

Mi sento in dovere di dedicare questa pagina del presente elaborato alle persone che mi hanno supportato nella redazione dello stesso.

Ringrazio la mia famiglia, i miei genitori e mio fratello, perché senza di loro non avrei mai potuto intraprendere questo percorso di studi, per tutti i sacrifici che comporta sostenere un figlio che studia lontano da casa e dagli affetti.

Ringrazio la mia ragazza per tutte le volte che ha saputo ascoltarmi, sostenermi e spronarmi nei momenti più difficili e per avermi dato i migliori suggerimenti.

# Sommario

1. Introduzione .....	5
2. Rassegna della letteratura .....	7
2.1 Modelli storici .....	7
2.1.1 Principio di differenziazione minima di Hotelling.....	7
2.1.2 Modelli di attrazione di Huff.....	8
2.1.3 Metodi analogici di Applebaum.....	8
2.2 Using betweenness metrics to investigate the geographical distribution of retailers .....	8
2.3 Driver che influenzano le decisioni di localizzazione nel settore del retail .....	9
2.3.1 Struttura della Popolazione .....	10
2.3.2 Fattori Economici .....	11
2.3.3 Concorrenza.....	12
2.3.4 Livello di Saturazione .....	14
2.3.5 Caratteristiche del Negozio.....	16
2.3.6 Magnet (Negozzi ancora) .....	17
2.4 Street centrality and densities of retail and services in Bologna, Italy.....	18
2.5 Location and Agglomeration: The Distribution of Retail and Food Businesses in Dense Urban Environments.....	19
2.6 Lively Streets: Exploring the relationship between built environment and social behaviour	20
2.7 Correlation Analysis of Retail Space and Shopping Behaviour in a Commercial Street Based on Space Syntax: A Case of Shijiazhuang, China.....	22
3 Dati e oggetto dello studio.....	24
3.1 Descrizione dei Dati .....	24
3.2 Selezione dei Dati .....	28
3.2.1 Controllo delle attività commerciali: Esclusione di attività chiuse prima del 2020 .....	28
3.2.2 Verifica delle coordinate geografiche: Esclusione dei record senza informazioni geografiche .....	28

3.2.3 Controllo di consistenza dei dati: Identificazione e correzione di dati anomali o mancanti	29
4 Metodologia	30
4.1 Preparazione dei Dati	30
4.1.1 Creazione di Variabili Aggiuntive	30
4.2 Analisi Spaziale	32
4.2.1 Utilizzo di sistemi di coordinate: Assicurarsi che i dati siano coerenti e compatibili con i sistemi di coordinate utilizzati	32
4.2.2 Definizione di parametri topografici: Identificazione di indicatori topografici rilevanti (es. elevazione, inclinazione del terreno)	33
4.2.3 Mappatura: Creazione di mappe che rappresentino la distribuzione delle attività commerciali sul territorio	33
4.3 Analisi delle Attività Commerciali in Base al Codice ATECO e alla Tipologia di Strada	40
4.4 Analisi Statistica	41
4.5 Verifica dei Risultati	43
4.6 Strumenti e Software Utilizzati	44
5 Risultati	46
6 Conclusioni	50
7 Bibliografia	53
8 Indice delle figure	55
9 Appendice	57
9.1 Database	57
9.2 Tabelle	59
9.3 Diagramma ER e query	65
9.4 Codici Python	68

# 1. Introduzione

Nel contesto urbano contemporaneo, la relazione tra la topologia della città e le diverse tipologie di attività commerciali che si sviluppano sul territorio rappresenta un tema di crescente interesse e rilevanza. Le città sono organismi complessi in continua evoluzione, caratterizzati da una varietà di elementi che vanno dalla morfologia urbana alla distribuzione delle attività economiche. La comprensione di come la struttura fisica e spaziale di una città influenzi e sia influenzata dalle attività che vi si svolgono è cruciale per la pianificazione urbana, lo sviluppo economico e la qualità della vita dei cittadini.

L'obiettivo di questa ricerca è indagare la correlazione esistente tra la topologia urbana e la tipologia delle attività commerciali presenti sul territorio. La legislazione urbana e commerciale può giocare un ruolo significativo nel determinare le opportunità e le sfide che le imprese affrontano nello svolgere le proprie attività all'interno di una determinata area urbana. Le normative urbanistiche, ad esempio, possono influenzare la densità commerciale, la tipologia di attività consentite in determinate zone, nonché le modalità di accesso al mercato per nuovi operatori economici. Uno degli aspetti fondamentali da considerare in questa analisi è il concetto di avviamento. L'avviamento rappresenta un elemento cruciale per molte attività commerciali e imprenditoriali, definendosi come un bene immateriale a sé stante, con un valore economico determinabile il cui valore deriva dalla reputazione, dalla clientela acquisita, dalla posizione e dalle relazioni commerciali stabilite nel corso del tempo. Pertanto, comprendere come la topologia della città e le disposizioni normative influenzino la creazione, il mantenimento e lo sviluppo dell'avviamento delle attività commerciali è di fondamentale importanza per comprendere la distribuzione e la diversificazione delle attività sul territorio urbano.

L'avviamento è il valore aggiunto di un'azienda grazie alla reputazione, clientela fedele e posizione strategica. Viene considerato un patrimonio immateriale che contribuisce alla redditività e competitività. La posizione in città è cruciale. Il Codice civile italiano fornisce regole per valutare e proteggere l'avviamento nelle transazioni aziendali.

La posizione geografica dell'azienda può influenzare significativamente il valore dell'avviamento. Ad esempio, se l'azienda si trova in una posizione strategica, come una zona di grande passaggio o vicino a centri commerciali, la sua visibilità e accessibilità possono aumentare il valore dell'avviamento in quanto la presenza in una zona di alto traffico può attirare più clienti e

contribuire alla crescita del business. Inoltre, la posizione geografica può determinare il mercato di riferimento dell'azienda, ad esempio, se l'azienda si trovasse in una zona residenziale, potrebbe avere una clientela diversa rispetto a un'azienda situata in una zona industriale. Questo può influenzare la base di clienti fedeli e la reputazione dell'azienda. La posizione geografica può anche influenzare la concorrenza. Se un'azienda opera in un'area con pochi concorrenti diretti potrebbe avere un vantaggio competitivo. D'altra parte, se è in una zona saturata di aziende simili, l'avviamento potrebbe essere meno significativo. Alcune zone geografiche poi sono associate a prestigio o status elevato per cui un'azienda situata in una zona di lusso o in un quartiere esclusivo potrebbe beneficiare di un'immagine positiva e di un valore aggiunto.

In questo contesto, la presente ricerca si propone di esaminare in che modo la struttura e la topografia della città, influenzino la distribuzione spaziale delle attività economiche e l'avviamento delle imprese. Attraverso un'analisi multidisciplinare che integra approcci geografici ed economici, nella fattispecie attraverso la correlazione tra elementi fisici dell'immobile e delle strade su cui insistono le attività commerciali e la classificazione delle stesse in gruppi sulla base del loro codice ATECO si mira a delineare una comprensione più approfondita dei meccanismi che guidano l'interazione tra la morfologia urbana e le dinamiche economiche locali.

## 2. Rassegna della letteratura

### 2.1 Modelli storici

Nel contesto degli studi sulla localizzazione delle imprese, vengono esaminati tre modelli chiave che hanno contribuito significativamente alla comprensione delle strategie di posizionamento. Questi modelli costituiscono i fondamenti per comprendere le strategie di localizzazione delle imprese, mettendo in evidenza l'importanza di variabili come l'incertezza competitiva, l'attrattività del negozio e gli analoghi, secondo la definizione data da Applebaum, nella definizione delle posizioni aziendali.

#### 2.1.1 Principio di differenziazione minima di Hotelling

Il principio di differenziazione minima, proposto da Hotelling, affronta la scelta della posizione di due imprese in una città lineare. Si basa su assunzioni quali la domanda inelastica, la distribuzione uniforme degli abitanti lungo la città e la presenza di un costo di trasporto. Hotelling dimostra che, nonostante una distribuzione equa sia socialmente ottimale, l'interazione strategica, ovvero quella situazione che si verifica quando due o più agenti decisionali sono influenzati dall'esito delle proprie scelte e quelle degli altri agenti come descritto nella teoria dei giochi, porta a un clustering subottimale al centro della città. Questo evidenzia il ruolo critico dell'interazione strategica nella determinazione delle posizioni aziendali.

Hotelling constata che il profitto di A aumenta se si avvicina a B. I due negozi hanno quindi interesse ad avvicinarsi. Hotelling suggerisce che ciò spiega anche la tendenza all'uniformazione dei prodotti o dei programmi dei partiti politici (differenza tra Democratici e Repubblicani).

### 2.1.2 Modelli di attrazione di Huff

Huff nel trattato "A Probabilistic Analysis of Shopping Center Trade Areas" (1964) analizza la scelta della localizzazione commerciale attraverso un modello gravitazionale. Esso considera la massa (dimensione del negozio) e la distanza (dal consumatore) come variabili influenti. Il modello suggerisce che maggiore è l'attrattività del negozio, maggiore è la distanza che i consumatori sono disposti a percorrere. Ciò sottolinea l'importanza dell'attrattività del negozio nella definizione della sua influenza sulla decisione di acquisto.

### 2.1.3 Metodi analogici di Applebaum

Il lavoro di Applebaum introduce un metodo per determinare le aree commerciali e stimare la penetrazione di mercato dei negozi esistenti. Il modello utilizza trade area concentriche e un'area commerciale principale per rappresentare la densità di clientela. Per trade area si intende un'area commerciale definita come l'area essenziale più vicina al punto vendita. Applebaum propone l'uso di analoghi, ossia negozi esistenti con caratteristiche simili, come standard di misura per prevedere le prestazioni di nuovi punti vendita. Questo approccio fornisce previsioni accurate basate sull'esperienza pregressa.

## 2.2 Using betweenness metrics to investigate the geographical distribution of retailers

Nel mondo del commercio al dettaglio, dunque, l'accessibilità e l'attrattività di una posizione sono strettamente legate alla disposizione spaziale di altri negozi e consumatori. La ricerca "*Using betweenness metrics to investigate the geographical distribution of retailers*" pubblicata dai professori Buzzacchi, Zotteri e Taramino (2020) ha dimostrato che un luogo diventa più attraente per i rivenditori quando è attraversato frequentemente dalle rotte seguite dai consumatori. Lo studio si basa sul fatto che esistono almeno due tipi di rotte seguite dai consumatori: gli spostamenti giornalieri tra la residenza e il luogo di lavoro e i viaggi dedicati allo shopping tra i negozi.

In questo studio, vengono esaminati gli impatti di entrambi questi tipi di rotte dei consumatori sui modelli commerciali a Torino. I risultati mostrano che i viaggi giornalieri tra residenza e lavoro non contribuiscono significativamente a incrementare l'attrattività di un negozio lungo il percorso, a differenza dei viaggi dedicati allo shopping. In particolare, viene evidenziato che i vantaggi che un negozio può ottenere dalla sua posizione lungo le rotte dipendono dal tipo di prodotti che vende.

Per effettuare queste analisi sono stati considerati tre indici:

- *Population density index*
- *Retailer density index*
- *Betweenness indexes*

utilizzati per molteplici scopi nel campo del design urbano. Il population density index, il retailer density index e il betweenness index sono stati applicati, tra gli altri, per studiare il loro effetto sulla dinamica residenziale (Omer, 2005), sul volume di pedoni (Kang, 2015) e sulla rete di trasporti pubblici (Porta e Scheurer, 2006)

L'obiettivo dell'analisi riportata in questa tesi è dunque definire anche come le rotte non lavorative sono definite sulla base della topografia della città di Torino.

### 2.3 Driver che influenzano le decisioni di localizzazione nel settore del retail

La distribuzione geografica dei rivenditori è influenzata da tre fasi: analisi della posizione dei retailers già presenti sul mercato, la selezione della posizione e l'uscita dal mercato.

Assumiamo che i rivenditori tendano a localizzarsi dove possono massimizzare i profitti. Secondo uno studio condotto da Turhan et al. (2013) i fattori chiave nella scelta della posizione dei punti vendita sono finalizzati alla massimizzazione delle prestazioni del punto vendita stesso, concentrandosi su criteri legati alle prestazioni del negozio stesso. La scelta della posizione si basa sulla prospettiva di ottenere le migliori prestazioni attese. La qualità dei modelli utilizzati dipende dalla capacità di prevedere gli obiettivi di performance del rivenditore, spesso espressi in termini di volumi di vendite, profitti del negozio, quote di mercato e elasticità al prezzo.

Turhan et al. (2013) classificano i fattori che influenzano la scelta della posizione in sei categorie: struttura della popolazione, fattori economici, concorrenza, livello di saturazione, caratteristiche del negozio e magnetismo [FIGURA 1].

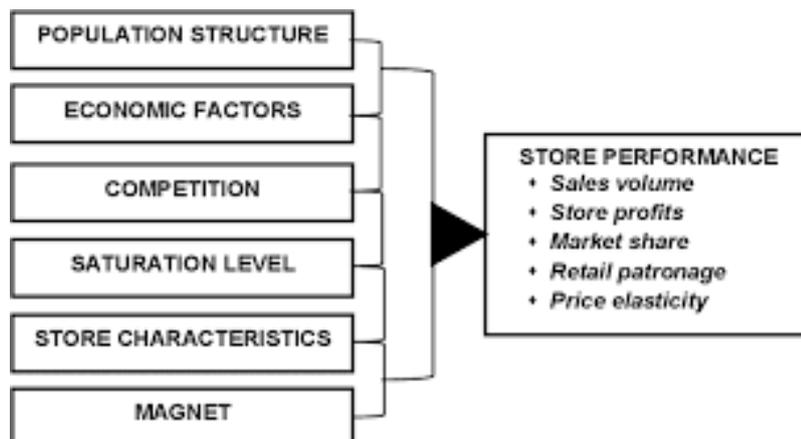


Figura 1: Driver che influenzano la scelta sulla base delle store performances

### 2.3.1 Struttura della Popolazione

Questo fattore prende in considerazione la composizione demografica della popolazione nella zona di interesse. Ad esempio, la presenza di un target di clientela specifico, come famiglie con reddito elevato, giovani professionisti o anziani, può influenzare la decisione di localizzazione. Le caratteristiche demografiche, come la dimensione della famiglia, il livello di istruzione e la distribuzione dell'età, sono prese in considerazione per massimizzare l'attrattiva del negozio per il pubblico di riferimento.

Alcuni studi suggeriscono che comprendere la struttura demografica della popolazione è fondamentale per la segmentazione del mercato. Ad esempio, una ricerca condotta da Kotler e Armstrong (2016) evidenzia l'importanza di adattare l'offerta di prodotti e servizi alle caratteristiche demografiche della popolazione di riferimento.

Gli studi condotti da Porta et al. (2009) e Sevtsuk (2014) esplorano l'importanza del flusso di persone all'interno della città e il suo impatto sui punti vendita, dimostrando quindi come sia necessario tenere conto non solamente della struttura della popolazione residente ma anche di quella che si sposta per seguire le tratte dello shopping

Attraverso l'utilizzo dell'indice di "betweenness", le ricerche evidenziano la capacità di un determinato punto geografico di figurare all'interno delle rotte ottimali che collegano altri due punti, consentendo una stima della frequenza di passaggio di consumatori in quel luogo durante

gli spostamenti all'interno dell'area geografica presa in considerazione, prendendo in considerazione i tragitti più efficienti e diretti.

Il lavoro di Buzzacchi et al. (2020) approfondisce ulteriormente questo contesto, enfatizzando due percorsi di consumo distinti: i percorsi dei pendolari, che collegano la residenza ai luoghi di lavoro, e le uscite dedicate allo shopping tra i negozi. Lo studio evidenzia che, il successo dei rivenditori sembra essere più strettamente legato agli spostamenti delle persone per fare acquisti rispetto agli spostamenti quotidiani casa-lavoro.

Questi risultati suggeriscono che la posizione strategica di un punto vendita in prossimità di aree con un elevato flusso di consumatori orientati allo shopping può essere cruciale per il successo del negozio.

### 2.3.2 Fattori Economici

Questa categoria comprende vari aspetti economici che possono influenzare la scelta della posizione. Ad esempio: la disponibilità di lavoro, il reddito medio della zona, il tasso di disoccupazione e la stabilità economica locale sono considerati. Inoltre, fattori come la capacità di spesa della popolazione e la presenza di zone commerciali floride possono essere determinanti nella decisione di localizzazione.

Ricerche come quelle di Berry e Waldfogel (2001) indicano che i fattori economici, come il reddito medio della zona, possono influenzare significativamente il comportamento di acquisto dei consumatori e la redditività dei punti vendita.

Secondo Karande & Lombard (2005), variabili come il tipo di abitazioni, il loro prezzo, la percentuale di proprietà delle case e le vendite pro-capite forniscono un quadro della situazione economica dei residenti nella zona considerata per la localizzazione del retailer. Sugeriscono, inoltre, che i costi opportunità, in termini di tempo dedicato all'acquisto, sono positivamente correlati al salario della persona. Pertanto, i retailer tendono a posizionarsi in zone economicamente ricche, dove il valore del costo opportunità è più alto per individui ad alto reddito.

