,076	2,098	1	0,147	0,895

TABELLA 58 - OUTPUT COMPLESSIVO GIOCATTOLI

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	108,189	1	0,000	0,999
Distanza	0,014	0,011	1,819	1	0,177	1,014
A1	-0,009	0,009	0,980	1	0,322	0,992
A2	0,008	0,005	2,402	1	0,121	1,008
Librerie	0,586	0,546	1,152	1	0,283	1,797
Librerie*Mq	0,001	0,001	3,227	1	0,072	1,001
Distanza*Librerie	-0,245	0,138	3,154	1	0,076	0,783
A1*Librerie	-0,049	0,146	0,114	1	0,735	0,952
A2*Librerie	-0,063	0,040	2,452	1	0,010	0,939

TABELLA 59 - OUTPUT COMPLESSIVO LIBRERIE

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	99,919	1	0,000	0,999
Distanza	0,012	0,011	1,230	1	0,267	1,012
A1	-0,010	0,009	1,214	1	0,271	0,990
A2	0,007	0,005	1,646	1	0,199	1,007
Minimercati	-0,077	0,652	0,014	1	0,906	0,926
Minimercati*Mq	-0,003	0,001	7,450	1	0,006	0,997
Distanza*Minimercati	0,240	0,113	4,554	1	0,003	1,272
A1*Minimercati	0,111	0,158	0,488	1	0,485	1,117
A2*Minimercati	-0,025	0,076	0,107	1	0,744	0,975

TABELLA 60 - OUTPUT COMPLESSIVO MINIMERCATI

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	106,123	1	0,000	0,999
Distanza	0,010	0,011	0,849	1	0,357	1,010
A1	-0,009	0,009	1,086	1	0,297	0,991
A2	0,006	0,005	1,309	1	0,253	1,006
Mobili	-0,465	0,349	1,770	1	0,183	0,628
Mobili*Mq	0,000	0,000	0,594	1	0,441	1,000
Distanza*Mobili	0,144	0,061	5,692	1	0,010	1,155
A1*Mobili	0,021	0,108	0,037	1	0,848	1,021
A2*Mobili	0,039	0,030	1,711	1	0,191	1,040

TABELLA 61 - OUTPUT COMPLESSIVO MOBILI

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	109,065	1	0,000	0,999
Distanza	0,013	0,011	1,587	1	0,208	1,013
A1	-0,009	0,009	1,025	1	0,311	0,991
A2	0,006	0,005	1,584	1	0,208	1,007
Ottica	-0,513	0,801	0,410	1	0,522	0,599
Mq*Ottica	0,000	0,004	0,000	1	0,989	1,000
Distanza*Ottica	-0,003	0,161	0,000	1	0,987	0,997
A1*Ottica	-0,019	0,074	0,062	1	0,803	0,982
A2*Ottica	0,035	0,070	0,248	1	0,619	1,035

TABELLA 62 - OUTPUT COMPLESSIVO OTTICA

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	110,170	1	0,000	0,999
Distanza	0,016	0,011	2,180	1	0,140	1,016
A1	-0,010	0,009	1,147	1	0,284	0,991
A3	0,006	0,005	1,213	1	0,271	1,006
Profumeria	-0,194	0,298	0,421	1	0,516	0,824
Mq*Profumeria	0,000	0,001	0,120	1	0,729	1,000
Distanza*Profumeria	-0,038	0,053	0,527	1	0,015	0,962
A1*Profumeria	-0,069	0,114	0,365	1	0,546	0,933
A3*Profumeria	0,021	0,025	0,671	1	0,413	1,021

TABELLA 63 - OUTPUT COMPLESSIVO PROFUMERIA

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	126,540	1	0,000	0,999
Distanza	0,014	0,011	1,648	1	0,199	1,014
A1	-0,015	0,011	1,732	1	0,188	0,985
A2	0,005	0,005	0,929	1	0,335	1,005
Quotidiani	-1,837	0,644	8,140	1	0,004	0,159
Mq*Quotidiani	0,001	0,001	2,419	1	0,011	1,001
Distanza*Quotidiani	-0,054	0,111	0,238	1	0,626	0,947
A1*Quotidiani	0,324	0,295	1,205	1	0,272	1,382
A2*Quotidiani	0,018	0,057	0,094	1	0,759	1,018

TABELLA 64 - OUTPUT COMPLESSIVO QUOTIDIANI

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	123,675	1	0,000	0,999
Distanza	0,012	0,011	1,379	1	0,240	1,012
A1	-0,015	0,011	1,827	1	0,176	0,985
A2	0,005	0,005	0,800	1	0,371	1,005
Tabacchi	-1,970	0,762	6,688	1	0,010	0,139
Mq*Tabacchi	0,002	0,005	0,141	1	0,708	1,002
Distanza*Tabacchi	0,057	0,116	0,244	1	0,622	1,059
A1*Tabacchi	-0,197	0,389	0,258	1	0,612	0,821
A2*Tabacchi	-0,005	0,067	0,005	1	0,946	0,995

TABELLA 65- OUTPUT COMPLESSIVO TABACCHI

Variabili nell'equazione						
	B	SE	Wald	gl	Sign.	Exp(B)
Mq	-0,001	0,000	107,658	1	0,000	0,999
Distanza	0,015	0,011	2,080	1	0,149	1,015
A1	-0,008	0,009	0,864	1	0,353	0,992
A2	0,008	0,005	2,595	1	0,107	1,008
Telefonia	0,841	0,370	5,171	1	0,023	2,319
Mq*Telefonia	-0,002	0,002	1,171	1	0,279	0,998
Distanza*Telefonia	-0,079	0,072	1,220	1	0,269	0,924
A1*Telefonia	-0,043	0,040	1,153	1	0,283	0,958
A2*Telefonia	-0,080	0,036	4,857	1	0,015	0,923

TABELLA 66 - OUTPUT COMPLESSIVO TELEFONIA