Big Data Co-occurrence Analysis Using Foursquare's Open Places

GROUP 15: Gianantonnio Sanrocco (s3532542), Ilja Kaniscev (s2810832),
Lars Swennenhuis (s2836645), Alexandra Albiño (s3485935), Ferdinand de Jong (s2801612)

28 January 2025

1 Introduction

Points-of-interest (POI) are specific points of urban landscapes, serving as hubs of economic, social and cultural activity. Spatial relationships between different POI's often reflect the different underlying cultural and economic dynamics. Studying these POI's and their relationships can provide valuable information. For example, cafes may often appear near office spaces, while entertainment venues may cluster around restaurants. Understanding the co-occurrence patterns between the POI's is crucial for a number of reasons. First, these patterns provide valuable information about social, cultural and economic differences in different countries or regions. For example, POI clusters in densely populated cities in Europe may differ from those in the United States or Japan. Second, these patterns can provide valuable insights for businesses and marketers. Analyzing which POI's frequently co-occur, gives opportunities to optimize location selection. Businesses can make data-driven decisions when selecting locations for their new venues. Furthermore, they can create marketing strategies designed for a specific region. Additionally, co-occurrence patterns are valuable for urban planning and development. By identifying areas where complementary POI's are lacking, planners can address gaps in infrastructure and improve accessibility.

Despite the importance of these patterns, there remains a gap in understanding how these co-occurrence patterns vary across countries. While previous research has analyzed POI's within individual cities or regions, few to none studies have explored the cross-country differences. This study aims to address this gap by answering the central question: **How do co-occurrence patterns of POI categories differ across countries for targeted POI's?** To achieve this, several sub-research questions will be explored:

1. **What are the most common country-specific co-occurrence patterns for targeted POIs?**
2. **Are there specific categories that frequently co-occur with targeted POIs in certain countries but not in others?**
3. **How does the spatial radius or density of targeted POIs correlate with the diversity of nearby POI categories across countries?**
4. **How does the closure (`date_closed`) of a targeted POI impact the longevity or closure of its co-occurring POIs in different countries?**

Our main POI for this research was a grocery store. This POI was selected because of the global presence of this category, which allows us to see the cross-cultural insights of this POI. Furthermore, grocery stores often co-occur with a wide variety of other POI's making them an ideal candidate for studying spatial relationships. The countries for each research question were selected based on factors such as the number of available grocery stores and the continent in which the country is located. Using the Euclidean distance, we analyzed the nearby places. We did this for two ranges: 1 kilometer and 2 kilometers. **These were the key results for the research questions:**

1. **Co-occurrences and Cultural Differences in Urban Planning** Analysis of POI (Point of Interest) co-occurrences reveals patterns in urban planning and cultural distinctions across different regions with relation to grocery stores:
   - **Brazil & Russia**: Similar patterns in housing and office spaces.
   - **United States**: Office spaces stand out prominently.
   - **Europe (Germany)**: Dominated by business and professional services.

- **Oceania (Australia)**: High presence of cafés and office spaces.
- **Asia (Japan)**: Predominantly catering services, including Japanese restaurants and bars.

2. **Increased POI Diversity at Larger Spatial Ranges**
   - A broader spatial range (2 km) exhibits greater diversity in POI categories.
3. **Dominance of Grocery Stores in Smaller Spatial Ranges**
   - Within a 1 km radius, offices emerge as the most dominant POI category.
4. **Impact on Restaurants and Hospitality Industry**
   - Restaurants were the most frequently closed POI category across six selected countries.
   - Cafés, bars, and other restaurants were among the most affected categories.

Finally, we have a URL for the code repository: https://gitlab.utwente.nl/computer-systems-project/2022-2023/student-projects/cs22-26/mbd-project

2 Related work

A large number of studies have been carried out with the aim of studying points of interest. A study conducted by Lin et al. (2025) highlights the importance of using databases of Points of Interest not only for the comprehensive study of the quantity, type, or variety and density of POIs that may exist in a specific area, but also proposes analyzing the interactions and contextual relationships between POIs. [5] However, it does not do so across different countries but is instead focused on the city of London. This study transforms the sequence of Points of Interest (POIs) along streets into words and applies the Latent Dirichlet Allocation (LDA) model to identify topics based on the observed patterns. For instance, a street with a sequence like cinema, mini-golf, and arcade would correspond to the topic of entertainment. In this way, the study introduces a novel perspective for analyzing the relationships between POIs within a given area.

Similar to previous work, the study by Zhou et al. (2025) focuses on identifying spatial co-location patterns, which refer to the tendency of certain POI categories to appear together in specific spatial arrangements. [8] This study introduces a novel approach by employing embedding techniques and contextual similarity, integrating the local context of POIs within a category (within a block) with the global context (across the entire island) to uncover more intricate relationships. While the analysis is confined to a single geographic location, Xiamen Island, it examines the spatial co-location patterns of five randomly selected categories (chinese restaurants, fast food restaurants, ATMs, leisure venues, and malls), offering valuable insights into their distribution and interactions.

On the other hand, the study conducted by Dong et al. (2024) examines POI co-occurrence patterns focusing on 297 cities in China to explore how imbalances in urban development influence their distribution. [2] Utilizing the Word2Vec model, the study highlights functional disparities between more developed cities, characterized by a high concentration of advanced POIs such as high-end restaurants, airports, and shopping malls, and less developed cities, where POIs are more dispersed and infrastructure is notably lacking. This approach shifts the focus from small interactions between categories to a macro-level understanding of how POI distribution reflects economic and functional disparities across different cities in the same country.

Chen et al. (2020) examined the spatial organization of urban functions, activities that take place in urban spaces, in 25 cities in China using POI data and co-location pattern mining. [1] They identified frequently co-located POI categories within a certain distance, where categories such as retail, catering, and recreation tend to cluster together, and where mixed urban functions positively influence mobility, reduce energy consumption and promote vibrant neighborhoods.

Similarly, Koohpayma et al. (2019) used POIs to analyze how they influence parking violations in Tehran, Iran, a region characterized by high traffic levels and significant pollution. [4] The findings indicate that certain POIs, such as hospitals, have a major impact on these violations, primarily due to the lack of parking spaces in congested urban areas, while others, like universities, have less impact.

Based on the reviewed related work, it is evident that existing studies primarily focus on analyzing points of interest and co-occurrence factors within specific regions or cities. This highlights a gap in the literature, which this project aims to address by providing a comprehensive cross-country analysis of co-occurrence patterns related to supermarkets.

3 Dataset

To answer the stated research question, large datasets are needed with information about many places in a large area (at least the size of a country). For this research, the Foursquare Open Source Places dataset has been used. [3] This dataset contains entries on locations throughout the world, although it is more concentrated on the USA and western countries. The countries most represented in the data are (respectively): USA, India, Turkey, UK and Germany.

All entries have a unique id, which functions as a primary key, and a name, which is the display name of the location. There are a number of fields containing geographical information, such as `country` (never null), `latitude`, `longitude`, `address`, `po_box`, `region` etc. Some or many of these may be null. Additionally, some fields provide information in the time dimension. Nearly all entries have a `date_created` field and all have a `date_refreshed` field. These fields enclose information as to when the entry was first added and when the entry was last accessed (by a human or a crawler). A few entries have a `date_closed` field to indicate that the location is no longer active.

A few fields contain contact information of the location. A `tel` (ephone number) is the most common, with nearly half the entries having one. Alternatively, social media accounts or an email or website are provided. An important feature of the dataset are the category fields, which allow filtering commercial locations, or a specific type of service, such as Health and Medicine, Dining and Drinking or Business and Professional Services. This information is particularly important for our research. All entries have a `placemaker_url` attached as well, which relates the entry to a feature on the Foursquare website. Lastly, some locations have been blessed with a bounding box or a well known binary (WKB) field to describe the shape of the location.