Dubé et al. (2016) sostengono che la decisione finale sulla localizzazione di un'azienda è correlata all'ambiente economico circostante, che dovrebbe essere influenzato dalla distanza da centri principali, come centri urbani o luoghi con elevata accessibilità e concentrazione di attività.

### 2.3.3 Concorrenza

La presenza e la tipologia di concorrenti nella zona sono fattori cruciali. L'analisi della concorrenza coinvolge la valutazione della presenza di altri rivenditori simili o complementari nella vicinanza. Nella selezione della posizione i rivenditori devono considerare l'ambiente competitivo come un fattore che influisce sulle prestazioni del negozio. La concorrenza, sia diretta che indiretta, è quindi un aspetto cruciale. Inoltre, la reputazione e la forza dei concorrenti possono influire sulla decisione di localizzazione, cercando di differenziarsi o di sfruttare la concorrenza in modo strategico.

L'analisi della concorrenza è un tema ampiamente trattato nella letteratura.

Studi come quelli di Porter (1980) hanno introdotto il concetto delle forze competitive che modellano l'ambiente competitivo di un'industria, influenzando le strategie di posizionamento delle imprese. Le 5 forze di Porter, concepite da Michael Porter, costituiscono un modello analitico ampiamente utilizzato per valutare l'attrattiva e la competitività di un settore industriale. Esse forniscono un quadro concettuale utile per comprendere il contesto competitivo in cui le imprese operano e per identificare le strategie più adeguate per competere efficacemente. In primo luogo, la minaccia dei nuovi entranti rappresenta la valutazione della facilità con cui nuove imprese possono penetrare nel mercato di un settore specifico. L'incremento dell'accessibilità all'ingresso nel settore può aumentare la competizione esistente, portando a una diminuzione dei profitti per le aziende consolidate. Un altro elemento critico è la minaccia dei prodotti sostitutivi, che valuta la possibilità che i consumatori scelgano alternative ai prodotti o servizi offerti dal settore in esame. La presenza di prodotti sostitutivi può influenzare negativamente la domanda dei prodotti principali del settore, influenzando prezzi e quote di mercato. Il potere contrattuale dei fornitori riflette la capacità dei fornitori di imporre condizioni, aumentare i prezzi o ridurre la qualità dei beni o servizi forniti. Se i fornitori detengono un forte potere contrattuale, possono esercitare pressioni significative sulle imprese del settore, minando la loro redditività. Inoltre, il potere contrattuale dei clienti rappresenta la capacità dei clienti di influenzare i prezzi, le condizioni di vendita e la qualità dei prodotti o servizi offerti dal settore. L'abbondanza di opzioni per i clienti può comportare una maggiore pressione sui prezzi e sulla qualità dei prodotti. Infine, l'intensità della rivalità tra concorrenti esistenti valuta il livello di competizione tra le imprese operanti nel settore. Un'elevata rivalità può ridurre la redditività, poiché le imprese competono per conquistare quote di mercato, clienti e risorse. In generale, il contesto geografico può determinare l'accesso alle risorse, la disponibilità di

alternative, la dimensione del mercato di riferimento e le dinamiche competitive locali, tutti fattori che influenzano le cinque forze di Porter e, di conseguenza, la competitività complessiva di un'azienda.

Le giuste posizioni dei punti vendita sono essenziali per attirare clienti e possono persino compensare strategie di marketing poco efficaci. Una posizione strategica può costituire un forte vantaggio competitivo, poiché è considerata un elemento "unico" nel mix di marketing al dettaglio e non può essere facilmente imitata dai concorrenti (Zentes et al., 2007).

Gli studi di Hoch et al. (1995) indicano che i fattori competitivi rappresentano la maggior parte della variazione dell'elasticità del prezzo tra i negozi. La concorrenza diretta implica la competizione con rivenditori che offrono gli stessi prodotti, mentre la concorrenza indiretta considera anche rivenditori che offrono prodotti diversi ma competono per il potere di consumo della stessa area geografica (Turhan et al., 2013).

La letteratura affronta la questione della concorrenza spaziale e sottolinea che, nella selezione della posizione di un nuovo punto vendita, è essenziale considerare diversi fattori legati alla concorrenza. Questi includono la distanza spaziale tra i negozi, la dimensione e il numero dei punti vendita concorrenti, le alternative di acquisto e il volume delle vendite dei concorrenti (Fatima et al., 2017).

Tuttavia, la letteratura mostra una lacuna in termini di approfondimento sulla determinazione della distanza tra i negozi concorrenti e sulle strategie adottate dai rivenditori in base a questa distanza. Uno studio condotto da Karade e Lombard (2005) affronta questa lacuna, introducendo il concetto di associazione spaziale. Utilizzando questa metrica, misurano la competizione nello spazio tra diverse distribuzioni di punti, fornendo una panoramica più dettagliata delle dinamiche competitive. Il loro studio si basa sulla distanza come misura di competizione e rivela che i rivenditori possono adottare strategie di prossimità o di lontananza a seconda della distanza spaziale tra loro. I risultati evidenziano che i rivenditori situati entro un certo raggio gli uni dagli altri adottano una strategia di prossimità, mentre quelli a distanze maggiori implementano una strategia di lontananza.

In conclusione, la letteratura dimostra che la concorrenza spaziale è un aspetto cruciale nella scelta della posizione di un punto vendita, e le strategie adottate dai rivenditori possono variare in base alla distanza spaziale tra di essi.

### 2.3.4 Livello di Saturazione

Il livello di saturazione riflette quanto il mercato nella zona è già saturo di punti vendita simili o di prodotti simili. Se il mercato è altamente saturo, la concorrenza potrebbe essere più intensa e le opportunità di crescita potrebbero essere limitate. D'altra parte, in mercati meno saturi, potrebbe esserci maggiore spazio per la crescita e l'espansione. Risulta importante considerare anche l'agglomerazione nei centri commerciali poiché in questo caso la presenza di un elevato numero di attività che offrono prodotti simili può influenzare positivamente le opportunità di crescita dato che permette di ridurre il costo sostenuto dal cliente per raggiungere diversi negozi creando un effetto di destinazione.

Ricerche come quelle di Glaeser et al. (2000) hanno esaminato l'effetto della densità di rivenditori su vari aspetti economici, tra cui la competitività, l'innovazione e la produttività.

Il lavoro di Turhan et al. (2013) propone l'indice di saturazione della vendita al dettaglio (IRS) come una metrica significativa per valutare l'attrattiva di un particolare mercato nel settore retail. Questo indice è fondamentale perché fornisce al retailer un'indicazione sulla probabilità di aumentare la propria profittabilità aprendo un punto vendita in una specifica località.

L'IRS si basa sul rapporto tra la domanda di un prodotto o servizio e l'offerta disponibile in relazione a una specifica area che si intende prendere in esame. La formula per calcolare l'IRS è la seguente:

$$IRS = \frac{(P)(A. E.)}{S}$$

Dove:

- *IRS* è l'indice di saturazione del settore retail per una specifica area geografica.
- *P* è il numero di persone nell'area, che rappresentano i potenziali clienti per uno specifico prodotto in vendita.
- *A. E.* è la spesa media per quella particolare linea di prodotti nell'area presa in esame.
- *S* è la quantità totale di spazio dedicato alla vendita di quel prodotto in tutti i negozi di quell'area geografica, espressa in piedi quadrati.

L'IRS offre agli analisti e agli imprenditori nel settore del commercio al dettaglio uno strumento cruciale per valutare l'attrattiva di un mercato. Più precisamente, fornisce una misura

quantitativa della relazione tra domanda e offerta per una specifica linea di prodotti o servizi. Le imprese possono utilizzare l'IRS per prendere decisioni strategiche riguardo al posizionamento dei loro punti vendita. Ad esempio, se l'IRS di un'area è elevato, indica una forte domanda rispetto all'offerta, suggerendo un potenziale per ottenere profitti elevati. Al contrario, un IRS basso può indicare una saturazione di mercato o una domanda insufficiente per sostenere ulteriori attività.

Poiché l'IRS può variare in base a diversi fattori, come le tendenze di mercato, le preferenze dei consumatori, la concorrenza locale e le politiche economiche è importante che le aziende monitorino regolarmente l'IRS e adattino le loro strategie di conseguenza. Alcuni studiosi hanno suggerito di affinare l'IRS utilizzando misure basate sul numero e sulla spesa delle famiglie anziché sulle singole persone. Questo approccio potrebbe fornire una rappresentazione più accurata della domanda e delle abitudini di spesa nelle comunità.

L'IRS riflette il grado di soddisfazione della domanda di beni e servizi in un'area geografica da parte delle strutture retail attive. Dunne & Lusch (2008) spiegano che l'offerta di negozi al dettaglio può essere vista come il numero di negozi nel mercato o come il numero di metri quadrati dedicato alla vendita di quel prodotto, mentre la domanda è identificata dalle vendite totali dei negozi nel mercato.

L'IRS rappresenta quindi uno strumento prezioso nel toolkit strategico del commercio al dettaglio, consentendo alle imprese di valutare le opportunità di mercato e di prendere decisioni informate per massimizzare i profitti e soddisfare le esigenze dei clienti.

### 2.3.5 Caratteristiche del Negozio

Questa categoria considera le specifiche caratteristiche del negozio stesso. Ciò include la dimensione del negozio, la sua visibilità, l'accessibilità, il parcheggio, e altre caratteristiche fisiche.

La scelta di posizionarsi in una zona ad alto traffico o in un centro commerciale può essere guidata dalle caratteristiche del negozio e dalle esperienze di shopping che si desidera offrire.

Studi sulla psicologia del consumatore, come quelli di Kotler e Keller (2016), evidenziano l'importanza di fattori come la visibilità del negozio e l'accessibilità nella formazione delle decisioni di acquisto dei consumatori.

Secondo Turhan et al. (2013), tre categorie fondamentali di aspetti del negozio sono la facilità di accessibilità, gli attributi dell'immagine del negozio e i costi:

- La facilità di accesso che si riferisce alla capacità delle persone di trovare il negozio facilmente e rapidamente. Questo aspetto è cruciale nella selezione dell'ubicazione del negozio, soprattutto considerando che molti consumatori si spostano in auto. Le strade, le vie e i parcheggi devono essere progettati per facilitare l'accesso alla zona commerciale. Studi come quelli di Setvsuk (2014) e Buzzacchi et al. (2020) hanno approfondito il flusso di persone all'interno delle città, sottolineando l'importanza dell'accessibilità per i negozi.
- L'efficacia dell'ubicazione su percorsi di viaggio dipende anche dal tipo di merce venduto. I negozi che offrono prodotti di prima necessità traggono vantaggio dall'essere situati su rotte di viaggio frequentate, mentre quelli che vendono prodotti soggetti a shopping comparativo possono prosperare in centri commerciali situati anche in aree suburbane. La tipologia di prodotto influenza anche la scelta tra centri urbani e zone periferiche o industriali, come dimostra l'esempio del caso IKEA che si insedia in zone periferiche per motivi logistici ed economici legati anche al layout degli store di tipo cash and carry.
- Gli attributi dell'immagine del negozio, come atmosfera, layout e assortimento di prodotti, influenzano significativamente le prestazioni del punto vendita. Migliorare l'atmosfera o ampliare gli assortimenti può impattare positivamente sui flussi di clienti e sulle spese, ma i rivenditori devono valutare attentamente l'equilibrio tra investimenti e profitti.

Fox et al. (2007) aggiungono anche la tipologia di prodotto venduto come un elemento chiave da considerare proponendo un modello di aggregazione spaziale che considera l'agglomerazione di diverse tipologie di negozi (agglomerazione inter-type) o cluster di negozi dello stesso tipo (agglomerazione intra-type). Questo modello offre una prospettiva sulla disposizione spaziale dei negozi in base alla tipologia di prodotti che offrono.

Complessivamente, la considerazione attenta di questi fattori, inclusi accessibilità, tipologia di prodotto, attributi dell'immagine del negozio e costi, è essenziale per i rivenditori che cercano di ottenere vantaggi competitivi e massimizzare le prestazioni del punto vendita.

### 2.3.6 Magnet (Negozi ancora)

Questa categoria si riferisce alla capacità di un negozio di attirare clienti. Può includere la presenza di negozi "ancora", ovvero negozi di destinazione che possono attirare un flusso costante di clienti nella zona. La co-localizzazione con negozi di successo o attrattivi può aumentare l'appeal complessivo della posizione.

Secondo Timmermans (1986), i rivenditori dovrebbero considerare la presenza di "negozi ancora" come un vantaggio, poiché la vicinanza a tali negozi può favorire un maggiore traffico di consumatori. Di conseguenza, nella scelta della posizione del nuovo punto vendita, i rivenditori potrebbero beneficiare dall'utilizzare la presenza di "negozi ancora" come un criterio di valutazione. La presenza di questi negozi può essere un elemento determinante per attirare clienti nel proprio negozio e massimizzare le opportunità di vendita.

La co-localizzazione di negozi "ancora" è stata discussa in studi sulla teoria dei centri commerciali, ad esempio da Green e Malpezzi (2003), che hanno esaminato il ruolo dei negozi di destinazione nella pianificazione commerciale.

## 2.4 Street centrality and densities of retail and services in Bologna, Italy

Per analogia con lo scopo della tesi risulta particolarmente interessante lo studio “*Street centrality and densities of retail and services in Bologna, Italy*” di Porta, Strano, Iacoviello, Messori, Latora, Cardillo, Wang e Scellato (2009).

In questo studio viene esaminata la relazione tra centralità delle strade e densità di attività commerciali nella città di Bologna Italia. I risultati indicano che le attività commerciali presenti a Bologna tendono a concentrarsi nelle aree con una migliore centralità che viene definita come:

- La *Closeness centrality* ( $C^c$ ) misura quanto un nodo è vicino a tutti gli altri nodi lungo i percorsi più brevi della rete, definita come:

$$C_i^c = \frac{N - 1}{\sum_{j=1, j \neq i}^N d_{ij}}$$

Dove  $N$  è il numero totale di nodi nella rete, e  $d_{ij}$  è la distanza più breve tra i nodi  $i$  e  $j$ . In altre parole, la Closeness centrality per un nodo è l'inverso del rapporto tra distanza media da questo nodo a tutti gli altri nodi.

La  $C^c$  può essere interpretata come la prossimità di un nodo all'altro e cattura anche il concetto di accessibilità di un luogo. Considerando un reticolo in cui gli archi tra due nodi corrispondono a delle tratte percorribili si potrebbe dire che più un nodo è vicino ad altri nodi, più è accessibile. La famiglia di misure di centralità di vicinanza è stata ampiamente utilizzata nell'analisi urbana e regionale. In sostanza, riflette il costo per un cliente di superare le separazioni spaziali tra luoghi con popolazione e attività. Per costo in questo caso non si intende effettivamente il costo economico per spostarsi da un nodo all'altro ma quanto più la complessità di spostamento e il conseguente effort che questo implica per chi deve effettuare questo spostamento.

- *Betweenness centrality* ( $C^b$ ) si basa sull'idea che un nodo sia più trafficato quando è attraversato da un numero maggiore di percorsi più brevi che collegano tutte le coppie di nodi nella rete.  $C^b$  è definita come:

$$C_i^b = \frac{1}{(N - 1)(N - 2)} \sum_{j=1, k=1, k \neq j \neq i}^N \frac{n_{jk}(i)}{n_{jk}}$$

Dove  $n_{jk}$  è il numero di percorsi più brevi tra i nodi  $j$  e  $k$ , e  $n_{jk}(i)$  è il numero di questi percorsi più brevi che contengono il nodo  $i$ .  $C^b$  cattura una proprietà speciale di un luogo in una città: non agisce come origine o destinazione per i viaggi, ma come un punto di passaggio.  $C^b$  rappresenta il volume di traffico di passaggio di un nodo. Un luogo con una migliore betweenness centrality può beneficiare di questa importante proprietà.

- *Straightness centrality* ( $C^s$ ) deriva dall'idea che l'efficienza della comunicazione tra due nodi aumenta quando c'è meno deviazione dal percorso più breve tra loro rispetto alla linea virtuale retta che li collega, ovvero una maggiore "rettilinearità" del percorso più breve.

$C^s$  è definita come:

$$C_i^s = \frac{1}{N-1} \sum_{j=1, j \neq i}^N \frac{d_{ij}^{Eucl}}{d_{ij}}$$

Dove  $d_{ij}^{Eucl}$  è la distanza euclidea tra i nodi  $i$  e  $j$ , ovvero la lunghezza della connessione virtuale retta.

## 2.5 Location and Agglomeration: The Distribution of Retail and Food Businesses in Dense Urban Environments

Allo stesso modo dei precedenti studi anche "Location and Agglomeration: The Distribution of Retail and Food Businesses in Dense Urban Environments" di Sevtsuk nel 2014 offre interessanti spunti su come le caratteristiche specifiche di una posizione influiscono sulle probabilità di trovare attività commerciali, il che può aiutare la pianificazione di cluster di determinate attività commerciali. In sostanza, l'obiettivo dell'articolo risulta essere quello di fornire informazioni pratiche utili sia per i retailer che per i legislatori sulle migliori strategie di posizionamento per le attività commerciali di vendita al dettaglio e punti di ristoro, al fine di migliorare la vivibilità e

la percorribilità a piedi nei quartieri studiati. In quest'ottica i retailer potrebbero ottenere maggiori profitti e gli enti regolatori potrebbero aumentare la vivibilità dei quartieri promuovendo, attraverso leggi ed incentivi, lo sviluppo di determinate categorie di attività commerciali.

Dai risultati si nota come diverse categorie di negozi mostrano preferenze di localizzazione differenti, con evidenze di clustering competitivo tra negozi meno frequentati come quelli di elettronica, hobby, libri e musica. L'accesso ai passanti risulta essere un fattore cruciale per il successo degli esercizi commerciali, spesso più importante dell'accesso diretto da parte di residenti o dipendenti.

La capacità di acquisto dei residenti e le loro entrate non giocano un ruolo significativo per quanto riguarda la struttura degli edifici, con altri fattori come l'accesso alle stazioni della metropolitana che risultano più influenti. L'analisi individua edifici con caratteristiche ideali per attività commerciali dal punto di vista topografico ma attualmente prive di attività commerciali avviate, suggerendo opportunità potenziali per l'adattamento della pianificazione urbana e suggeriscono che le autorità di pianificazione possono influenzare la densità commerciale attraverso linee guida di sviluppo, incentivi fiscali e miglioramenti dell'accesso al trasporto pubblico.

In definitiva, l'articolo propone un approccio diverso alla pianificazione urbana, suggerendo che comprendere le condizioni che favoriscono lo sviluppo spontaneo di attività commerciali è essenziale per una progettazione urbana efficace.

## 2.6 Lively Streets: Exploring the relationship between built environment and social behaviour

Lo studio "Lively Streets: Exploring the relationship between built environment and social behaviour" di V. Metha (2008) tratta in maniera più ampia la stessa tematica dello studio proposto in questa tesi, cercando di dimostrare una relazione positiva tra alcune caratteristiche topografiche e la vitalità delle strade commerciali. Questa si può riflettere ovviamente anche in una maggiore concentrazione della attività commerciali nelle vie che soddisfano queste caratteristiche.

L'autore identifica undici caratteristiche specifiche dell'ambiente stradale basate sulla rassegna della letteratura, interviste e osservazioni estensive. Tali caratteristiche includono sedute commerciali come dehors o tavoli all'aperto, facciate stradali articolate, vetrine personalizzate, luoghi comunitari, vetrine ben visibili, ombra degli alberi, utilizzi indipendenti, sedute pubbliche come panchine o arredo urbano pubblico in generale, arredi e manufatti fisici, larghezza del marciapiede e varietà di attività commerciali. Viene mostrata una correlazione positiva tra queste caratteristiche e la vitalità delle strade commerciali di quartiere.

Attraverso un'analisi multivariata, si dimostra che le sedute commerciali, le sedute pubbliche, i luoghi di ritrovo della comunità, la personalizzazione delle vetrine e la larghezza del marciapiede hanno un impatto significativo sulla vitalità della strada. Alcune caratteristiche, come la permeabilità delle vetrine, mostrano una correlazione elevata ma non significativa nella regressione.

Successivamente, viene condotta un'analisi fattoriale che riduce le undici caratteristiche a quattro fattori principali: qualità dell'uso del suolo, qualità fisiche, sedute commerciali e luoghi comunitari. Questi quattro fattori spiegano il 73% della varianza nella vitalità delle strade commerciali di quartiere. In particolare, le sedute commerciali, la larghezza del marciapiede, la presenza di sedute pubbliche e luoghi comunitari emergono come elementi chiave per sostenere l'attività sociale e la vitalità della strada.

Viene dunque evidenziata l'importanza di una serie di caratteristiche stradali nell'incoraggiare attività sociali e commerciali sulle strade urbane, fornendo indicazioni preziose per la progettazione e la gestione degli spazi pubblici.

## 2.7 Correlation Analysis of Retail Space and Shopping Behaviour in a Commercial Street Based on Space Syntax: A Case of Shijiazhuang, China

Ulteriore forza ai risultati ottenuti durante lo sviluppo di questa tesi è fornita dallo studio “Correlation Analysis of Retail Space and Shopping Behavior in a Commercial Street Based on Space Syntax: A Case of Shijiazhuang, China” di Haitao Lian e Gaomei Li (2023).

La relazione tra la vitalità degli spazi commerciali e l'attrattività delle vetrine è una considerazione importante per il successo delle attività commerciali. È stato dimostrato che le attività preferiscono insediarsi su immobili con vetrine ben visibili e attraenti. Questo perché le vetrine svolgono un ruolo cruciale nell'attrarre l'attenzione dei potenziali clienti e nel comunicare l'immagine e l'offerta del negozio.

Nel contesto dello studio citato, i risultati mostrano che i parametri di visibilità, come la connettività dello spazio e i coefficienti di profondità e spazio delle vetrine dei negozi, influenzano positivamente la vitalità degli spazi commerciali nelle strade urbane. La connettività dello spazio si riferisce alla capacità di un ambiente di facilitare o ostacolare la comunicazione e il movimento tra le persone, gli oggetti e gli elementi presenti nello spazio stesso. Si tratta di un concetto ampio che può includere diverse dimensioni, come la disposizione fisica degli oggetti, la facilità di accesso e la percezione degli spazi aperti o chiusi.

I coefficienti di profondità e spazio delle vetrine dei negozi sono termini usati nel contesto del visual merchandising e del design degli spazi commerciali. Questi coefficienti si riferiscono alla profondità e allo spazio visivo disponibile all'interno e intorno alle vetrine dei negozi.

Il coefficiente di profondità si riferisce alla distanza fisica dalla vetrina alla parete posteriore del negozio. Una maggiore profondità può consentire una migliore visualizzazione dei prodotti esposti e offrire più opportunità per il display creativo. Il coefficiente di spazio si riferisce all'area visiva intorno alla vetrina, inclusi gli spazi aperti o congesti, l'illuminazione ambientale e qualsiasi elemento architettonico che possa influenzare la percezione dello spazio intorno alla vetrina stessa. Un coefficiente di spazio adeguato può contribuire ad attirare l'attenzione dei passanti e a migliorare l'attrattività complessiva del negozio. Le vetrine ben integrate e visibili sono fondamentali per attirare clienti e aumentare l'attività commerciale. Questo concetto è importante non solo per gli operatori commerciali, ma anche per gli urbanisti e i progettisti urbani che lavorano alla progettazione e allo sviluppo di spazi urbani. Considerare la visibilità e l'attrattività delle vetrine durante la pianificazione e il design degli spazi commerciali può

contribuire a creare ambienti urbani più vivaci e accoglienti, che a loro volta possono aumentare l'attrattività e la vitalità di un'area. Inoltre, questo studio suggerisce che modelli analitici come la sintassi dello spazio possono essere utilizzati per comprendere meglio le relazioni tra gli elementi spaziali e il comportamento degli acquirenti, offrendo così un quadro più completo per la progettazione e la gestione degli spazi urbani.

## 3 Dati e oggetto dello studio

### 3.1 Descrizione dei Dati

Il database della camera di commercio di Torino è un ampio archivio di informazioni relative alle attività commerciali autorizzate nell'area urbana della città. Le varie categorie di dati presenti nel database forniscono dettagli su diversi aspetti delle attività commerciali, dai dettagli dell'autorizzazione agli attributi degli indirizzi associati. Possiamo raggruppare le famiglie di variabili in diverse categorie per comprendere meglio il contenuto del database:

- Informazioni sull'autorizzazione e sull'attività commerciale
- Informazioni sull'indirizzo e sulla localizzazione

Queste categorie forniscono una panoramica completa delle informazioni contenute nel database, consentendo agli utenti di analizzare e comprendere meglio il panorama del commercio all'interno dell'area considerata.

Nel corso di questa tesi, è stata data grande rilevanza all'analisi della correlazione tra le attività commerciali presenti nel territorio di Torino e la topografia della città stessa. L'obiettivo è quello di esaminare in che modo la distribuzione spaziale delle attività commerciali sia influenzata dalla configurazione topografica di Torino e come queste dinamiche possano riflettersi nel tessuto economico della città.

Un aspetto cruciale dello studio è stato l'utilizzo del database, che include diversi campi significativi. In particolare, facendo affidamento sull'identificatore unico "ID\_AUTORIZZAZIONI" come chiave di accesso per tracciare e collegare le informazioni specifiche di ciascuna autorizzazione nel dataset, grazie alla presenza tra le variabili di quella che indica la data di chiusura dell'attività commerciale ("DATA\_CHIUSURA") è stato possibile distinguere tra attività commerciali aperte e chiuse, offrendo così la possibilità di analizzare le attività all'interno di un range temporale definito.

La categorizzazione delle attività economiche attraverso il codice "C\_ATECO\_2002" è stata fondamentale per clusterizzare le attività in base alla loro destinazione commerciale. Questa classificazione ha fornito uno strumento potente per individuare tendenze settoriali e comprendere la diversificazione economica della città. Inoltre, la presenza di coordinate geografiche "Coordinata\_X" e "Coordinata\_Y", espresse attraverso il sistema Gauss-Boaga con

Vertice geodetico di Roma Monte Mario, ha introdotto una componente spaziale, consentendo di mappare le attività sulla città e visualizzare la loro distribuzione geografica.

Per contestualizzare questo studio, è stata considerata la topografia di Torino come elemento chiave [FIGURA 2]. L'ipotesi di base è dunque che la disposizione delle attività commerciali sul territorio urbano potrebbe essere influenzata da fattori quali la presenza di zone pianeggianti o collinari, oltre che dalla struttura urbana della città sia dal punto di vista stradale che dalla struttura stessa degli immobili.

Nel corso di questa tesi, è stata approfondita l'analisi dettagliata dei quartieri di Torino, concentrando lo studio sulla tipologia degli immobili e sulle sue implicazioni che questa ha sulla distribuzione delle attività commerciali. Questa approfondita esplorazione ha rappresentato un elemento cruciale per comprendere la connessione tra la struttura urbana e la varietà delle attività economiche nelle diverse aree della città.

A partire da San Salvario, attraverso l'utilizzo di Google Maps, è stata rilevata una presenza significativa di immobili con vetrine lungo le principali arterie stradali. Questa caratteristica ha reso il quartiere particolarmente favorevole a attività commerciali orientate al pubblico, come bar, ristoranti e negozi al dettaglio. L'abbondanza di vetrine ha giocato un ruolo chiave nella visibilità di queste attività, contribuendo alla loro vitalità economica. Considerando il quartiere Crocetta, si nota una combinazione interessante di residenze e spazi commerciali con vetrine, concentrati soprattutto nelle zone centrali. Questa distribuzione ha potenzialmente favorito la presenza di attività di lusso o servizi mirati alla comunità residente, sfruttando l'attrattiva delle vetrine per attirare l'attenzione dei passanti. Il Quadrilatero Romano è caratterizzato da strade strette e edifici storici, queste caratteristiche hanno creato un ambiente ideale per attività commerciali che beneficiano della visibilità offerta dalle vetrine. In caso di assenza delle stesse sarebbe difatti molto complesso individuare le attività a causa della limitata visibilità offerta dall'ampiezza delle strade. Negozi al dettaglio, gallerie d'arte e locali gastronomici sono emersi come componenti vitali di questo quartiere, con la struttura urbana che ha contribuito a definirne l'identità commerciale. L'analisi si è estesa poi a Lingotto, dove la riconversione di aree industriali ha portato a una varietà di immobili con vetrine molto ampie. Questi spazi commerciali, da uffici a negozi, hanno sfruttato la visibilità offerta dalle vetrine per inserirsi con successo nella nuova identità multifunzionale del quartiere. Mirafiori, storicamente legato all'industria automobilistica, ha mostrato una miscela di immobili industriali e residenziali, potenzialmente riflettendo la diversità delle attività commerciali presenti.

Le caratteristiche degli immobili con vetrine hanno quindi svolto un ruolo cruciale nella definizione delle attività commerciali in ogni quartiere di Torino, influenzando la visibilità, l'accessibilità e l'attrattiva di determinate zone per il commercio. Questa analisi fornisce un quadro più completo della connessione tra la struttura urbana e l'economia locale.

L'integrazione di dati spaziali e settoriali, come il codice ATECO, permette di ottenere una visione approfondita della correlazione tra attività commerciali e topografia a Torino. L'obiettivo di questo studio è non solo contribuire alla comprensione della dinamica economica della città, ma fornire anche una base per future analisi e decisioni urbanistiche mirate a sviluppare in modo sostenibile le attività commerciali a Torino.



*Figura 2: Comune di Torino*

Volendo effettuare un'analisi più approfondita ho deciso di esaminare nel dettaglio il quartiere Crocetta della Circoscrizione 1 di Torino, situato a sud rispetto al centro storico cittadino [FIGURA 3].

Storicamente considerata una delle zone residenziali di maggior prestigio, raggiunse il suo massimo sviluppo tra il XIX secolo e gli anni 1930, mantenendo la fama di quartiere di medio-alta borghesia.

Le vie principali del quartiere (Corso Re Umberto, Corso Galileo Ferraris, Corso Alcide de Gasperi, Corso Duca degli Abruzzi, Corso Carlo e Nello Rosselli, Corso Luigi Einaudi, Corso Stati Uniti, Corso Vittorio Emanuele II, Corso Filippo Turati) si contraddistinguono dalle altre, oltre che per ampiezza, per una vasta disponibilità di vetrine d'esposizione che dovrebbero renderle particolarmente attrattive per attività commerciali di dimensioni medio grandi dedite alla vendita di beni. Al contrario le vie interne e meno ampie sono caratterizzate dalla presenza di diversi di dehors e dovrebbero risultare più attrattivi per attività commerciali nel settore della ristorazione o per negozi di dimensioni medio piccole.

Peculiare invece risulta la via Paolo Sacchi su cui insiste la stazione ferroviaria di Porta Nuova e che quindi può offrire un solo lato per le attività commerciali.

A seguito di questa prima analisi superficiale si è deciso quindi di mettere in relazione il database della camera di commercio con una tabella contenente l'analisi topografia delle vie del quartiere Crocetta per comprendere meglio le implicazioni che la struttura degli immobili del quartiere e della loro disposizione hanno sulle scelte commerciali e confermare alcune delle ipotesi sopra citate: in particolare, si intende comprendere se e come la collocazione geografica di varie categorie di attività può essere influenzata dalla configurazione degli edifici e delle strade su cui si trovano.



Figura 3: Quartiere Crocetta

## 3.2 Selezione dei Dati

La fase di selezione dei dati è cruciale per garantire che l'analisi si basi su informazioni affidabili e pertinenti. Di seguito sono descritte le attività relative al controllo delle attività commerciali, alla verifica delle coordinate geografiche e al controllo di consistenza dei dati.

### 3.2.1 Controllo delle attività commerciali: Esclusione di attività chiuse prima del 2020

Per garantire la rilevanza temporale delle attività commerciali considerate nell'analisi sono state considerate attività ancora aperte dal 2020 in poi, è fondamentale quindi escludere tutte quelle che hanno cessato la loro attività prima del 2020. Questo passaggio assicura che i dati riflettano la situazione commerciale attuale. La procedura segue i seguenti passaggi:

- Identificazione delle attività chiuse: Utilizzo di un filtro temporale per individuare le attività commerciali la cui data di chiusura è precedente al 2020.
- Esclusione delle attività chiuse: Creazione di un subset di dati che include solo le attività commerciali attive nel 2020 e successivi.

Per effettuare queste analisi è stato sviluppato un programma python che permette di confrontare la data presente nel database come "DATA\_CHIUSURA" con la data fissa "01/01/2020". Nel caso in cui "DATA\_CHIUSURA" sia antecedente alla data fissa viene effettuato un secondo controllo sul dato "DESC\_STATO\_AUTORIZZAZIONE" e nel caso in cui queste risultino volturate ("VOLTURATE") vengono comunque incluse nel database filtrato.

Per effettuare questa analisi ho inoltre considerato che le attività con DATA\_CHIUSURA impostata a null siano attualmente attive

### 3.2.2 Verifica delle coordinate geografiche: Esclusione dei record senza informazioni geografiche

L'analisi spaziale richiede coordinate geografiche accurate per mappare la distribuzione delle attività commerciali. È essenziale escludere i record senza informazioni geografiche per garantire una rappresentazione corretta e significativa del territorio. La procedura coinvolge i seguenti passaggi:

- Identificazione dei record senza coordinate: Ricerca di record privi di informazioni geografiche.
- Esclusione dei record senza coordinate: Creazione di un sottoinsieme di dati che include solo le attività commerciali con coordinate geografiche valide.