The full dataset has a size of 12.9 GB as parquet files (compressed). The dataset can also be accessed using an API available in Spark, meaning that it is not necessary to download the entire dataset to interact with it. More information about the dataset is available in the Foursquare documentation. [7]

4 Method

To start with, the datasets with places and categories were uploaded to HDFS.

4.1 Methodology addressing common country-specific co-occurrence patterns and differences

The targeted point of interest (POI) is identified as a 'Grocery store'. To ensure accurate identification, the categories were manually analyzed to understand the categorization of places. Based on this analysis, it was decided to filter categories using the following strings: %grocer%, %supermarket%, and %food store%, which effectively captured relevant categories.

The resulting categories were:

Place Categories
Meat and Seafood Store
Imported Food Store
Supermarket
Health Food Store
Grocery Store

Table 1: Table showing categories selected.

In order to decide on the countries within which the grocery stores are going to be searched, we have picked the representative countries from 6 continents: Europe, Asia, Oceania, Africa, South and North America. The countries were selected based on the largest amount of grocery stores available within a continent. Once the countries scope was identified by having the most efficient way of narrowing down the scope of analyzed countries and at the same time, having an ultimate continent representative, the search for the nearest places around grocery stores could start.

In order to analyze the most frequent places surrounding grocery stores per country, a radius of 1 km has been chosen as a representative value, capturing the area within walking distance from the targeted points of interest (POIs). By leveraging the presence of coordinates for each place, we used the Euclidean distance in order to limit the nearby places next to targeted POI only within 1 km. Along with identifying all nearby places within radius, the corresponding location's categories were fetched to further base the analysis on.

Below, is the resulting schema for each grocery store processed with their nearby places array, featuring the location with corresponding categories that it belongs to.

```
root
 |-- fsq_place_id: string (nullable = true)
 |-- name: string (nullable = true)
 |-- country: string (nullable = true)
 |-- nearby_places: array (nullable = false)
 |    |-- element: struct (containsNull = false)
 |    |    |-- place_id: string (nullable = true)
 |    |    |-- place_name: string (nullable = true)
 |    |    |-- place_categories: array (nullable = false)
 |    |    |    |-- element: struct (containsNull = false)
 |    |    |    |    |-- category_id: string (nullable = true)
 |    |    |    |    |-- category_name: string (nullable = true)
```

Figure 1: Nearby places schema of POI

4.2 Methodology for radius analysis and POI closures analysis

The script to address research question with different radius uses PySpark to analyze supermarket data and nearby points of interest (POIs). It begins by loading Parquet files into DataFrames, exploding array columns to break down category IDs into individual rows. A join operation enriches the places data with category names. The dataset is then filtered to focus on supermarkets, grouping and counting them by country to identify the top 20 countries with the highest supermarket counts. For the top 10 countries, the script computes nearby POI categories within a 1-km radius of each supermarket. Using mathematical distance calculations, it identifies and aggregates the diversity and frequency of nearby categories. These results are saved as serialized files, providing a detailed view of category co-occurrence patterns and supermarket surroundings for further analysis.

As regards the script for POI closures analysis, to calculate closure counts for each category in specified countries (Germany, France, Italy, Poland, Spain, and Turkey), the script filters the dataset to include only POIs with closure dates (`date_closed`) in these countries. It then groups the data by category names using the groupBy method and applies count to determine the number of closures per category. The results are sorted in descending order using orderBy, highlighting the categories with the highest closure counts. This analysis reveals that restaurants are most impacted by closures across the target regions.

To analyze the impact of restaurant closures on nearby categories, the script first isolates restaurants and identifies POIs within a 2-km radius of each restaurant. This is done by calculating the Haversine distance between restaurant and POI coordinates, filtering for POIs within the specified range. Only POIs that closed after the restaurant's closure are considered. Finally, the script groups the data by category name and counts the affected POIs for each category, providing insights into which categories are most influenced by restaurant closures. The top 10 categories by impact are visualized to illustrate the effects of restaurant closures on their surroundings.

5 Results

In this section, we will detail the results that were achieved using the specified methods. Many graphs were produced, which prevented us from visualizing the results in a concise manner, so these graphs are located in the appendix. See them here: 10.1 and 10.2.

5.1 Co-occurrence patterns for targeted POIs

5.1.1 Country selection

The list of countries that we found to be most populated with Grocery stores per continent listed in Table 2.

Continent	Top Country	Count
North America	US	1921
South America	BR	400
Europe	DE	856
Asia	JP	593
Asia	RU	323
Oceania	AU	137
Africa	ZA	39

Table 2: Top Grocery store counts per continent

5.1.2 Execution performance

Execution real time: 4m30s with 2 executors (FindNearbyPlaces.py)

5.1.3 POI co-occurence patterns

For each country, we collected the 7 most co-occurring commercial location types. To create a better sense of place, we also generated a map of the surroundings of locations for every country in the analysis.

5.1.4 Australia

Top 7 categories surrounding Grocery stores in Australia are cafes and office being the most frequent ones, later followed by hair salons, taking up 3rd most frequent position. Restaurants are on the 4th place, resulting to 27, which is is 2.5 less than occurrence of Cafes. Furniture and Home stores, together with Law offices and Coffee shop are featured the same amount of times around groceries shops, being placed on 3-7 places. See 4 and 10.1.1

5.1.5 Brazil

As for Brasil, the dominant places around Grocery shops are offices (1st place), apartments or condos (2nd place). Hair salons are also quite frequent, being on the 3rd spot. 4th place is featured structure, which represents landmarks or outdoor places, for instance monuments or swimming pools. Top 5 most frequent places are doctor's offices which is more frequent than appearance of bars by a bit. Less frequent in top 7 are restaurants, which is more than 3 times inferior to the first position. See 6 and 10.1.2

5.1.6 Germany

In Germany, grocery stores are predominantly surrounded by business and professional service establishments. Clothing shops hold the second position. Law offices and construction supply stores share the third and fourth spots with equal representation. They are followed by hair salons and insurance agencies, which also occur with the same frequency. See 8 and 10.1.3

5.1.7 Japan

In Japan, the top 5 frequent categories around Grocery stores within 1km are restaurants, bars and cafes, which are all together 5 times more frequent than hair salons and offices combined. See 10 and 10.1.4

5.1.8 Russia

From the chart of Russia, it can be assumed that Grocery stores are most surely being built in living areas to which Apartment or Condo category on the first place corresponds to. Second and third places are taken by offices and hair salons which resulted in the same frequency amount. Structures, banks, pharmacies and cafes are top 4-7 most frequent locations around Russian Grocery stores with approximately same appearance frequency. See 12 and 10.1.5

5.1.9 USA

This graph shows the top 7 categories of places surrounding grocery stores in the United States. Offices dominate the surroundings with a count of 626, significantly surpassing other categories. Structures take second place with 339 occurrences, followed by doctor's offices with 240. Financial services (231) and apartments or condos (229) occupy the fourth and fifth positions, respectively. Restaurants (228) and hair salons (219) round out the list, reflecting a diverse mix of professional, residential, and service-oriented establishments near grocery stores. See 14 and 10.1.6

5.1.10 South Africa

This graph shows the top 7 categories of places surrounding grocery stores in South Africa. The data highlights a balanced presence of construction supplies stores, coworking spaces, doctor's offices, and burger joints, each appearing with similar frequencies. Additionally, fast food restaurants, cafés, and neighborhoods are also commonly found near grocery stores. Due to the limited data points, these results provide only a general indication of the types of establishments typically located in proximity to grocery stores in the region. See 16 and 17

5.2 POI Density vs. Nearby Category Diversity Across Countries

Execution time: 4m30s with 2 executors

In this research, we aim to analyze the patterns and diversity of nearby points of interest (POI) categories in relation to supermarkets or grocery stores across different countries. Specifically, we focus on examining the diversity of POI categories within 1 km and 2 km radius. The 1 km radius allows us to capture co-occurrence patterns within a short range, offering insights into the immediate neighborhood dynamics of supermarkets. Meanwhile, the 2 km radius provides a broader perspective, enabling us to explore how the distribution and variety of POI categories change as the distance increases. This dual-distance approach enhances our understanding of spatial patterns and the diversity of urban landscapes surrounding supermarkets globally.