Per effettuare queste analisi è stato sviluppato un programma python che permette di leggere i valori presenti in `Coordinata_X` e `Coordinata_Y`, controllare che non siano impostati a null o non siano numerici ed eventualmente escludere i record che non rispettano questi controlli. Inoltre, sono state escluse tutte le attività che non rientrano nel perimetro del quartiere Crocetta

### 3.2.3 Controllo di consistenza dei dati: Identificazione e correzione di dati anomali o mancanti

La presenza di dati anomali o mancanti può compromettere l'attendibilità dell'analisi. Il controllo di consistenza dei dati mira a individuare e, se possibile, correggere tali problematiche. La procedura include i seguenti passaggi:

- Gestione dei dati mancanti: riferendomi principalmente al codice ATECO, nel caso in cui manchi il dato non è possibile distinguere la tipologia dell'attività commerciale, risulterebbe quindi poco rilevante per la nostra analisi in quanto le attività sono state raggruppate in sei diverse tipologie sulla base del codice ATECO . Questi gruppi sono "convenience store", "shopping goods", "supermarkets", "personal services", "bars and restaurants" e "recreational and entertainment centers".

- **Correzione dei dati anomali:** sempre in riferimento al codice ATECO sono state effettuate alcune correzioni per rendere coerenti i dati facendo riferimento alla classificazione delle attività economiche ATECO 2002.

Questi passaggi nella selezione dei dati assicurano che il dataset sia coerente, privo di informazioni obsolete e pronto per le fasi successive dell'analisi, contribuendo così a una rappresentazione accurata del contesto commerciale nel territorio considerato.

## **4 Metodologia**

### 4.1 Preparazione dei Dati

La fase di preparazione dei dati risulta cruciale per creare un dataset pronto per l'analisi, compreso il trattamento dei dati mancanti e l'arricchimento con variabili aggiuntive.

#### **4.1.1 Creazione di Variabili Aggiuntive**

La creazione di variabili aggiuntive mira ad arricchire il dataset con informazioni rilevanti per l'analisi. Questo processo può includere sia l'estrazione di nuove informazioni che l'aggiunta di indicatori specifici.

##### *4.1.1.1 Estrazione di informazioni: Derivazione di nuove variabili pertinenti per l'analisi*

L'estrazione di informazioni rappresenta un processo cruciale nell'analisi dei dati, in quanto consente di arricchire il dataset con nuove variabili che offrono una comprensione più approfondita del fenomeno in esame. Questo procedimento non è limitato alla semplice estrazione di dati esistenti, ma implica la creazione di nuove variabili che possono fornire ulteriori contestualizzazioni e insight.

Per identificare potenziali variabili da estrarre, è necessaria un'analisi approfondita dei dati disponibili. Ad esempio, nel nostro contesto, potremmo valutare la possibilità di classificare i dati secondo diversi raggruppamenti di codici ATECO, un sistema di classificazione utilizzato per identificare le attività economiche. Questo processo di identificazione richiede un'esplorazione attenta dei dati per individuare le opportunità di estrazione di nuove informazioni che potrebbero essere rilevanti per l'analisi.

Una volta identificate le potenziali variabili da considerare, è necessario applicare tecniche di derivazione per creare le nuove variabili. Nel nostro caso, possiamo utilizzare la classificazione dei codici ATECO come punto di partenza per creare nuove categorie o gruppi di interesse. Attraverso l'incrocio dei dati disponibili con questa classificazione, possiamo identificare sei gruppi di potenziale interesse che potrebbero offrire un'ulteriore comprensione delle dinamiche del fenomeno in esame.

#### *4.1.1.2 Aggiunta di indicatori: Utilizzo di indicatori che possano arricchire l'analisi*

L'integrazione di indicatori rappresenta un passaggio fondamentale nel processo di analisi dei dati, poiché consente di arricchire il dataset con nuove variabili che ampliano la comprensione della distribuzione e della densità delle attività commerciali nel territorio.

In primo luogo, è essenziale selezionare indicatori rilevanti che riflettano gli aspetti specifici del contesto commerciale in esame. Ciò implica un'attenta identificazione di indicatori che possano offrire insight significativi, come la presenza di vetrine o l'estensione delle strade, elementi che possono influenzare la visibilità e l'accessibilità delle attività commerciali.

Successivamente, è necessario procedere con il calcolo degli indicatori selezionati. Questo processo richiede un'analisi dettagliata dei dati disponibili, inclusa un'indagine empirica sulle strade del quartiere. Attraverso questa analisi empirica, è stato determinato se vi fosse una presenza rilevante di vetrine o dehors lungo le strade considerate. Per "presenza rilevante", si è considerata la presenza di vetrine o dehors in almeno il 50% dei numeri civici della via, indicando una distribuzione significativa di attività commerciali lungo la strada.

La preparazione dei dati è stata una fase cruciale, che ha coinvolto la gestione dei dati mancanti e l'aggiunta di variabili pertinenti al dataset. Questo approccio ha contribuito a creare un dataset più robusto e informativo, fornendo una base solida per l'analisi successiva e consentendo una comprensione più approfondita delle dinamiche commerciali nel territorio considerato.

## 4.2 Analisi Spaziale

L'analisi spaziale è fondamentale per comprendere la distribuzione geografica delle attività commerciali e la loro relazione con la topografia del territorio. Questa sezione dettaglia gli step relativi all'utilizzo di sistemi di coordinate, alla mappatura, alla definizione dei parametri topografici e all'analisi statistica della correlazione.

### 4.2.1 Utilizzo di sistemi di coordinate: Assicurarsi che i dati siano coerenti e compatibili con i sistemi di coordinate utilizzati.

L'omogeneità dei sistemi di coordinate è essenziale per garantire che le analisi spaziali siano accurate. Questo passo coinvolge:

- Verifica dei sistemi di coordinate: Esame dei sistemi di coordinate utilizzati nei dati geografici per assicurarsi che siano coerenti e compatibili.
- Conformità dei dati: Trasformazione, se necessario, dei dati in modo che siano tutti nel medesimo sistema di coordinate.

Il sistema di riferimento scelto in questo caso sono le "Coordinate Montemario", un sistema di coordinate geografiche utilizzato principalmente nell'area di Roma, Italia. Il Monte Mario è una collina situata a nord-ovest del centro storico di Roma e rappresenta uno dei punti più alti della città. Le coordinate Montemario sono spesso utilizzate come punto di riferimento per determinare la posizione e l'orientamento di luoghi e strutture all'interno di Roma e delle sue vicinanze. Esse sono espresse in termini di latitudine e longitudine, e forniscono un modo conveniente per identificare specifici luoghi o per calcolare le distanze tra diverse località. Le coordinate geografiche del Monte Mario sono approssimativamente 41°56'N di latitudine e 12°27'E di longitudine. Il sistema di coordinate Montemario si basa sul sistema di riferimento geografico WGS84 (World Geodetic System 1984), che è ampiamente utilizzato a livello globale per rappresentare la superficie terrestre in coordinate geografiche.

#### 4.2.2 Definizione di parametri topografici: Identificazione di indicatori topografici rilevanti (es. elevazione, inclinazione del terreno)

L'identificazione di parametri topografici è cruciale per comprendere l'influenza del terreno sulle attività commerciali.

Questo processo implica:

- Selezione dei parametri topografici: Identificazione di indicatori pertinenti come estensione e presenza di vetrine o dehors, calcolati attraverso l'acquisizione dei dati topografici.
- Acquisizione dei dati topografici: Raccolta o acquisizione di dati topografici da fonti affidabili come il LARTU, che ha fornito l'intera mappa del comune di Torino permettendo così di calcolare l'estensione delle strade, e l'utilizzo di mappe digitali, in questo caso Google Maps per verificare la presenza di dehors e vetrine.

#### 4.2.3 Mappatura: Creazione di mappe che rappresentino la distribuzione delle attività commerciali sul territorio.

La mappatura offre una visualizzazione chiara della distribuzione delle attività commerciali, facilitando l'identificazione di pattern spaziali.

La procedura comprende:

- Selezione delle variabili da mappare: in questo caso sono state scelte sei categorie di strade con determinate caratteristiche, differenti sulla base della loro estensione all'interno del quartiere e della presenza o meno di elementi architettonici come vetrine e dehors. Queste hanno portato ad una classificazione che ha portato ad avere le seguenti tipologie: strade con estensione inferiore a 500m e meno del 50% dei numeri civici con vetrine lungo la loro estensione (a cui è stato assegnato il peso 2), strade con estensione inferiore a 500m e più del 50% dei numeri civici con vetrine lungo la loro estensione (a cui è stato assegnato il peso 3), strade con estensione compresa tra 500m e 1000m e meno del 50% dei numeri civici con vetrine lungo la loro estensione (a cui è

stato assegnato il peso 5), strade con estensione compresa tra 500m e 1000m e più del 50% dei numeri civici con vetrine lungo la loro estensione (a cui è stato assegnato il peso 8), strade con estensione maggiore di 1000m e meno del 50% dei numeri civici con vetrine lungo la loro estensione (a cui è stato assegnato il peso 13), strade con estensione maggiore di 1000m e più del 50% dei numeri civici con vetrine lungo la loro estensione (a cui è stato assegnato il peso 21).

Il 15,41% delle strade nel quartiere Crocetta di Torino, determinato in base alla loro estensione in metri lineari e alla valutazione della presenza o assenza di vetrine lungo il percorso, è classificato nella categoria 2. Allo stesso modo il 19,18% delle strade è stato classificato con peso 3, il 18,53% con peso 5, il 20,03% con peso 8, il 3,66% con peso 13 e il 23,19% con peso 21.

È emerso che non vi è una significativa differenza tra la percentuale delle strade che hanno lunghezze differenti: il 26,85% ha una lunghezza inferiore a 500 metri, il 38,56% è compreso tra 500 e 1000 metri e il 34,58% ha una lunghezza superiore ai 1000 metri. Risulta fondamentale chiarire che, quando si parla di uno di questi valori si intende dire che considerando l'estensione dell'intera rete stradale del quartiere Crocetta in metri lineari la percentuale di strade che può essere categorizzata come sopra descritto risulta essere rispettivamente del 26,85%, 38,56% e 34,58%.

Risulta interessante, inoltre, che solamente il 9,73% delle strade con vetrine ha una lunghezza inferiore a 500 metri.

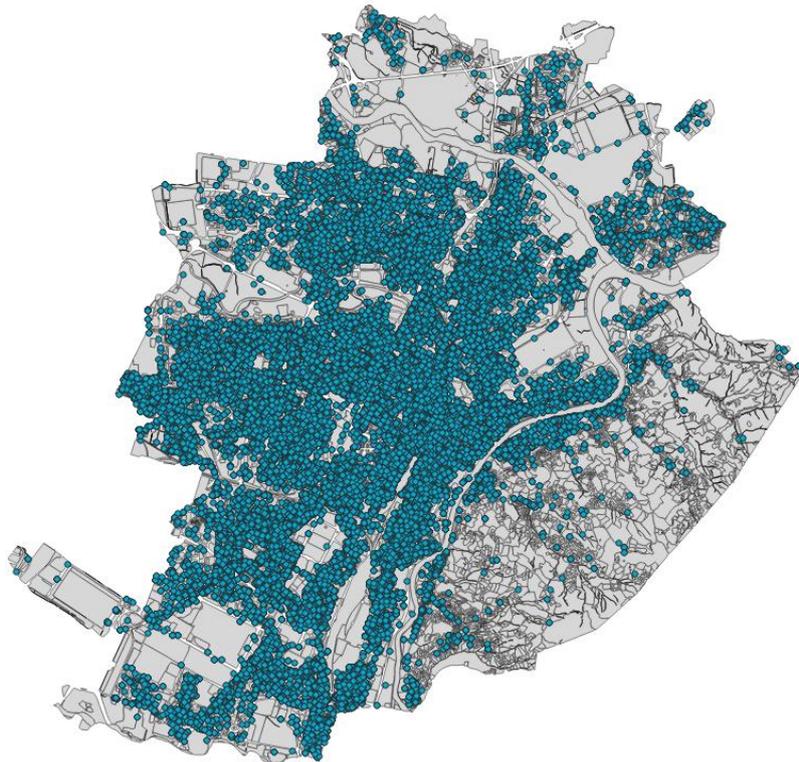
D'altra parte, le strade senza vetrine rappresentano il 62,40% del totale. Tra queste, l'37,17% ha una lunghezza inferiore a 500 metri, mentre il 32,10% ha una lunghezza compresa tra 500 e 1000 metri e il 30,73% superiore ai 1000 metri. L'analisi dei dati conferma quindi che nel quartiere Crocetta di Torino le strade con vetrine sono generalmente più lunghe rispetto a quelle senza vetrine. Questo suggerisce che le strade più brevi tendono a ospitare meno vetrine, probabilmente a causa dello spazio limitato disponibile per le attività commerciali lungo il percorso.

I dati evidenziano una distribuzione asimmetrica della lunghezza delle strade e della presenza di vetrine nel quartiere Crocetta, con una concentrazione relativamente bassa di strade brevi con vetrine.

- Selezione della rappresentazione cartografica: Utilizzo di tecniche di visualizzazione appropriate (in questo caso punti sulla mappa) per evidenziare i pattern spaziali.

Di seguito le rappresentazioni estratte:

- tutte le attività presenti sul territorio di Torino [FIGURA 4]



*Figura 4: Attività presenti sul comune di Torino*

- tutte le attività cui codice ATECO rientra tra quelli scelti per l'analisi [FIGURA 6] ("recreational and entertainment centers", "convenience store", "shopping goods", "personal services", "restaurants and bars" e "supermarkets") all'interno del quartiere di Crocetta sia come punti di interesse che come heatmap [FIGURA 5] per individuare in maniera rapida le zone più rilevanti



Figura 5: Heatmap delle attività presenti nel quartiere Crocetta



Figura 6: Attività presenti nel quartiere Crocetta

- tutte le attività cui codice ATECO rientra nel campo dei "convenience store" all'interno del quartiere di Crocetta [FIGURA 7]



Figura 7: Attività nel campo dei "convenience store" all'interno del quartiere di Crocetta

- tutte le attività cui codice ATECO rientra nel campo dei "recreational and entertainment centers" all'interno del quartiere di Crocetta [FIGURA 8]



Figura 8: Attività nel campo dei "recreational and entertainment centers" all'interno del quartiere di Crocetta

- tutte le attività cui codice ATECO rientra nel campo dei "shopping goods" all'interno del quartiere di Crocetta [FIGURA 9]



Figura 9: Attività nel campo dei "shopping goods" all'interno del quartiere di Crocetta

- tutte le attività cui codice ATECO rientra nel campo dei "personal services" all'interno del quartiere di Crocetta [FIGURA 10]



Figura 10: Attività nel campo dei "personal services" all'interno del quartiere di Crocetta

- tutte le attività cui codice ATECO rientra nel campo dei "restaurants and bars" all'interno del quartiere di Crocetta [FIGURA 11]



Figura 11: Attività nel campo dei "restaurants and bars" all'interno del quartiere di Crocetta

- tutte le attività cui codice ATECO rientra nel campo dei "supermarkets" all'interno del quartiere di Crocetta [FIGURA 12]



Figura 12: Attività nel campo dei "supermarkets" all'interno del quartiere di Crocetta

### 4.3 Analisi delle Attività Commerciali in Base al Codice ATECO e alla Tipologia di Strada

Nel contesto di un'indagine sulle attività commerciali, abbiamo suddiviso tali attività in sei distinti gruppi, utilizzando come criterio di classificazione il loro codice ATECO (Classificazione delle Attività Economiche).

Questi gruppi sono stati definiti come segue:

- **Bars and Restaurants:** Questo primo gruppo comprende locali di ristorazione, caffetterie, bar e altri esercizi simili. Si tratta di luoghi in cui le persone si recano per consumare cibo e bevande.
- **Personal Services:** In questa categoria rientrano attività come parrucchieri, estetisti, centri benessere e altre prestazioni di servizi personali.
- **Convenience store:** I negozi di vicinato costituiscono un importante punto di riferimento per la comunità locale. Qui troviamo piccoli esercizi commerciali come alimentari, tabaccai e farmacie. Questi negozi forniscono beni di uso.
- **Shopping Goods:** Questa categoria comprende i negozi che vendono prodotti di consumo, come abbigliamento, elettronica, libri e articoli per la casa. Sono spazi in cui i clienti possono acquistare beni di vario genere per soddisfare le loro esigenze quotidiane.
- **Supermarkets:** I supermercati sono grandi punti vendita alimentari che offrono una vasta gamma di prodotti. Rappresentano una scelta comune per gli acquisti di generi alimentari e beni di largo consumo.
- **Recreation and Entertainment Centers:** Questa categoria include strutture come cinema, teatri, sale giochi, palestre e altre destinazioni di svago. Questi luoghi offrono intrattenimento e svago per il pubblico e spesso richiedono spazi ampi e attrezzature specifiche.

Successivamente, sono state ulteriormente divise queste attività in base a sei diverse tipologie di strade. Ogni tipologia di strada è stata associata a un peso specifico, che varia da 2 a 21. Questo peso tiene conto dell'estensione della strada su cui l'attività commerciale è ubicata e della presenza o meno di vetrine o dehors. Ad esempio, alle strade principali con alta visibilità è stato assegnato un peso maggiore rispetto a strade secondarie meno frequentate.

L'analisi delle attività commerciali in relazione alle tipologie di strade fornisce informazioni preziose per la pianificazione urbana, lo sviluppo economico e la comprensione della distribuzione geografica delle attività stesse.