To address the following problem, another set of countries has been reselected. The countries involved in our analysis are the top-10 countries with the most supermarkets (or grocery stores) in our dataset. This allows us to make reliable analysis because we can get valuable insights about different surrounding POI categories. Table 3.

Country	Number of Supermarkets
US	1921
DE	856
FR	628
JP	593
GB	540
IT	442
BR	400
PL	396
ES	369
TR	364

Table 3: Number of Supermarkets by Country

5.2.1 Italy (IT)

1 km Range: Food-related POIs dominate, such as restaurants, cafés and pizzerias. There is a wide representation of professional, law offices and business categories present. Hotels and bars are almost twice less frequent.
2 km Range: Professional services such as offices and business services increase significantly, indicating an urban sprawl toward economic and administrative infrastructure. Grocery stores and restaurants maintain their prominence, but the relative distribution flattens, suggesting more balanced POI diversity in the larger radius.

5.2.2 Japan (JP)

1 km Range: Focus on food and leisure, with categories like Japanese restaurants, sake bars, and grocery stores dominating. These categories reflect dense culinary and cultural activities within smaller areas.
2 km Range: A substantial increase in food-related POIs (Japanese restaurants, bars) is observed, but the introduction of broader categories like offices and professional services suggests an extension of economic activity beyond immediate urban cores.

5.2.3 Poland (PL)

1 km Range: Business services and retail dominate, reflecting a focus on essential and economy services. There is some representation of construction-related POIs, indicating a practical infrastructure focus.
2 km Range: Increasing inclusion of categories emerges with the rise of legal services, hotels, and financial services. Although retail remains dominant, there is a significant increase in automotive repair shops.

5.2.4 Turkey (TR)

1 km Range: Residential categories such as apartments, co-working spaces, and housing developments are dominant. Hair salons and cafes are also notable.
2 km Range: There are no significant changes in the more occurred categories, they are proportionally increased in frequency.

5.2.5 United States (US)

1 km Range: Offices, health-related POIs (doctor's offices), and residential structures dominate, showcasing a balance of essential and residential services.
2 km Range: Offices and professional services rise sharply, while grocery stores maintain their dominance.

5.2.6 Brazil (BR)

1 km Range: Offices, in general are dominant, but also with strong representation in healthcare-related POIs like dentist and doctor's offices. Residential POIs also feature prominently.
2 km Range: Office spaces and apartments continue to dominate, but new categories like arts and entertainment and automotive repair shops gain prominence.

5.2.7 Germany (DE)

1 km Range: Business, professional, legal services dominate along with retailing. Health and beauty care are present but secondary.
2 km Range: Professional services grow substantially, with offices, advertising agencies, and law offices showing notable increases. Germany transitions from essential retail dominance to a more business-oriented infrastructure in the larger radius.

5.2.8 Spain (ES)

1 km Range: Food-related categories (restaurants, tapas restaurants, Cafè, Bar) dominate, reflecting Spain's cultural emphasis on dining and leisure.
2 km Range: Financial and hospitality categories such as banks and hotels see significant growth. Cafés and restaurants remain prominent, but the increase in professional and leisure-related POIs signals a well-rounded suburban expansion.

5.2.9 France (FR)

1 km Range: Offices, hair salons and bars dominate, with some representation of French-specific categories like French restaurants.
2 km Range: Offices become the most dominant category, followed by restaurants and hair salons. The addition of hotels highlights suburban diversification.

5.2.10 United Kingdom (GB)

1 km Range: Professional services like offices and financial services dominate, followed by retailing and construction supply stores.
2 km Range: Offices and professional services grow significantly, with new categories such as legal services and advertising agencies appearing. The charts indicate a balanced expansion of cultural, professional, and economic infrastructure with a higher range.

5.3 Closure of POIs

Execution time: 0m33s with 2 executors

In this section we are taking into account for our research question the countries with more data around Europe.

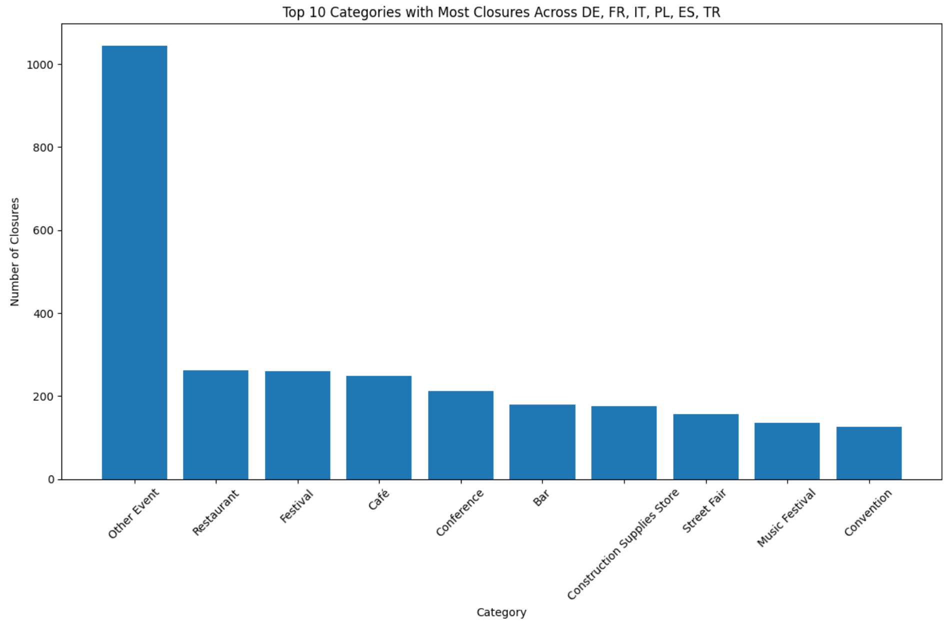


Figure 2: Top 10 Categories with Most Closures Across DE, FR, IT, PL, ES, TR

Figure 1 shows the top 10 categories with the highest number of closures in Germany, France, Italy, Poland and Turkey. The data highlights that restaurants, festivals, and cafes have the highest number of closures (around 200 each). These are followed by conferences, bars, construction supplies stores, and street fairs, while music festivals and conventions have the fewest closures among the top 10 categories.

This distribution suggests that categories closely tied to social gatherings and events are particularly vulnerable, reflecting the impact of reduced activity in areas dependent on foot traffic and community engagement.

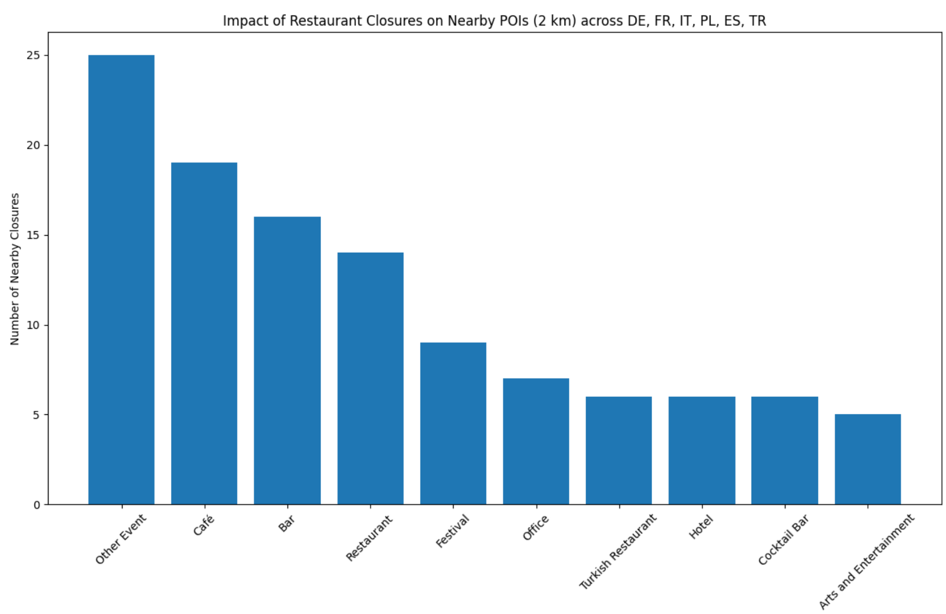


Figure 3: Impact of Restaurant Closures on Nearby POIs (2km) across DE, FR, IT, PL, ES, TR

Figure 2 shows how restaurant closures within a 2 km radius in Germany, France, Italy, Poland, Spain, and Turkey significantly impact nearby POIs. The most affected categories are cafes and bars, followed by restaurants, festivals, and offices. Turkish restaurants, hotels, and cocktail bars also show similar levels of impact, highlighting their shared vulnerability.

This demonstrates how interconnected these places are, as restaurants often act as a magnet, bringing in people and supporting nearby businesses like cafes and bars.

6 Discussion

6.1 POI co-occurrence for countries

For every country, something can be said about the distribution. In Australia, for example, a lot of catering is present nearby grocery stores, cafés in particular. In Brazil, a lot of offices and a large concentration of housing is located nearby, suggesting a loose approach to urban planning. In Germany, various types of store are located close to grocery stores. This may indicate that stores are often located together in shopping centres. In Japan, mainly many different types of catering are situated close to grocery stores, which may have to do with a culture of eating less at home, and more with others, perhaps colleagues or friends at other places. In Russia, in some ways like Brazil, housing and office space are most common close to grocery stores. This may again suggest loose urban planning. In the US, office space is particularly present close to grocery stores, with structures (landmarks or outdoor places) as a distant second. Perhaps this is an indication of the workaholic culture in the US. Structures may be explained by a nationalistic attitude among the citizens of this superpower. Sadly, for South Africa (and all other African countries) very little data was available. A variety of locations are located nearby grocery stores, but little can be said conclusively about this.

6.2 Co-occurrence pattern regional differences

Australia stands out in the list of countries that were selected, as the café occurs most close to grocery shops, where it is 5th for Japan, and does not enter the top 5 for other countries at all. It is unclear why this happens, and why somewhat similar countries like the US and South Africa do not have such a strong correlation. The example screenshot perhaps gives a plausible explanation, as the grocery store there is located in the old city centre, where cafés are more common. This may be the case for many other grocery stores. Germany is the only country in the comparison that has a large amount of locations classified as Business and Professional Services nearby its grocery stores. An unsurprising result is the fact that Japanese restaurants are most prevalent in Japan. It is unclear why there has been differentiated between Restaurant and Japanese Restaurant in the dataset. In the provided example, the grocery store

6.3 Co-occurrence diversity and distance

Some interesting trends can be noted in the results. Across all countries, the 2 km radius consistently introduces new categories, reflecting greater diversity. Categories such as financial services, hair salons, and legal services often appear only in the 2 km range. Offices and business and professional services are universally dominant in the smaller 1 km radius, highlighting the frequent co-occurrence of these categories. Countries like Germany, the UK, and the US show significant growth in office spaces and business services in the 2 km radius, while Spain and Turkey (and Japan) show increased representation of cultural and leisure-oriented POIs, such as restaurants, cafes, and entertainment, in the larger radius. This may be indicative of the classic protestant - catholic or North-European - South-European divide that runs across Europe, a cultural divide which sociologist Max Weber drew attention to in his famous work about the protestant work ethic. [6]

6.4 Closure of POIs

In figure 1, results shows that in terms of the number of POI closures in Germany, France, Italy, Poland, and Turkey, the highest numbers are concentrated in restaurants, festivals, and cafes (around 200 closures each). These categories are consistently impacted across all five countries, primarily because restaurants, festivals, and cafes are key spaces for social interaction and relationships. For instance, in Germany, bars and restaurants serve as important venues for work-related meetings or casual meetings with friends, reflecting the country's strong beer culture. In Italy and France, it is very common to visit bars or restaurants after work to meet friends and enjoy an Aperol Spritz or a glass of wine. In Poland and Turkey, social gatherings often revolve around traditional foods, such as tea and baklava, which are deeply ingrained in their cultural habits. These similarities highlight the shared importance of these POIs as social hubs, making them particularly vulnerable to disruptions, which explains their significant presence among the closures.

Moreover, the results in Figure 2 show the POIs affected by the closure of a restaurant within a 2 km radius in the aforementioned countries. The analysis highlights that the categories most impacted by restaurant closures are cafes, bars, and restaurants. Since these POIs are closely tied to social activities, it can be said that restaurants act as a magnet, attracting people to a particular area. For example, in these countries, there is a strong after-office culture, where it is common for people to visit a bar before or after dining out, or to go to a cafe after a meal. If a restaurant closes, these complementary visits are also likely to decline, which may lead to closures of nearby bars and cafes due to reduced foot traffic.

Therefore, in response to the sub-research question, it can be concluded that the closure of a restaurant does indeed have a significant impact, primarily leading to the closure of nearby bars, restaurants, and cafes. This is due to the reduced foot traffic and customer flow that typically follows when a restaurant shuts down, affecting the surrounding businesses that rely on the same audience.

7 Conclusion

This study analyzed the co-occurrence patterns of POI's across multiple countries using data from Foursqaure. By examining grocery stores and their surrounding POI's, we identified key spatial trends that provide insights into cultural dynamics and cross-country differences. Our findings show that Brazil and Russia show similarities in co-occurrence patterns related to grocery stores with housing and office spaces as the most frequently occurring POI's. Office spaces stand out prominently in the United States. In Europe, particularly in Germany, business and professional services are dominant. Oceania, represented by Australia, has a strong presence of cafés and office spaces, whereas Japan's urban landscape is heavily influenced by catering services, including Japanese restaurants and bars. Moreover, larger spatial ranges (2 km) consistently show a higher diversity of POI categories, while offices and businesses and professional services remain dominant within a 1 km radius. Cross-country variations in co-occurrence patterns reflect the underlying social and economical differences. For example, in the 2 km radius, office spaces and business stand out even more among the most common related POIs in the UK, Germany, and the US. In contrast, Japan, Spain, and Turkey show a higher presence of cultural and leisure-oriented POIs. Furthermore, the impact of closures across six selected countries reveals that restaurants were the most frequently closed category. The broader hospitality sector, including cafés, bars, and other restaurants, were the most affected by the closure of a restaurant.