#### 4.4 Analisi Statistica

L'analisi statistica permette l'identificazione dei pattern e tendenze nelle distribuzioni spaziali delle attività commerciali in relazione alla topografia del territorio. Ad esempio, potrebbe rilevare se ci sono cluster di attività in determinate zone o se ci sono relazioni significative tra l'altitudine, la pendenza del terreno e la densità di attività commerciali. Allo stesso modo l'analisi statistica può aiutare a determinare se esiste una correlazione tra determinate caratteristiche topografiche e la presenza di attività commerciali. Tuttavia, è importante notare che la correlazione non implica necessariamente una relazione causale, ma può suggerire l'identificazione di variabili chiave che possono influenzare la distribuzione delle attività. Attraverso l'analisi statistica è possibile, inoltre, costruire modelli predittivi che possono stimare la probabilità di presenza di attività commerciali in base alla topografia del territorio e ad altre variabili pertinenti. Questi modelli possono essere utilizzati per fare previsioni e pianificare lo sviluppo commerciale in determinate aree. Questo processo implica un'analisi inferenziale ovvero l'utilizzo di test statistici per identificare eventuali relazioni significative.

- Per valutare la correlazione tra le due variabili e come questa impatti sulla distribuzione geografica delle attività commerciali, è stato adottato un approccio simile al coefficiente di correlazione di Pearson. Le variabili considerate sono il 'gruppo', definito attraverso il codice ATECO che classifica le attività in sei categorie, e il 'peso', che indica l'importanza attribuita a una determinata strada in base a parametri come estensione e presenza di vetrine o dehors. Tuttavia, trattandosi di due variabili categoriche per il 'gruppo' e per il 'peso', è necessario un metodo adeguato per calcolare la correlazione.

Una possibile soluzione è l'utilizzo dell'Analisi della Varianza (ANOVA) per determinare se esiste una relazione significativa tra il 'gruppo' e il 'peso'. L'ANOVA consente di valutare se vi siano differenze significative nelle medie del 'peso' tra i vari 'gruppi'. È importante notare che, sebbene l'ANOVA non fornisca un coefficiente di correlazione, può evidenziare la presenza di differenze significative tra i gruppi.

Per condurre questa analisi, ho impiegato la libreria `scipy` in Python per eseguire l'ANOVA

- Validazione delle correlazioni: Per verificare la validità statistica delle correlazioni identificate e valutare la loro significatività si prende in considerazione il p-value. Il p-value, o valore p, è un parametro statistico utilizzato per valutare la significatività di un risultato in un test di ipotesi. Nel contesto della validazione delle correlazioni, il p-value viene utilizzato per determinare se esiste una differenza significativa tra i pesi medi di gruppi diversi.

Ad esempio, consideriamo il nostro caso in cui abbiamo più gruppi di dati (le attività commerciali suddivise per codice ATECO) e vogliamo verificare se esiste una differenza significativa tra di essi rispetto a una determinata variabile di interesse (il peso attribuito alla strada in cui si trova l'attività commerciale).

Ecco come funziona il processo:

- Definizione delle ipotesi: si inizia formulando un'ipotesi nulla ( $H_0$ ), che generalmente afferma che non esiste alcuna differenza significativa tra i gruppi. L'ipotesi alternativa ( $H_1$ ) sostiene invece che esista una differenza significativa tra almeno due dei gruppi.
- Calcolo del test statistico: utilizzando il test statistico appropriato (ad esempio, test ANOVA per più gruppi o test t per due gruppi), viene calcolato il valore p.
- Interpretazione del p-value: Se il p-value è inferiore a un certo livello di significatività predefinito (tipicamente 0.05), possiamo respingere l'ipotesi nulla e concludere che esiste una differenza significativa nei pesi medi tra almeno due dei gruppi. In altre parole, una bassa probabilità di ottenere un risultato simile sotto l'ipotesi nulla suggerisce che i dati osservati non sono dovuti al caso, ma sono probabilmente il risultato di un effetto reale.
- Considerazioni aggiuntive: è importante notare che il p-value da solo non fornisce informazioni sulla forza o la direzione della relazione tra le variabili, ma indica solo se la differenza osservata è statisticamente significativa. Inoltre, il p-value dipende anche dalla dimensione del campione: campioni più grandi tendono ad avere p-value più piccoli per lo stesso effetto.

In sintesi, il p-value è uno strumento cruciale nell'analisi statistica che fornisce una valutazione della significatività dei risultati dei test di ipotesi, consentendo agli analisti di trarre conclusioni affidabili riguardo alla presenza o all'assenza di differenze significative tra gruppi di dati.

L'analisi spaziale fornisce un quadro dettagliato della distribuzione geografica delle attività commerciali e della loro relazione con la topografia del territorio, contribuendo così a una comprensione più approfondita dei fattori spaziali che possono influenzare il contesto commerciale.

#### 4.5 Verifica dei Risultati

La statistica F è una misura utilizzata in diverse analisi statistiche, tra cui l'ANOVA (Analysis of Variance), per valutare se ci sono differenze significative tra le medie di due o più gruppi.

Nell'ambito dell'ANOVA, la statistica F è calcolata confrontando la variazione tra le medie dei gruppi con la variazione all'interno dei gruppi. In sostanza, misura quanto la variazione nei dati è attribuibile alla differenza tra i gruppi rispetto alla variazione casuale all'interno dei singoli gruppi.

Se la statistica F è grande e il valore p associato è inferiore a un livello di significatività prefissato (comunemente 0.05), si conclude che almeno uno dei gruppi ha una media significativamente diversa dagli altri. In altre parole, ci sono differenze significative tra almeno due dei gruppi.

D'altro canto, se la statistica F è bassa e il valore p è alto, non ci sono sufficienti prove per concludere che ci siano differenze significative tra le medie dei gruppi, e quindi non possiamo rigettare l'ipotesi nulla di uguaglianza delle medie.

In sintesi, la statistica F fornisce un modo per determinare se ci sono differenze significative tra le medie dei gruppi e fornisce una base per decidere se rigettare o accettare l'ipotesi nulla nell'analisi dei dati.

## 4.6 Strumenti e Software Utilizzati

Per effettuare le analisi descritte nella tesi sono stati utilizzati quattro software principali:

- Visual Studio code un editor di codice sorgente leggero, potente e gratuito sviluppato da Microsoft, progettato per essere un ambiente di sviluppo integrato (IDE) altamente personalizzabile attraverso cui sono stati sviluppati due file in codice python, un linguaggio di programmazione ad alto livello, avvalendosi della libreria scipy per eseguire l'ANOVA, delle librerie pandas e geopandas per poter effettuare operazioni sui file Excel contenenti i database della camera di commercio e delle strade analizzate. Nella fattispecie per effettuare l'ANOVA attraverso la libreria spicity è stata utilizzata la funzione la funzione `f_oneway()` che confronta le medie dei pesi (o di qualsiasi altra variabile numerica) tra due gruppi, determinando se le differenze tra le medie sono statisticamente significative. Restituisce due valori: `f_statistic`: il valore della statistica F calcolata dall'ANOVA e `p_value`: il valore p associato al test ANOVA, che indica la probabilità di ottenere un risultato altrettanto estremo o più estremo del risultato osservato, assumendo che l'ipotesi nulla sia vera (ovvero, che non ci siano differenze significative tra le medie dei gruppi). Per quanto riguarda il secondo file di filtraggio sono state invece utilizzate le operazioni di filtraggio sulla base della data di chiusura, del codice ATECO e delle coordinate riuscendo così ad eliminare i record con dati parziali o assenti.
- Microsoft Access, un'applicazione software di database relazionale sviluppata da Microsoft che consente agli utenti di creare e gestire database per archiviare, organizzare e recuperare dati in modo efficiente grazie alla quale si è riusciti a creare un database unico con tre diverse tabelle per riuscire a mettere in relazione i dati della camera di commercio di Torino, la classificazione delle strade e la classificazione dei codici ATECO suddivisi in gruppi. Le tre tabelle sono state messe in relazione tra di loro assegnando a un campo per tabella la proprietà di primary key, un identificativo univoco attraverso cui era possibile accedere al record ad esso legato. La primary key tabella contenente le informazioni della camera di commercio è rappresentata dall'identificativo univoco dell'attività commerciale, la primary key della tabella contenente le informazioni delle strade è determinata attraverso la denominazione corrente delle strade che è presente anche all'interno della tabella delle attività commerciali, infine la primary key della tabella dei codici ATECO è il codice ATECO stesso

in quanto univoco e presente nella tabella delle attività commerciali presenti sul territorio di Torino.

- Microsoft Excel, utilizzato principalmente per effettuare le analisi statistiche più basilari come la frequenza ed elaborare i relativi grafici attraverso la creazione di tabelle pivot.
- QGIS, acronimo di "Quantum GIS", che è un software open source per il Geographic Information System (GIS). È progettato per la creazione, la modifica, la visualizzazione, l'analisi e la pubblicazione di dati geospaziali fondamentali nel contesto di questa ricerca attraverso il quale è stato possibile estrarre i grafici spaziali sopra disponibili e calcolare alcune variabili come l'estensione delle strade del quartiere analizzato.

Oltre a questi principali software ho fatto riferimento anche a testi di letteratura per individuare quale test statistico fosse maggiormente indicato per la tipologia di dati a mia disposizione e mi sono avvalso dell'utilizzo di Google Maps per confermare la validità di alcuni dati e correggerne altri.

Inoltre, per quanto riguarda i database ho utilizzato il database nominato "commercio\_per\_politecnico\_dicembre\_2022" all'interno del quale erano contenute la maggior parte dei dati utilizzati e che costituiva la base di partenza della mia analisi e il database denominato "DB\_TOPOGRAFICO\_TO" fornitomi dal professor Antonio Cittadino per conto del LARTU - Laboratorio di Analisi e Rappresentazioni Territoriali ed Urbane.

## 5 Risultati

	bars and restaurants	personal services	convenience	shopping	supermarket	Recreational and Entertainment Centers
bars and restaurants		F: 1.1856 p: 0.2764	F: 9.6011 p: 0.0020	F: 0.5004 p: 0.4795	F: 0.1661 p: 0.6837	F: 7.0963 p: 0.0079
personal services			F: 2.3498 p: 0.1254	F: 0.1079 p: 0.7426	F: 0.6676 p: 0.4142	F: 5.8759 p: 0.0157
convenience				F: 3.4418 p: 0.0637	F: 1.6918 p: 0.1935	F: 4.0106 p: 0.0454
shopping					F: 0.4682 p: 0.4942	F: 6.3595 p: 0.0120
supermarket						F: 5.6449 p: 0.0202
Recreational and Entertainment Centers						

Figura 13: Tabella p-value e statistica F

Nella tabella sopra riportata [FIGURA 13] sono stati riportati i risultati dell'analisi statistica effettuata: il valore della statistica F calcolata dall'ANOVA e il valore p associato al test ANOVA. Risulta facile notare come i dati a nostra disposizione mostrano una significativa differenza nella distribuzione sul territorio delle attività commerciali tra le attività appartenenti al gruppo "Bars and Restaurants" e quelle appartenenti al gruppo "Convenience", evidenziati in verde nella tabella.

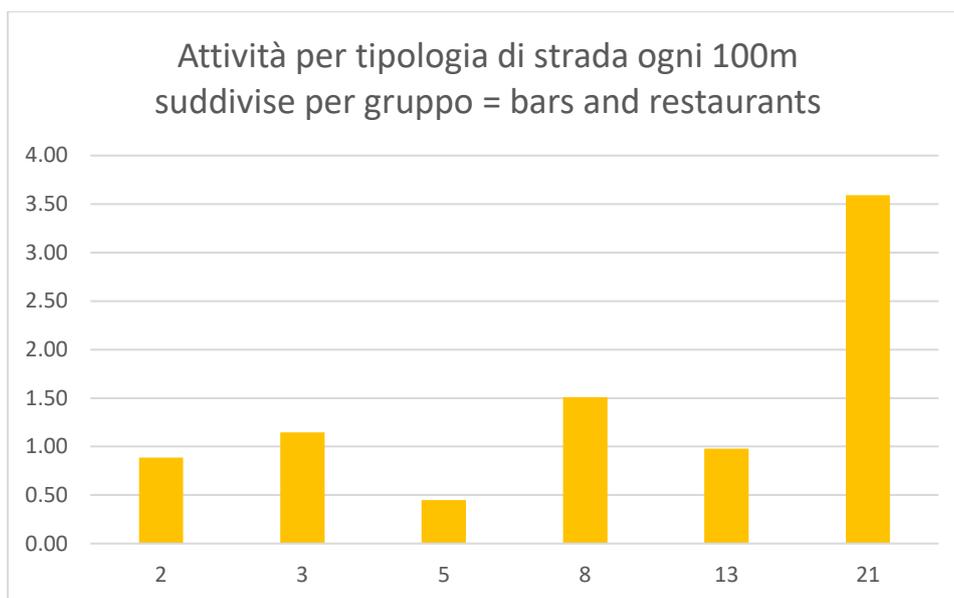


Figura 14: Grafico a barre di attività appartenenti al gruppo "bars and restaurants" suddivisi per tipologia di strada

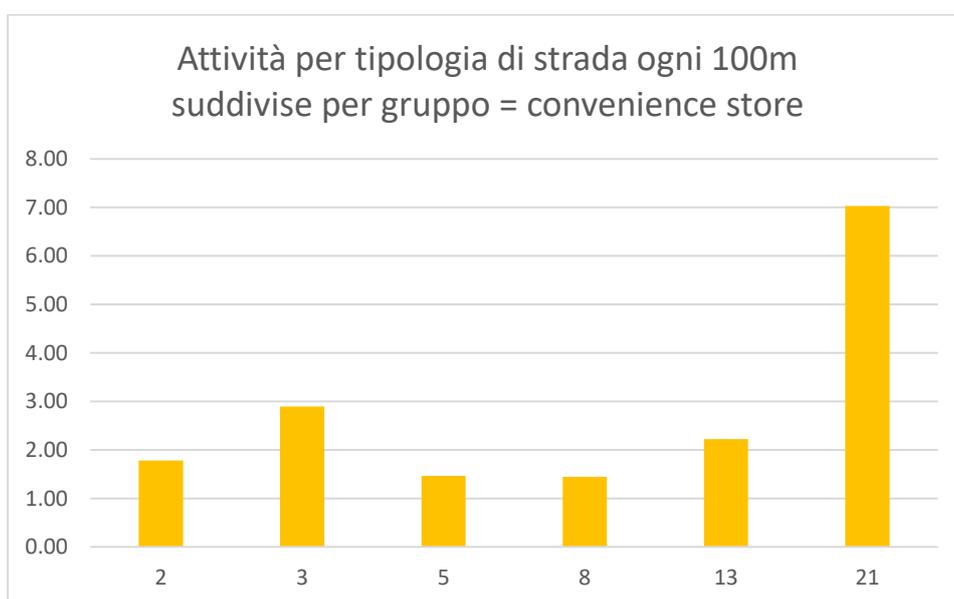


Figura 15: Grafico a barre di attività appartenenti al gruppo "convenience store" suddivisi per tipologia di strada

Le tabelle qui sopra [FIGURA 14][FIGURA 15] mostrano quante attività commerciali si trovano in media su una strada suddiviso per tipologia di strada. I valori riportati nelle tabelle sono stati normalizzati considerando la presenza delle attività ogni 100 metri lineari di strada. La principale differenza che emerge risulta essere riguardo le preferenze sulle strade con estensione compresa tra i 500 e i 1000 metri.

Le attività commerciali appartenenti al gruppo convenience store infatti non mostrano alcuna preferenza significativa tra strade con prevalenza di vetrine (peso 8) e non (peso 5), in effetti è possibile prevedere questo comportamento dalla tipologia delle attività commerciali stesse: un convenience store, o negozio di prossimità, è un tipo di negozio che offre una vasta gamma di prodotti di consumo di base, generalmente situato in aree residenziali o adiacenti ad aree ad alto traffico o vicino a luoghi di lavoro. La caratteristica distintiva di un convenience shop è la sua comodità e accessibilità, in quanto mira a soddisfare le esigenze immediate dei clienti offrendo una selezione di prodotti di uso quotidiano, come generi alimentari, bevande, articoli per la casa, prodotti per la cura personale e articoli da toeletta. Questi negozi sono progettati per essere rapidi e facili da visitare, consentendo ai clienti di fare acquisti rapidi e di ritirare gli articoli di cui hanno bisogno senza dover affrontare lunghe code o spostarsi troppo lontano dalla loro posizione attuale.

Al contrario per le attività commerciali appartenenti al gruppo bars and restaurants è possibile individuare una netta preferenza a favore delle strade con presenza di vetrine o dehors probabilmente perché le vetrine permettono ai clienti di vedere il numero di clienti già all'interno, infatti il numero di avventori può riflettere in modo approssimativo la reputazione del ristorante come evidenziato nel caso Starbucks (Youngme Moon e John Quelch, 2006), e i dehors offrono la possibilità di ampliare il numero di coperti a disposizione delle attività.