8 Limitations

Our research faces several limitations, with one of the most significant being the lack of documented data for certain locations. In some countries, limited data availability may lead to a misrepresentation of the actual co-occurrence patterns present within those regions. This can lead to geographical bias and can impact our results. Furthermore, the quality of the data also imposes a limitation. In the last research question, we analyze store closures using `date_closed`. From this data, we analyze POI closures and identify co-occurrence patterns, though they may not always align with a specific POI. However, the reasons for closure are unknown. A store may close due to bankruptcy, but it could also shut down for reasons unrelated to financial issues. Such closures should not be included, as they do not contribute to the co-occurrence patterns of interest. Lastly, there is a bias in counting the same category multiple times when POIs fall within the same radius. Instead of being counted once, a category may be counted twice—once for the first POI and again for another POI in the same area. This can impact the results, particularly in high-density areas where similar categories are clustered closely together.

8.1 Future work

Regarding the future work of this project, it would be valuable to consider expanding the dataset to include broader geographic regions, improving the representation of areas with limited data. Additionally, identifying the possible causes of POI closures would allow for a better understanding of their impact on co-occurrence patterns. Furthermore, incorporating socioeconomic data, such as income levels and population density, could provide deeper insights into the relationship between POI closures and the dynamics of geographic areas.

9 References

References

[1] Yimin Chen et al. "Understanding the spatial organization of urban functions based on co-location patterns mining: A comparative analysis for 25 Chinese cities". In: *Cities* 97 (Feb. 2020), p. 102563. ISSN: 0264-2751. DOI: 10.1016/j.cities.2019.102563. URL: https://www.sciencedirect.com/science/article/pii/S0264275119307553 (visited on 01/27/2025).

[2] Guangsheng Dong et al. "Differences in Urban Development in China from the Perspective of Point of Interest Spatial Co-Occurrence Patterns". en. In: *ISPRS International Journal of Geo-Information* 13.1 (Jan. 2024). Number: 1 Publisher: Multidisciplinary Digital Publishing Institute, p. 24. ISSN: 2220-9964. DOI: 10.3390/ijgi13010024. URL: https://www.mdpi.com/2220-9964/13/1/24 (visited on 01/27/2025).

[3] *foursquare/fsq-os-places · Datasets at Hugging Face*. Jan. 2025. URL: https://huggingface.co/datasets/foursquare/fsq-os-places (visited on 01/10/2025).

[4] Javad Koohpayma et al. "Spatial Analysis of Curb-Park Violations and Their Relationship with Points of Interest: A Case Study of Tehran, Iran". en. In: *Sustainability* 11.22 (Jan. 2019). Number: 22 Publisher: Multidisciplinary Digital Publishing Institute, p. 6336. ISSN: 2071-1050. DOI: 10.3390/su11226336. URL: https://www.mdpi.com/2071-1050/11/22/6336 (visited on 01/27/2025).

[5] Xuhui Lin, Tao Yang, and Stephen Law. "From points to patterns: An explorative POI network study on urban functional distribution". In: *Computers, Environment and Urban Systems* 117 (Apr. 2025), p. 102246. ISSN: 0198-9715. DOI: 10.1016/j.compenvurbsys.2024.102246. URL: https://www.sciencedirect.com/science/article/pii/S0198971524001753 (visited on 01/27/2025).

[6] Max Weber. *The Protestant Ethic and the Spirit of Capitalism.* Trans. by Talcott Parsons. London: Routledge, 2005. ISBN: 0-415-25559-7.

[7] *Places OS Data Schemas*. en. URL: https://docs.foursquare.com/data-products/docs/places-os-data-schema (visited on 01/17/2025).

[8] Xusheng Zhou et al. "A Method for Mining Spatial Co-location Patterns Based on Contextual Similarity Among Categories". en. In: *Journal of Geovisualization and Spatial Analysis* 9.1 (Jan. 2025), p. 9. ISSN: 2509-8829. DOI: 10.1007/s41651-024-00211-2. URL: https://doi.org/10.1007/s41651-024-00211-2 (visited on 01/27/2025).

10 Appendix

10.1 Appendix A: Additional plots Co-occurrences

10.1.1 Australia (AU)

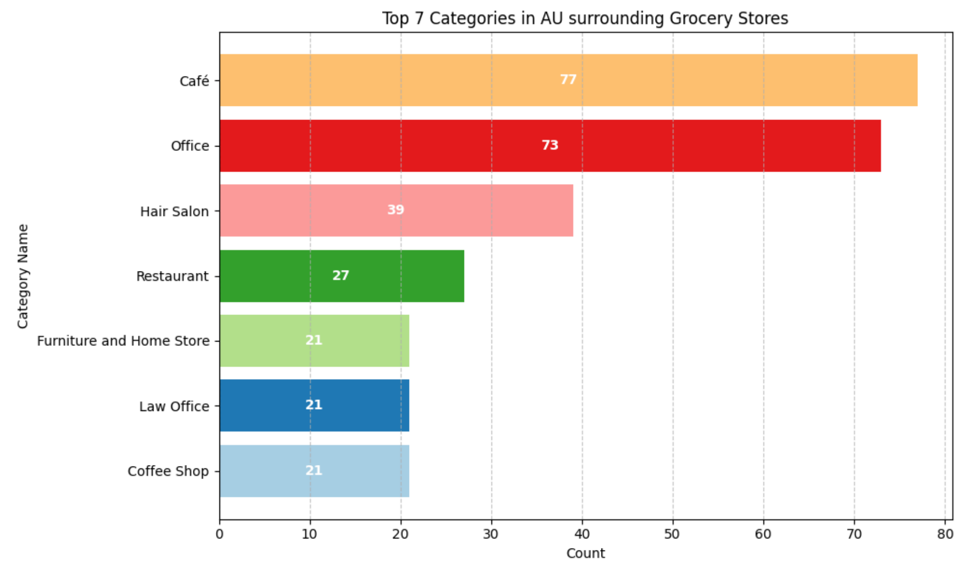


Figure 4: Top 7 categories in AU surrounding Grocery stores

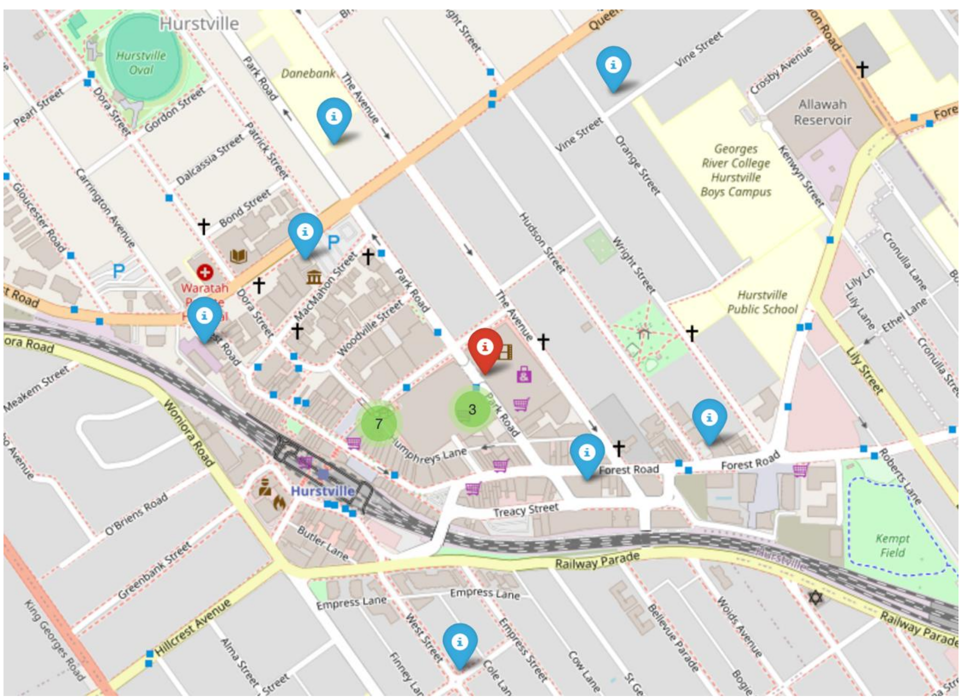


Figure 5: AU Grocery shop's example surrounding places (1km)