*Figura 16: Grafico a barre di attività appartenenti al gruppo "recreational and entertainment centers" suddivisi per tipologia di strada*

Risulta inoltre peculiare il comportamento delle attività appartenenti al gruppo "recreational and entertainment centers" [FIGURA 16]: è comune che le attività di intrattenimento e ricreative si concentrino in determinate zone della città, come il centro storico o i quartieri più vivaci e frequentati. Queste aree attraggono spesso un maggior numero di residenti e turisti, offrendo un ambiente più dinamico e vivace per le attività di intrattenimento.

Risulta quindi molto probabile che il quartiere Crocetta, preso in esame, non sia particolarmente adatto allo sviluppo di questa tipologia di attività come si nota dalla presenza di 36 sole attività (1,0676% sul totale delle attività del quartiere) senza alcuna preferenza riguardo le diverse tipologie di strada. Potrebbe essere dunque interessante effettuare lo stesso tipo di analisi su diversi quartieri per indagarne il comportamento.

In generale è comunque visibile una correlazione positiva tra gli immobili che offrono spazi come vetrine e dehors e la presenza di attività commerciali indicando questa caratteristica come particolarmente rilevante nell'avviamento di attività commerciali all'interno del quartiere.

## 6 Conclusioni

In questo lavoro di ricerca, è stata dunque esaminata attentamente la distribuzione delle attività commerciali all'interno del quartiere Crocetta, ponendo particolare attenzione su sei categorie principali: bar e ristoranti, negozi di prossimità, shopping goods, supermercati, servizi alla persona e centri di intrattenimento e ricreativi. Attraverso l'analisi dei dati raccolti e delle tendenze emerse, siamo giunti a diverse conclusioni significative che offrono spunti interessanti per comprendere la dinamica commerciale all'interno di questo contesto urbano.

Innanzitutto, è stata osservata una marcata differenza nella distribuzione delle attività tra i vari gruppi esaminati. Soprattutto le attività appartenenti al settore dei bar e ristoranti mostrano una chiara preferenza per le strade con vetrine o dehors, probabilmente a causa della visibilità e dell'attrattiva che questi elementi offrono alla clientela. D'altra parte, i negozi di convenienza non presentano una preferenza significativa tra le diverse tipologie di strade, indicando una diversa strategia di posizionamento e attrazione della clientela.

Un elemento peculiare che emerge dall'analisi riguarda i centri di intrattenimento e ricreativi, i quali tendono a concentrarsi in determinate zone della città, come il centro storico o i quartieri più vivaci. Tuttavia, nel caso del quartiere Crocetta, abbiamo notato una presenza limitata di queste attività, suggerendo che il contesto urbano potrebbe non essere particolarmente favorevole allo sviluppo di questa tipologia di business. Questa osservazione apre la porta a ulteriori indagini su altri quartieri per comprendere meglio le dinamiche di distribuzione e le preferenze delle attività commerciali.

L'analisi condotta ha evidenziato l'importanza della visibilità e dell'attrattiva estetica delle attività commerciali, nonché l'influenza del contesto urbano e della tipologia di clientela che ogni zona può attrarre. Tali risultati forniscono preziosi spunti per gli imprenditori e gli urbanisti interessati allo sviluppo e alla pianificazione del commercio locale.

Tuttavia, è importante sottolineare che questo studio presenta alcune limitazioni, tra cui la focalizzazione su un singolo quartiere e la mancanza di considerazione di altri fattori socioeconomici che potrebbero influenzare la distribuzione delle attività commerciali. Pertanto, future ricerche potrebbero approfondire questi aspetti e ampliare l'analisi ad altre aree urbane per ottenere una visione più completa e approfondita del fenomeno.

In conclusione, questo lavoro fornisce un contributo significativo alla comprensione della distribuzione delle attività commerciali all'interno del quartiere Crocetta e offre importanti

spunti per la progettazione di politiche urbane e strategie di sviluppo economico locale. La ricerca futura potrebbe approfondire ulteriormente le dinamiche osservate ed esplorare nuove prospettive per favorire un ambiente commerciale vibrante e dinamico all'interno delle città.

Il presente studio offre inoltre una solida base di conoscenze e informazioni che possono essere sfruttate per il posizionamento strategico e l'avviamento di nuove attività commerciali nel quartiere Crocetta e in contesti urbani simili. Le conclusioni e le tendenze emerse forniscono agli imprenditori e agli investitori una guida preziosa per prendere decisioni informate e ottimizzare il successo delle loro iniziative imprenditoriali. Di seguito sono riportati alcuni modi in cui questo studio può essere utilizzato per il posizionamento in fase di avviamento:

- **Scelta della Posizione:** Le analisi condotte indicano che le attività commerciali nel settore dei bar e ristoranti mostrano una chiara preferenza per le strade con vetrine o dehors. Pertanto, gli imprenditori che intendono avviare nuovi locali gastronomici potrebbero concentrarsi sulla ricerca di locali situati lungo tali arterie stradali per massimizzare la visibilità e l'attrattiva per i clienti.
- **Pianificazione dello Spazio e del Design:** L'importanza della visibilità e dell'attrattiva estetica delle attività commerciali sottolineata nello studio suggerisce che gli imprenditori dovrebbero prestare particolare attenzione alla progettazione e al layout del loro locale. Investire in vetrine ben curate, dehors accoglienti e un design interno accattivante potrebbe contribuire a catturare l'interesse dei potenziali clienti e a differenziarsi dalla concorrenza.
- **Analisi della Concorrenza:** Lo studio evidenzia anche la presenza limitata di centri di intrattenimento e ricreazione nel quartiere Crocetta. Questo potrebbe rappresentare un'opportunità per gli imprenditori interessati a questo settore di trovare una nicchia di mercato ancora poco sfruttata e adattare la propria offerta di intrattenimento alle esigenze e ai gusti della comunità locale. Risulta importante, tuttavia, effettuare ulteriori analisi per comprendere il motivo per cui questa tipologia di attività non è ancora presente all'interno del quartiere Crocetta.
- **Studio di Mercato:** Gli imprenditori possono utilizzare i dati e le tendenze emerse dalla ricerca per condurre uno studio di mercato più approfondito e comprendere meglio le preferenze e i comportamenti dei potenziali clienti nel quartiere Crocetta. Questo aiuterà a identificare opportunità di business non ancora sfruttate e a sviluppare strategie di marketing mirate per raggiungere il proprio pubblico di riferimento.

In sintesi, l'analisi condotta fornisce agli imprenditori una panoramica dettagliata delle dinamiche commerciali nel quartiere Crocetta, consentendo loro di prendere decisioni più informate e strategiche durante la fase di avviamento delle loro attività. Utilizzando queste informazioni in modo efficace, gli imprenditori possono massimizzare le probabilità di successo e contribuire alla vitalità economica e sociale della comunità locale.

## 7 Bibliografia

1. Hotelling (1929) - "Stability in Competition"
2. Huff (1964) - "Defining and Estimating a Trade Area".
3. Applebaum (1966) - "A Note on the Retail Location Model"
4. Buzzacchi et al. (2020) – “Using betweenness metrics to investigate the geographical distribution of retailers”.
5. Omer (2005) - "Location, Location, Location: The Co-location of Foreign Direct Investment in the U.S."
6. Kang (2015) - "The Role of Geographic Proximity in FDI Productivity Spillovers in China"
7. Porta e Scheurer (2006) – “Centrality Measures in Spatial Networks of Urban Streets”
8. Porta et al. (2009) – “Street centrality and densities of retail and services in Bologna, Italy”
9. Sevtsuk (2014) – “Location and agglomeration: The distribution of retail and food businesses in dense urban environments. Journal of Planning Education and Research.”
10. Turhan et al. (2013) – “Literature Review on Selection Criteria of Store Location Based on Performance Measures. Procedia – Social and Behavioral Sciences”
11. Kotler e Armstrong (2016) - "Principles of Marketing"
12. Berry e Waldfogel (2001) - "Product Quality and Market Size"
13. Karande & Lombard (2005) - "Effects of Store Image and Attitude toward Shelf Space on Purchase Behaviour"
14. Dubé et al. (2016) - "Effects of the Minimum Legal Drinking Age on Youth Substance Use and Traffic Injuries”.
15. Porter (1980) - "Competitive Strategy: Techniques for Analysing Industries and Competitors”.
16. Zentes et al. (2007) - "Strategic Retail Management: Text and International Cases"
17. Hoch et al. (1995) - "Store Brand Image and Store Brand Purchase Intention: Moderating Role of Perceived Risk"
18. Fatima et al. (2017) - “Retailing: An emerging Trend in India- Literature Review”

19. Glaeser et al. (2000) - "Consumer City"
20. Dunne e Lusch (2008) - "Retailing (2nd ed.)"
21. Kotler e Keller (2016) - "Marketing Management"
22. Fox et al. (2007) - "Quantifying Agglomeration Spillovers: Evidence from Million Dollar Plants"
23. Timmermans (1986) - "Shopping: From a Comparative Historical Perspective"
24. Green e Malpezzi (2003) - "A Primer on U.S. Housing Markets and Housing Policy"
25. Moon e Quelch (2006) – "Starbucks: Delivering Customer Service"
26. Lian e Li (2023) – "Correlation Analysis of Retail Space and Shopping Behaviour in a Commercial Street Based on Space Syntax: A Case of Shijiazhuang, China"
27. Mehta (2006) – "Lively Streets: Exploring the Relationship between Built Environment and Social Behaviour"

## 8 Indice delle figure

Figura 1: Driver che influenzano la scelta sulla base delle store performances.....	10
Figura 2: Comune di Torino.....	26
Figura 3: Quartiere Crocetta .....	27
Figura 4: Attività presenti sul comune di Torino .....	35
Figura 5: Heatmap delle attività presenti nel quartiere Crocetta .....	36
Figura 6: Attività presenti nel quartiere Crocetta .....	36
Figura 7: Attività nel campo dei "convenience store" all'interno del quartiere di Crocetta.....	37
Figura 8: Attività nel campo dei "recreational and entertainment centers" all'interno del quartiere di Crocetta.....	37
Figura 9: Attività nel campo dei "shopping goods" all'interno del quartiere di Crocetta.....	38
Figura 10: Attività nel campo dei "personal services" all'interno del quartiere di Crocetta.....	38
Figura 11: Attività nel campo dei "restaurants and bars" all'interno del quartiere di Crocetta .....	39
Figura 12: Attività nel campo dei "supermarkets" all'interno del quartiere di Crocetta.....	39
Figura 13: Tabella p-value e statistica F .....	46
Figura 14: Grafico a barre di attività appartenenti al gruppo "bars and restaurants" suddivisi per tipologia di strada .....	47
Figura 15: Grafico a barre di attività appartenenti al gruppo "convenience store" suddivisi per tipologia di strada.....	47
Figura 16: Grafico a barre di attività appartenenti al gruppo "recreational and entertainment centers" suddivisi per tipologia di strada .....	49
Figura 17: Grafico totale attività suddivise per categoria ATECO .....	59
Figura 18: Grafico totale attività suddivise per categoria ATECO e tipologia di strada.....	60
Figura 19: Grafico totale attività suddivise per tipologia di strada e categoria ATECO.....	60
Figura 20: Grafico a barre di attività appartenenti alla tipologia di strada 2 suddivisi per gruppo ATECO .....	61
Figura 21: Grafico a barre di attività appartenenti alla tipologia di strada 3 suddivisi per gruppo ATECO .....	61
Figura 22: Grafico a barre di attività appartenenti alla tipologia di strada 5 suddivisi per gruppo ATECO .....	62
Figura 23: Grafico a barre di attività appartenenti alla tipologia di strada 8 suddivisi per gruppo ATECO .....	62

Figura 24: Grafico a barre di attività appartenenti alla tipologia di strada 13 suddivisi per gruppo ATECO.....	63
Figura 25: Grafico a barre di attività appartenenti alla tipologia di strada 21 suddivisi per gruppo ATECO.....	63
Figura 26: Grafico a barre di attività appartenenti al gruppo "personal services" suddivisi per tipologia di strada .....	64
Figura 27: Grafico a barre di attività appartenenti al gruppo "shopping goods" suddivisi per tipologia di strada .....	64
Figura 28: Grafico a barre di attività appartenenti al gruppo "supermarket" suddivisi per tipologia di strada .....	65
Figura 29: Tabella riassuntiva attività commerciali ogni 100m suddivisi per categoria ATECO e tipologia di strada .....	65
Figura 30: Diagramma ER.....	66

## 9 Appendice

### 9.1 Database

Il Database denominato “commercio\_per\_politecnico\_dicembre\_2022” su cui si basa l’analisi effettuata è composto da 166.523 record e 32 categorie di dati diversi così definite:

1. ID\_AUTORIZZAZIONI: Un identificatore unico per le autorizzazioni o i record all'interno del database.
2. COD\_COMPARTO: Il codice associato a un settore o a una divisione specifica all'interno del contesto dell'autorizzazione.
3. DATA\_INIZIO: La data in cui inizia l'attività commerciale.
4. DATA\_CHIUSURA: La data in cui l'attività commerciale si chiude o viene volturata.
5. ID\_COD\_FISC: Il codice fiscale associato all'attività commerciale o al record.
6. C\_ATECO\_2002: Un codice che rappresenta la classificazione dell'attività economica dell'azienda o dell'entità associata. Questo codice permette di classificare le aziende in 17 macroaree diverse. Il codice è composto da un numero variabile tra 2 e 6: le prime due cifre identificano un’area di interesse, le successive 4 specificano ancora di più il settore in cui opera l’attività commerciale
7. DESCRIZIONE\_2002: La descrizione dell'attività economica basata sul codice C\_ATECO\_2002.
8. DESC\_DOTAZIONE: Descrizione delle risorse o delle dotazioni legate all'autorizzazione.
9. DOTAZIONE, DOTAZIONE1: Ulteriori informazioni sulle dotazioni o risorse.
10. DESC\_RIPARTIZIONI\_COMPARTI: Descrizione della suddivisione dei comparti o delle divisioni.
11. DESC\_DETTAGLI\_RIPARTIZIONI: Descrizione dettagliata della suddivisione dei comparti.
12. TIPOLOGIA\_MERC: Il tipo di merci o prodotti associati all'autorizzazione.
13. DESC\_CATEGORIA\_MERCEOLOGICA: La descrizione della categoria merceologica associata alle merci o ai prodotti.
14. DESC\_STATO\_AUTORIZZAZIONE: La descrizione dello stato dell'autorizzazione.

15. ID\_CIVICO: L'identificatore del civico dell'indirizzo associato all'autorizzazione.
16. CODICE\_CIRCOSCRIZIONE: Il codice della circoscrizione associato all'indirizzo.
17. CODICE\_QUARTIERE: Il codice del quartiere associato all'indirizzo.
18. DESC\_ZONA\_STATISTICA: La descrizione della zona statistica dell'indirizzo.
19. CODICE\_SEZIONE\_CENSIMENTO: Il codice della sezione di censimento associata all'indirizzo.
20. CODICE\_VIA: Il codice della via associato all'indirizzo.
21. NOME\_CORRENTE: Il nome corrente dell'indirizzo.
22. DENOMINAZIONE\_CORRENTE: La denominazione corrente dell'indirizzo.
23. NUMERO\_RADICE: Il numero civico principale dell'indirizzo.
24. BIS\_TER, INTERNO\_1, BIS\_INTERNO\_1, INTERNO\_2, BIS\_INTERNO\_2: Ulteriori informazioni sull'indirizzo, come eventuali bis, interni, ecc.
25. SCALA: Informazioni sulla scala o sul piano dell'edificio associato.
26. SECONDARIO: Dettagli aggiuntivi sull'indirizzo, come unità secondarie o appuntamenti.
27. FLAG\_PASSO\_CARRAIO: Un indicatore che può rappresentare se c'è un passo carraio associato.
28. CAP: Il codice di avviamento postale (CAP) associato all'indirizzo.
29. DESCRIZIONE\_STATO\_CIVICO: La descrizione dello stato del civico (ad esempio, attivo, inattivo, ecc.).
30. DESCRIZIONE\_TIPOLOGIA\_CIVICO: La descrizione della tipologia del civico (ad esempio, residenziale, commerciale, ecc.).
31. Coordinata\_X: Coordinate geografiche associate all'indirizzo.
32. Coordinata\_Y: Coordinate geografiche associate all'indirizzo.

## 9.2 Tabelle

Di Seguito una serie di tabelle che meglio rappresentano la distribuzione delle attività commerciali in relazione alla loro posizione geografica e al loro gruppo:

Qui è possibile notare come sia suddivise le attività commerciali prese in esame, in relazione anche alle caratteristiche proprie di ogni gruppo e della tipologia di strada.