10.1.2 Brazil (BR)

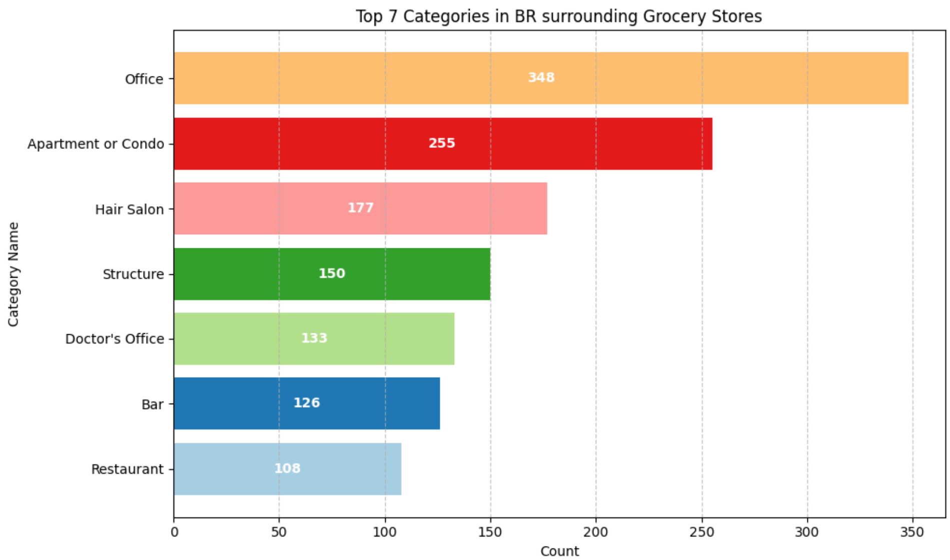


Figure 6: Top 7 categories in BR surrounding Grocery stores

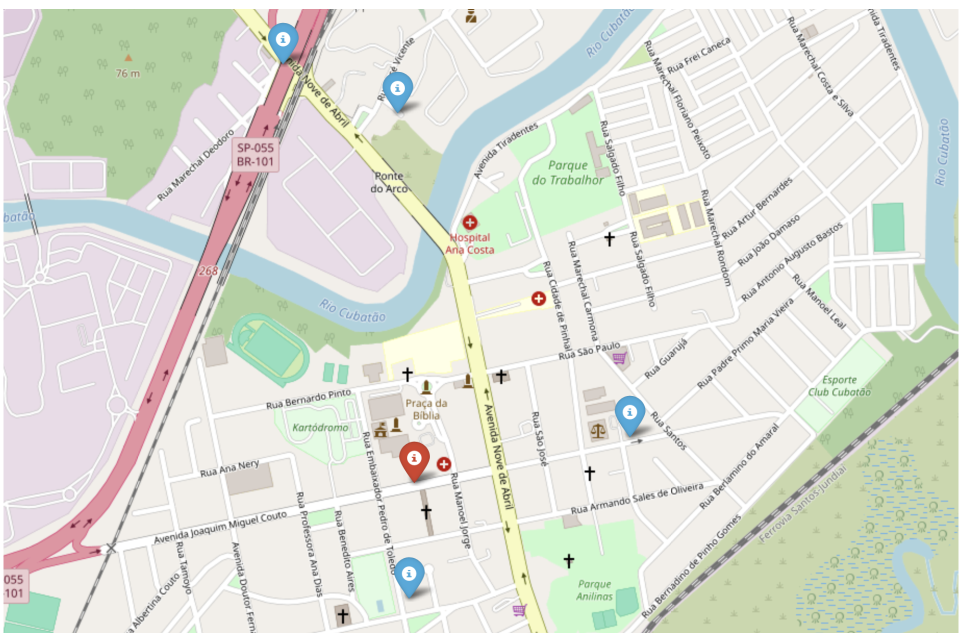


Figure 7: BR Grocery shop's example surrounding places (1km)

10.1.3 Germany (DE)

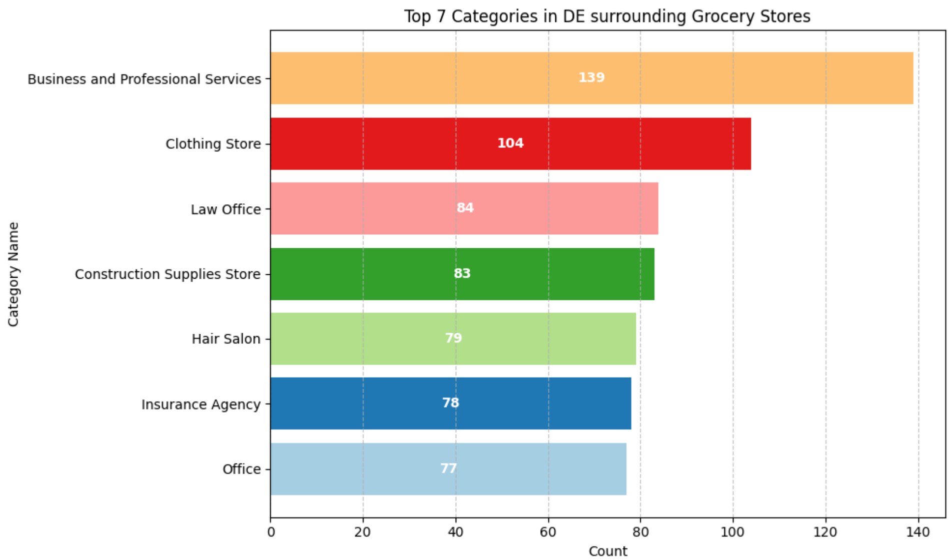


Figure 8: Top 7 categories in DE surrounding Grocery stores

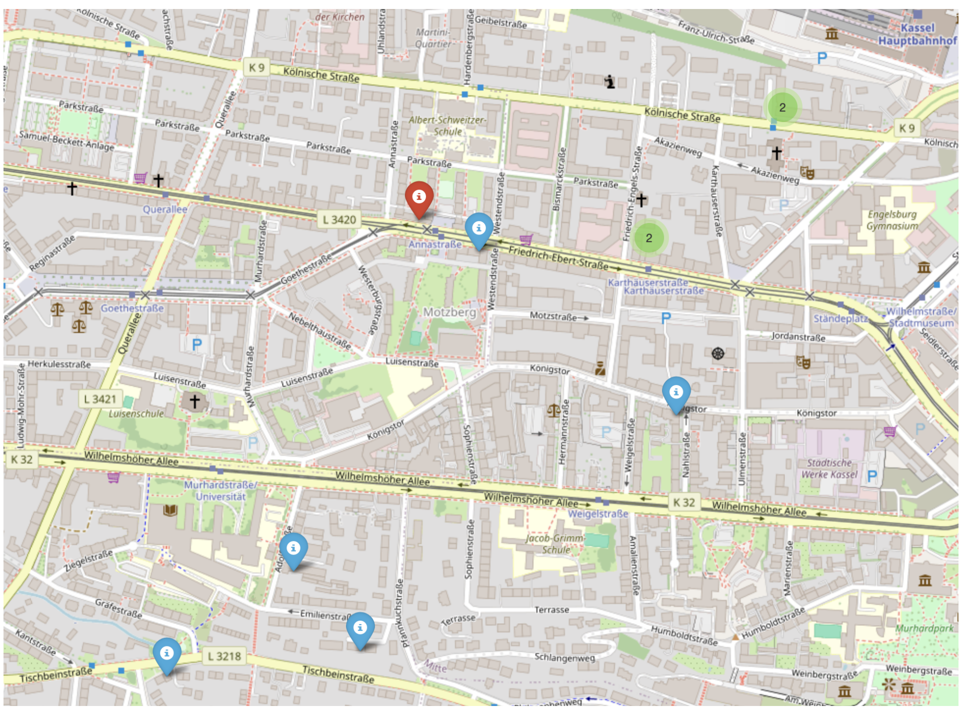


Figure 9: DE Grocery shop's example surrounding places (1km)

10.1.4 Japan (JP)

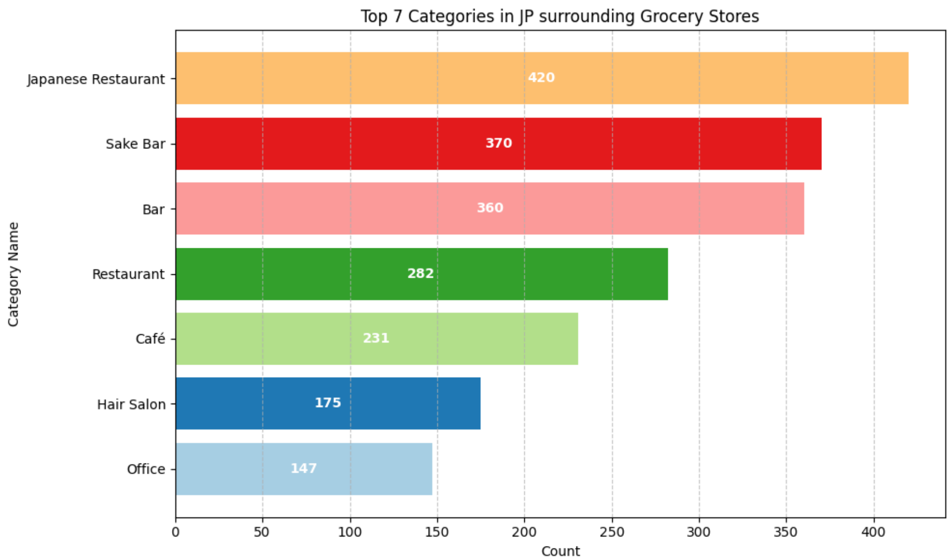


Figure 10: Top 7 categories in JP surrounding Grocery stores

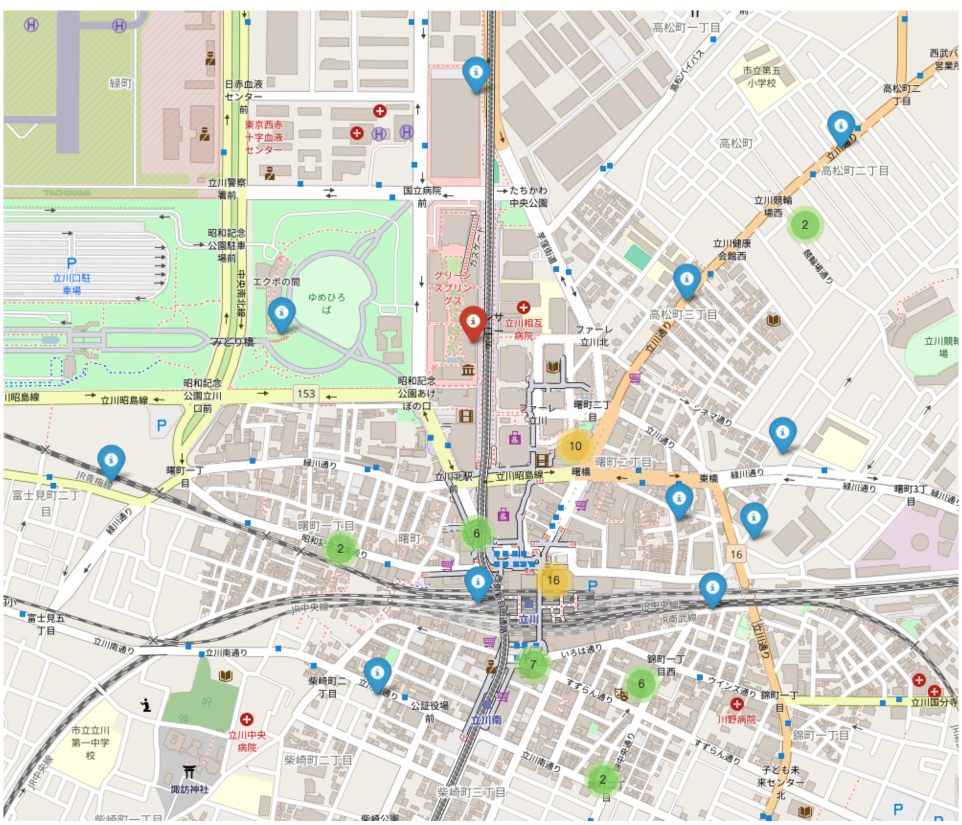


Figure 11: JP Grocery shop's example surrounding places (1km)