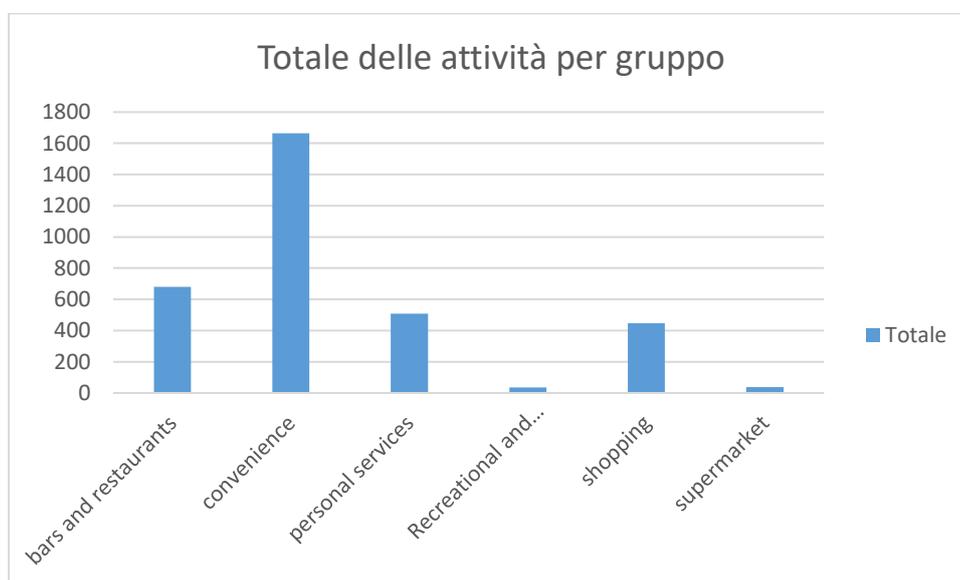


Figura 17: Grafico totale attività suddivise per categoria ATECO

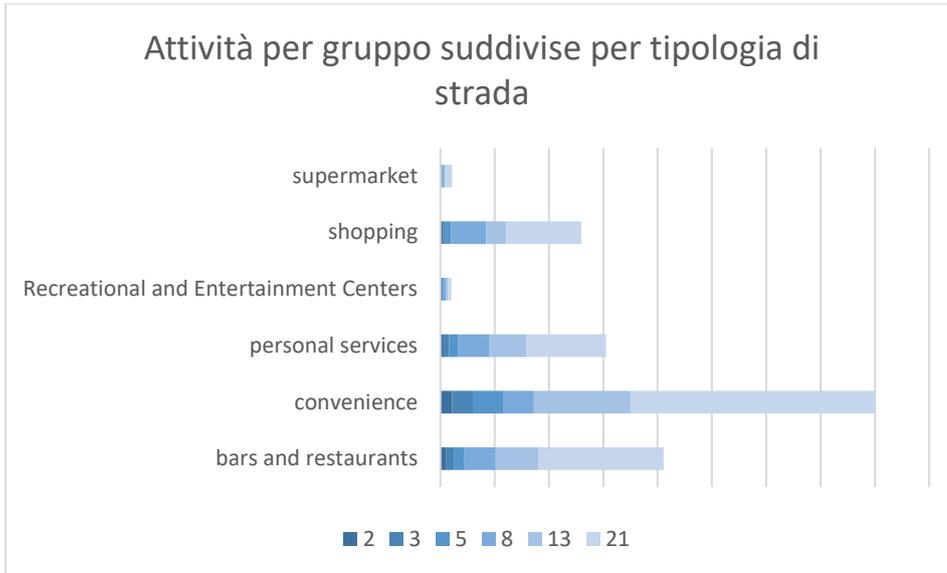


Figura 18: Grafico totale attività suddivise per categoria ATECO e tipologia di strada

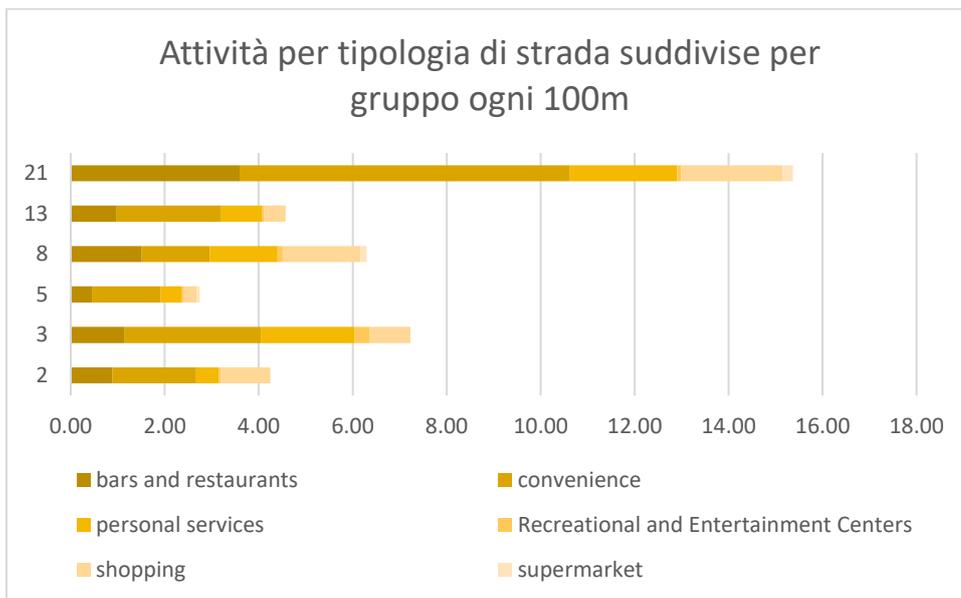


Figura 19: Grafico totale attività suddivise per tipologia di strada e categoria ATECO

Le seguenti tabelle mostrano quante attività commerciali si possono trovare su una determinata tipologia di strada suddivise per gruppo di appartenenza e normalizzati: difatti definisci quante attività per categoria si trovano in media su una determinata tipologia di strada

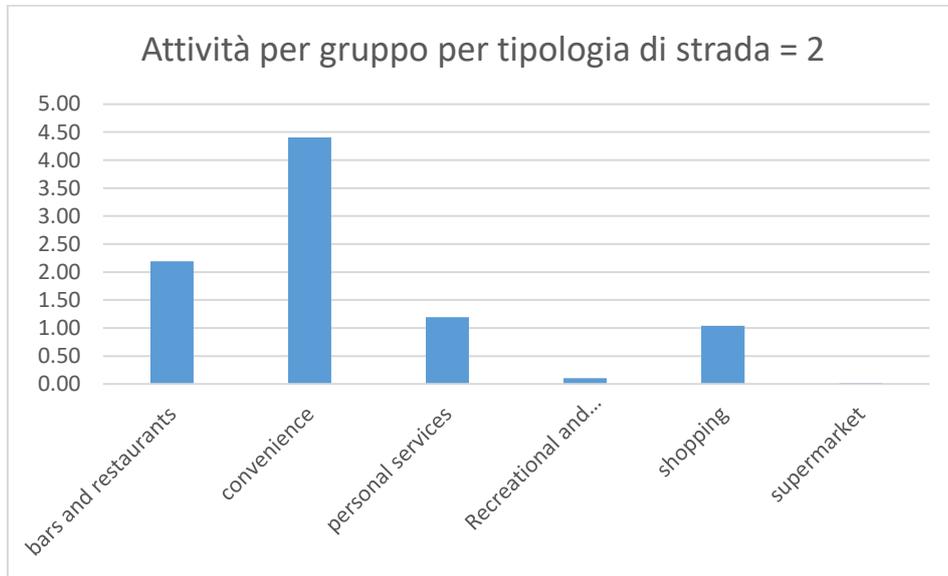


Figura 20: Grafico a barre di attività appartenenti alla tipologia di strada 2 suddivisi per gruppo ATECO

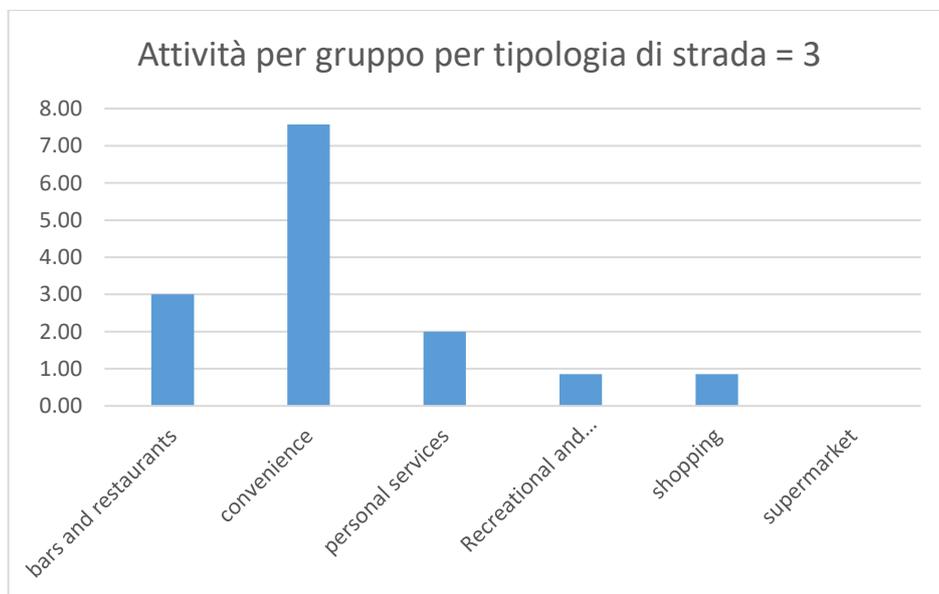


Figura 21: Grafico a barre di attività appartenenti alla tipologia di strada 3 suddivisi per gruppo ATECO

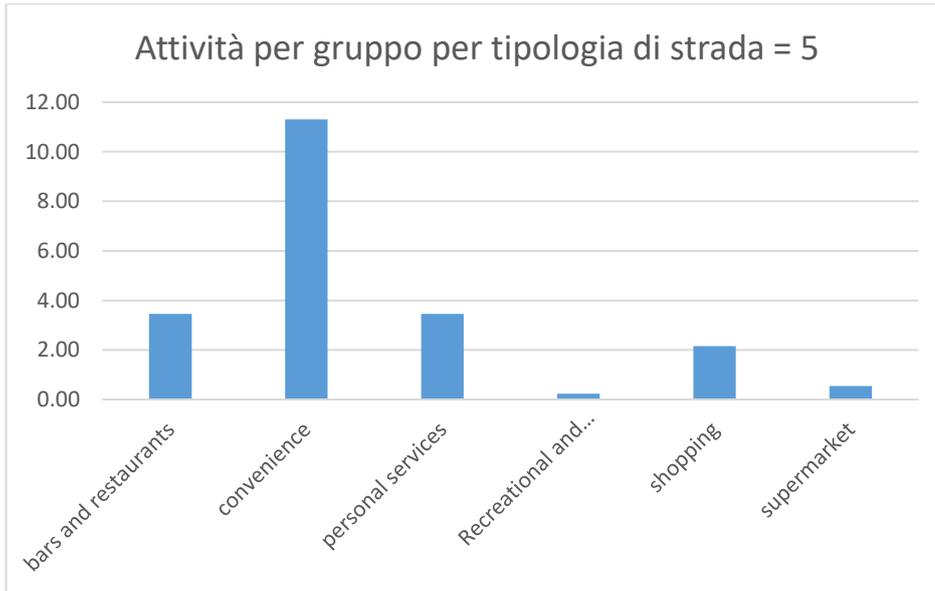


Figura 22: Grafico a barre di attività appartenenti alla tipologia di strada 5 suddivisi per gruppo ATECO

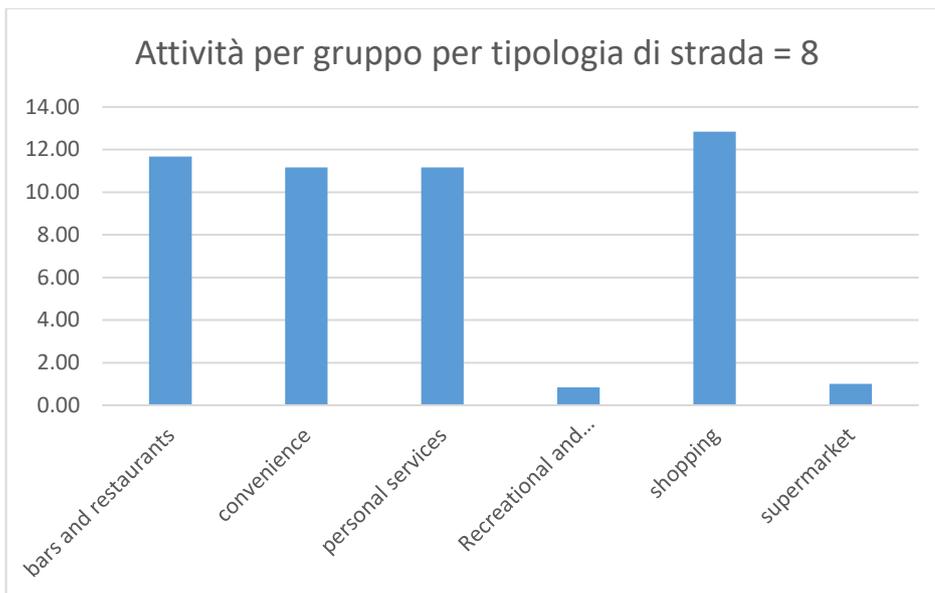


Figura 23: Grafico a barre di attività appartenenti alla tipologia di strada 8 suddivisi per gruppo ATECO

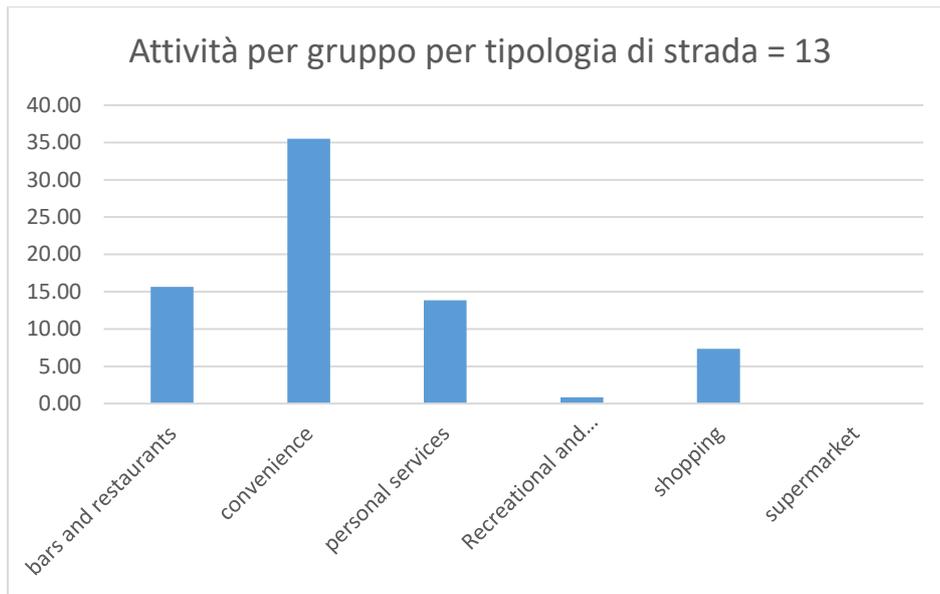


Figura 24: Grafico a barre di attività appartenenti alla tipologia di strada 13 suddivisi per gruppo ATECO

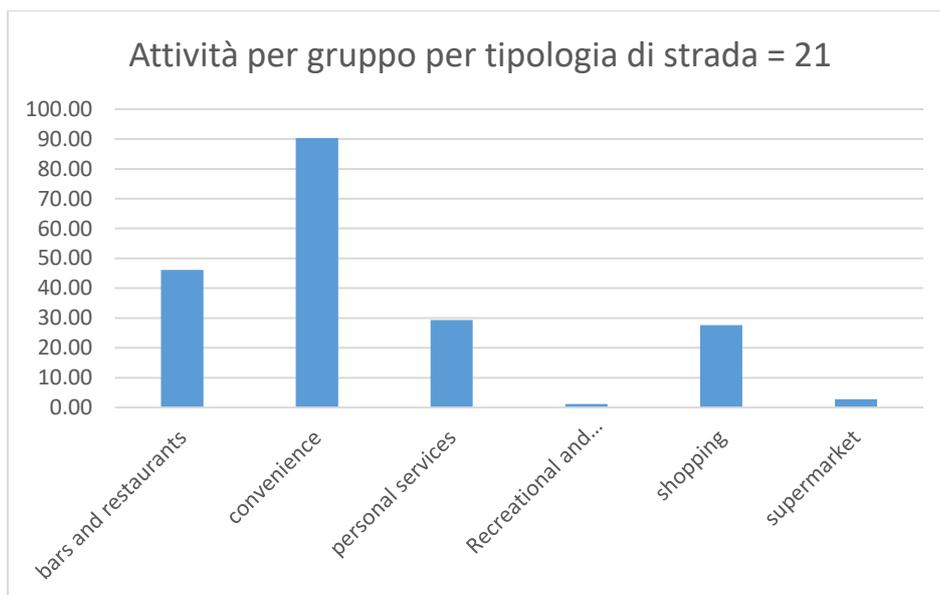


Figura 25: Grafico a barre di attività appartenenti alla tipologia di strada 21 suddivisi per gruppo ATECO

Le seguenti tabelle mostrano su quale tipologia di strada è più facile trovare le attività commerciali suddivise per gruppo di appartenenza

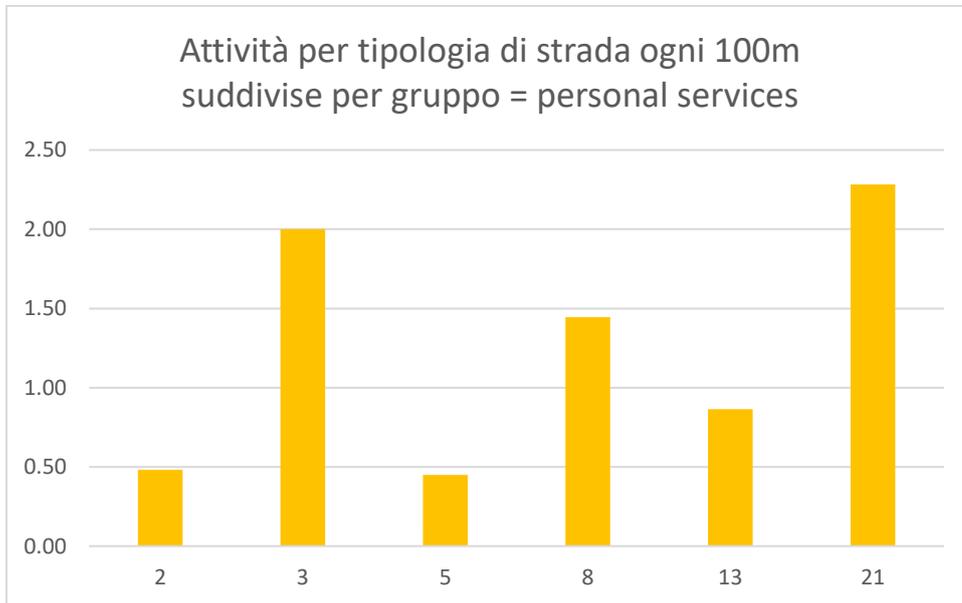


Figura 26: Grafico a barre di attività appartenenti al gruppo "personal services" suddivisi per tipologia di strada



Figura 27: Grafico a barre di attività appartenenti al gruppo "shopping goods" suddivisi per tipologia di strada

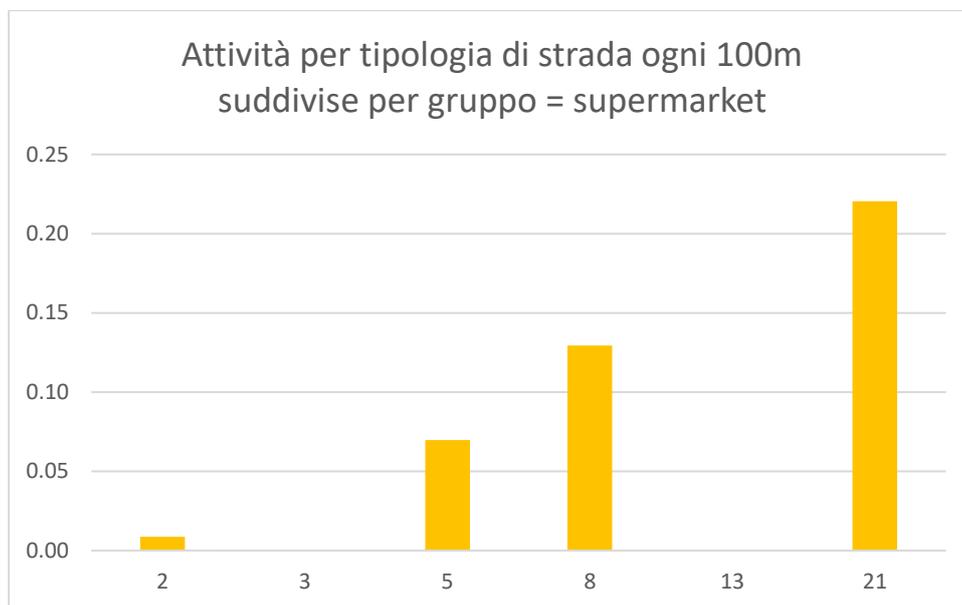


Figura 28: Grafico a barre di attività appartenenti al gruppo "supermarket" suddivisi per tipologia di strada

	2	3	5	8	13	21
bars and restaurants	0.89	1.15	0.45	1.51	0.98	3.59
convenience	1.78	2.89	1.47	1.45	2.22	7.03
personal services	0.48	2.00	0.45	1.45	0.86	2.28
Recreational and Entertainment Centers	0.04	0.33	0.03	0.11	0.05	0.09
shopping	1.04	0.86	0.28	1.66	0.46	2.15
supermarket	0.01	0.00	0.07	0.13	0.00	0.22

Figura 29: Tabella riassuntiva attività commerciali ogni 100m suddivisi per categoria ATECO e tipologia di strada

### 9.3 Diagramma ER e query

Rappresentazione grafica del database, composto da tre tabelle. Dal grafico è possibile individuare in "ID", "C\_ATECO\_2002" e "DENOMINAZIONE CORRENTE" le chiavi univoche di accesso alle tabelle per costruire le query per l'estrazione dei dati utili all'analisi.



Figura 30: Diagramma ER

Query per l'estrazione dei dati

```
SELECT c.ID, c.C_ATECO_2002, c.DATA_CHIUSURA, r.GROUP AS gruppo, cs.Peso,
c.Coordinata_X, c.Coordinata_Y, r.F2
FROM
  (commercio_per_politecnico_2022_ AS c
  INNER JOIN raggruppazioni_ateco AS r ON c.C_ATECO_2002 = r.C_ATECO_2002)
  INNER JOIN classificazione_strade AS cs ON c.DENOMINAZIONE_CORRENTE =
cs.[DENOMINAZIONE_CORRENTE]
WHERE
  c.C_ATECO_2002 IN (
    SELECT ra.C_ATECO_2002
    FROM raggruppazioni_ateco AS ra
    WHERE ra.[GROUP] IS NOT NULL
  )
  AND (c.DATA_CHIUSURA >= #1/1/2020# OR c.DATA_CHIUSURA IS Null)
  AND c.DENOMINAZIONE_CORRENTE IN (
    SELECT cl.[DENOMINAZIONE_CORRENTE]
    FROM classificazione_strade AS cl
  );
```

Esempio di record estratto dalla query

commercio_per_politecnico_2022_Query							
ID	C_ATECO_2002	DATA_CHIUSURA	Gruppo	peso	Coordinata_X	Coordinata_Y	F2
12	930201		personal services	8	1394264.462	4989770.192	Parrucchieri e barbieri

## 9.4 Codici Python

Operazioni sul database

```
import pandas as pd
import geopandas as gpd
import datetime
from pyproj import Proj, transform

# Carica il file Excel in un DataFrame
file_excel = pd.read_excel('DB/DB_filtrato_finale.xlsx',
sheet_name='commercio_per_politecnico_2022_')

file_excel['DATA_CHIUSURA'] =
pd.to_datetime(file_excel['DATA_CHIUSURA'], errors='coerce')

selettore_data = datetime.datetime(2020, 1, 1)

def compare_dates(date1, date2):
    if isinstance(date1, datetime.datetime):
        return date1 > date2
    return False

filtered_df = file_excel[
    (file_excel['DATA_CHIUSURA'].apply(lambda x: compare_dates(x,
selettore_data)))
]

filtered_df = filtered_df[
    (file_excel['Coordinata_X'].notnull()) &
    (file_excel['Coordinata_Y'].notnull()) &
    (file_excel['C_ATECO_2002'].notnull())
]

# Salva il DataFrame filtrato in un nuovo file Excel
filtered_df.to_excel('Database_filtrato.xlsx', index=False)

inProj = Proj(init='epsg:3003')
outProj = Proj(init='epsg:4326')

for index, row in filtered_df.iterrows():
    # Estrai le coordinate
    coordinata_y_str = str(row['Coordinata_Y']) # Converti a stringa
    coordinata_x_str = str(row['Coordinata_X']) # Converti a stringa
```

```

# Verifica che le coordinate siano stringhe non vuote
if coordinata_y_str.strip() != '' and coordinata_x_str.strip() !=
'':
    try:
        # Tentativo di conversione delle coordinate UTM in float
        lat2 = pd.to_numeric(coordinata_y_str, errors='coerce')
        lon2 = pd.to_numeric(coordinata_x_str, errors='coerce')

        # Verifica che le coordinate siano valori numerici validi
        if not pd.isnull(lat2) and not pd.isnull(lon2):
            # Converti le coordinate UTM in latitudine e
longitudine
            lon, lat = transform(inProj, outProj, lon2, lat2)

    except ValueError as e:
        # Gestisci il caso in cui la conversione in float non
riesce
        print(f"Errore: Impossibile convertire coordinate in float
per l'indice {index}. Errore: {e}")

print('File filtrato e salvato correttamente')

# Crea file CSV separati per ciascun valore unico di 'Gruppo_ATECO'
for gruppo, gruppo_df in file_excel.groupby('Gruppo'):
    # Crea un nome di file CSV basato sul valore di 'Gruppo_ATECO'
    nome_file = f'coordinate_ateco_{gruppo}.csv'

    # Salva il DataFrame del gruppo nel file CSV
    gruppo_df[['Coordinata_X', 'Coordinata_Y']].to_csv(nome_file,
index=False)

    print(f'Coordinate per GRUPPO={gruppo} salvate correttamente in
{nome_file}')

```

## Analisi ANOVA

```
from scipy.stats import f_oneway
import pandas as pd

## Specifica il percorso del tuo file Excel
percorso_file_excel = 'DB/DB_anova.xlsx'

# Carica i dati dal file Excel in un DataFrame pandas
data = pd.read_excel(percorso_file_excel)

# Eseguiamo l'ANOVA
f_statistic, p_value = f_oneway(data[data['gruppo'] == 'bars and
restaurants']['peso'],
                                data[data['gruppo'] == 'personal
services']['peso'],
                                data[data['gruppo'] ==
'convenience']['peso'],
                                data[data['gruppo'] ==
'shopping']['peso'],
                                data[data['gruppo'] ==
'supermarket']['peso'],
                                data[data['gruppo'] == 'Recreational
and Entertainment Centers']['peso'])
print("Statistiche su tutti i dati")
print("Statistiche F:", f_statistic)
print("Valore p:", p_value)

f_statistic, p_value = f_oneway(data[data['gruppo'] == 'bars and
restaurants']['peso'],
                                data[data['gruppo'] == 'personal
services']['peso'],)
print("\n\nStatistiche su bars and restaurants vs personal services")
print("Statistiche F:", f_statistic)
print("Valore p:", p_value)

f_statistic, p_value = f_oneway(data[data['gruppo'] == 'bars and
restaurants']['peso'],
                                data[data['gruppo'] ==
'convenience']['peso'],)
print("Statistiche su bars and restaurants vs convenience")
print("Statistiche F:", f_statistic)
print("Valore p:", p_value)

f_statistic, p_value = f_oneway(data[data['gruppo'] == 'bars and
restaurants']['peso'],
                                data[data['gruppo'] ==
'shopping']['peso'],)
```

```

print("Statistiche su bars and restaurants vs shopping")
print("Statistiche F:", f_statistic)
print("Valore p:", p_value)

f_statistic, p_value = f_oneway(data[data['gruppo'] == 'bars and
restaurants']['peso'],
                                data[data['gruppo'] ==
'supermarket']['peso'],)
print("Statistiche su bars and restaurants vs supermarket")
print("Statistiche F:", f_statistic)
print("Valore p:", p_value)

f_statistic, p_value = f_oneway(data[data['gruppo'] == 'bars and
restaurants']['peso'],
                                data[data['gruppo'] == 'Recreational
and Entertainment Centers']['peso'],)
print("Statistiche su bars and restaurants vs Recreational and
Entertainment Centers")
print("Statistiche F:", f_statistic)
print("Valore p:", p_value)

f_statistic, p_value = f_oneway(data[data['gruppo'] == 'personal
services']['peso'],
                                data[data['gruppo'] ==
'convenience']['peso'],)
print("\n\nStatistiche su personal services vs convenience")
print("Statistiche F:", f_statistic)
print("Valore p:", p_value)

f_statistic, p_value = f_oneway(data[data['gruppo'] == 'personal
services']['peso'],
                                data[data['gruppo'] ==
'shopping']['peso'],)
print("Statistiche su personal services vs shopping")
print("Statistiche F:", f_statistic)
print("Valore p:", p_value)

f_statistic, p_value = f_oneway(data[data['gruppo'] == 'personal
services']['peso'],
                                data[data['gruppo'] ==
'supermarket']['peso'],)
print("Statistiche su personal services vs supermarket")
print("Statistiche F:", f_statistic)
print("Valore p:", p_value)

```

```

f_statistic, p_value = f_oneway(data[data['gruppo'] == 'personal
services']['peso'],
                                data[data['gruppo'] == 'Recreational
and Entertainment Centers']['peso'],)
print("Statistiche su personal services vs Recreational and
Entertainment Centers")
print("Statistiche F:", f_statistic)
print("Valore p:", p_value)

f_statistic, p_value = f_oneway(data[data['gruppo'] ==
'convenience']['peso'],
                                data[data['gruppo'] ==
'shopping']['peso'],)
print("\n\nStatistiche su convenience vs shopping")
print("Statistiche F:", f_statistic)
print("Valore p:", p_value)

f_statistic, p_value = f_oneway(data[data['gruppo'] ==
'convenience']['peso'],
                                data[data['gruppo'] ==
'supermarket']['peso'],)
print("Statistiche su convenience vs supermarket")
print("Statistiche F:", f_statistic)
print("Valore p:", p_value)

f_statistic, p_value = f_oneway(data[data['gruppo'] ==
'convenience']['peso'],
                                data[data['gruppo'] == 'Recreational
and Entertainment Centers']['peso'],)
print("Statistiche su convenience vs Recreational and Entertainment
Centers")
print("Statistiche F:", f_statistic)
print("Valore p:", p_value)

f_statistic, p_value = f_oneway(data[data['gruppo'] ==
'shopping']['peso'],
                                data[data['gruppo'] ==
'supermarket']['peso'],)
print("\n\nStatistiche su shopping vs supermarket")
print("Statistiche F:", f_statistic)
print("Valore p:", p_value)

f_statistic, p_value = f_oneway(data[data['gruppo'] ==
'shopping']['peso'],
                                data[data['gruppo'] == 'Recreational
and Entertainment Centers']['peso'],)
print("Statistiche su shopping vs Recreational and Entertainment
Centers")

```

```

print("Statistiche F:", f_statistic)
print("Valore p:", p_value)

f_statistic, p_value = f_oneway(data[data['gruppo'] == 'Recreational
and Entertainment Centers']['peso'],
                                data[data['gruppo'] ==
'supermarket']['peso'],)
print("\n\nStatistiche su Recreational and Entertainment Centers vs
supermarket")
print("Statistiche F:", f_statistic)
print("Valore p:", p_value, "\n\n")

```

Risultato ANOVA

```

PS C:\PoliTo\LM\II Anno\Tesi> &
C:/Users/Lorenzo/AppData/Local/Programs/Python/Python311/python.exe
"c:/PoliTo/LM/II Anno/Tesi/ANOVA.py"
Statistiche su tutti i dati
Statistiche F: 3.5095906825399217
Valore p: 0.003627305379905054

```

```

Statistiche su bars and restaurants vs personal services
Statistiche F: 1.185564512741665
Valore p: 0.27644727498424004
Statistiche su bars and restaurants vs convenience
Statistiche F: 9.601063523292872
Valore p: 0.0019677349031444443
Statistiche su bars and restaurants vs shopping
Statistiche F: 0.5003872722929315
Valore p: 0.47947651255514245
Statistiche su bars and restaurants vs supermarket
Statistiche F: 0.16607829038074454
Valore p: 0.6837429218970883
Statistiche su bars and restaurants vs Recreational and Entertainment
Centers
Statistiche F: 7.096283197095146
Valore p: 0.007898644294540028

```

```

Statistiche su personal services vs convenience
Statistiche F: 2.349761699166461
Valore p: 0.12544792985853034
Statistiche su personal services vs shopping
Statistiche F: 0.10794151690806905
Valore p: 0.7425721476512626
Statistiche su personal services vs supermarket
Statistiche F: 0.6676447576681203
Valore p: 0.4142323092083562

```

Statistiche su personal services vs Recreational and Entertainment Centers

Statistiche F: 5.875876085437993

Valore p: 0.015675891881399145

Statistiche su convenience vs shopping

Statistiche F: 3.4417683352294075

Valore p: 0.0637070712665904

Statistiche su convenience vs supermarket

Statistiche F: 1.691800527064113

Valore p: 0.19354017730178516

Statistiche su convenience vs Recreational and Entertainment Centers

Statistiche F: 4.010647228774691

Valore p: 0.04537252316595747

Statistiche su shopping vs supermarket

Statistiche F: 0.46817133525453203

Valore p: 0.49415848280415764

Statistiche su shopping vs Recreational and Entertainment Centers

Statistiche F: 6.359461324292393

Valore p: 0.011997549497051613

Statistiche su Recreational and Entertainment Centers vs supermarket

Statistiche F: 5.644967131891484

Valore p: 0.020207072434670712