10.1.5 Russia (RU)

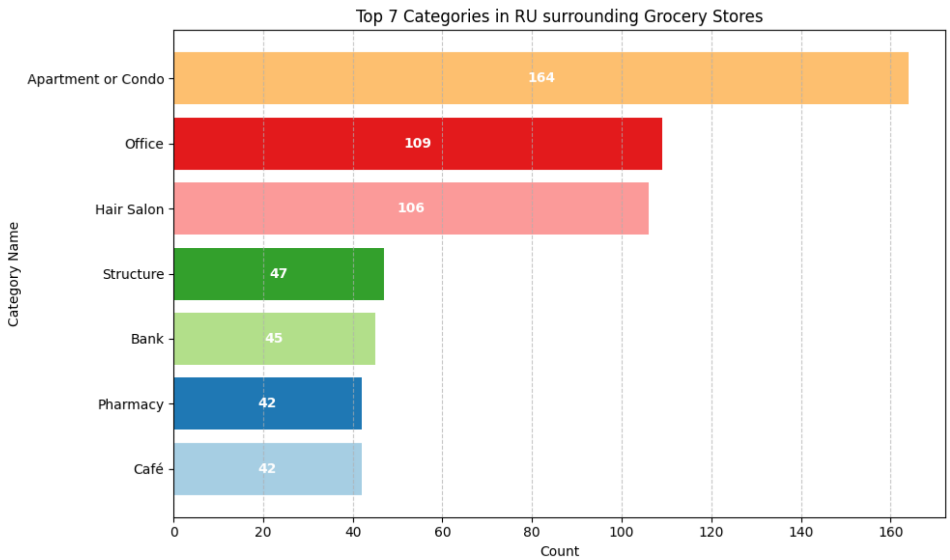


Figure 12: Top 7 categories in RU surrounding Grocery stores

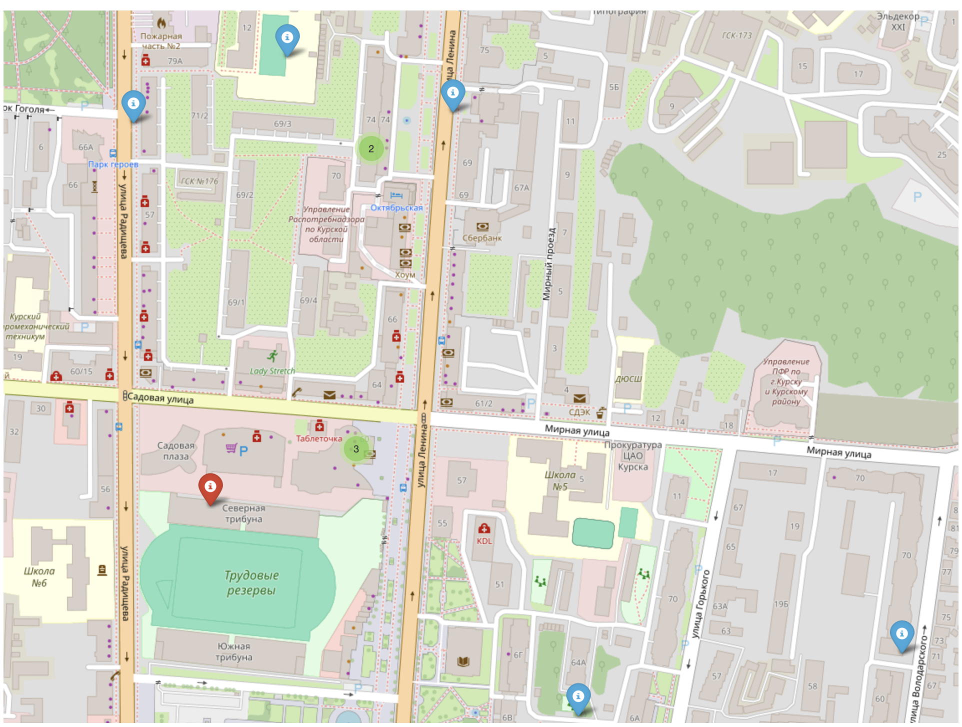


Figure 13: RU Grocery shop's example surrounding places (1km)

10.1.6 United States of America (US)

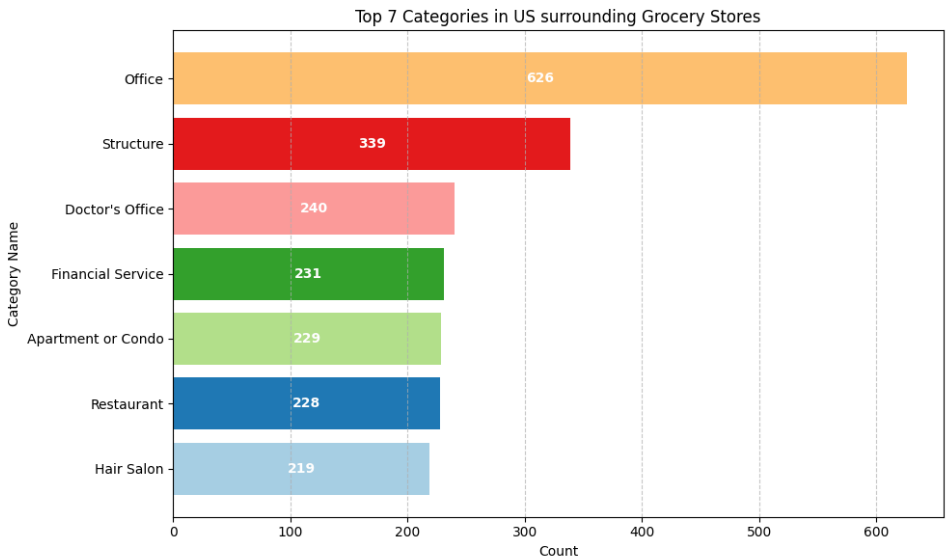


Figure 14: Top 7 categories in US surrounding Grocery stores

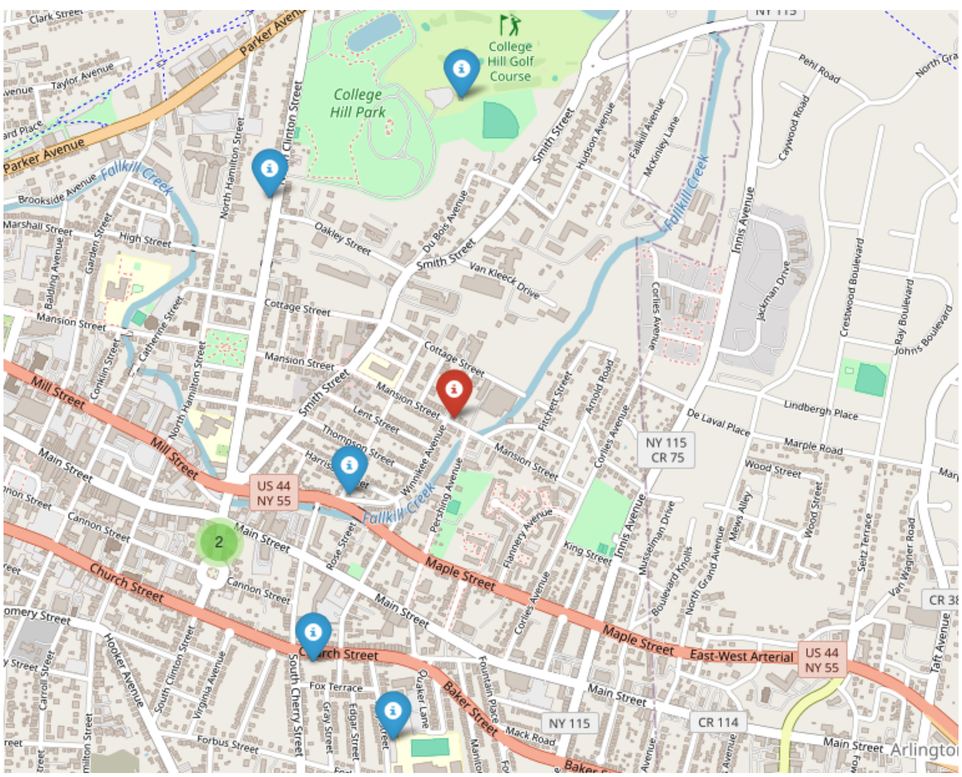


Figure 15: US Grocery shop's example surrounding places (1km)

10.1.7 South Africa (ZA)

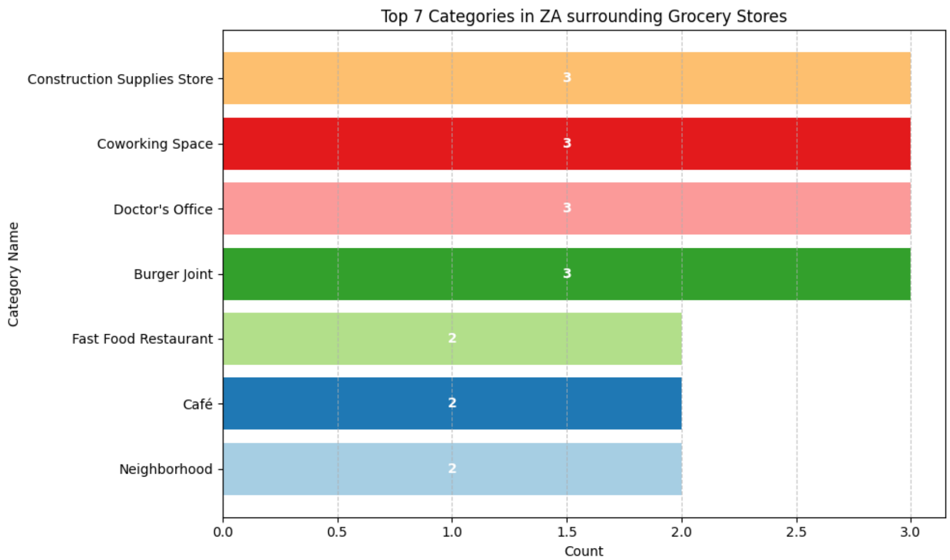


Figure 16: Top 7 categories in ZA surrounding Grocery stores

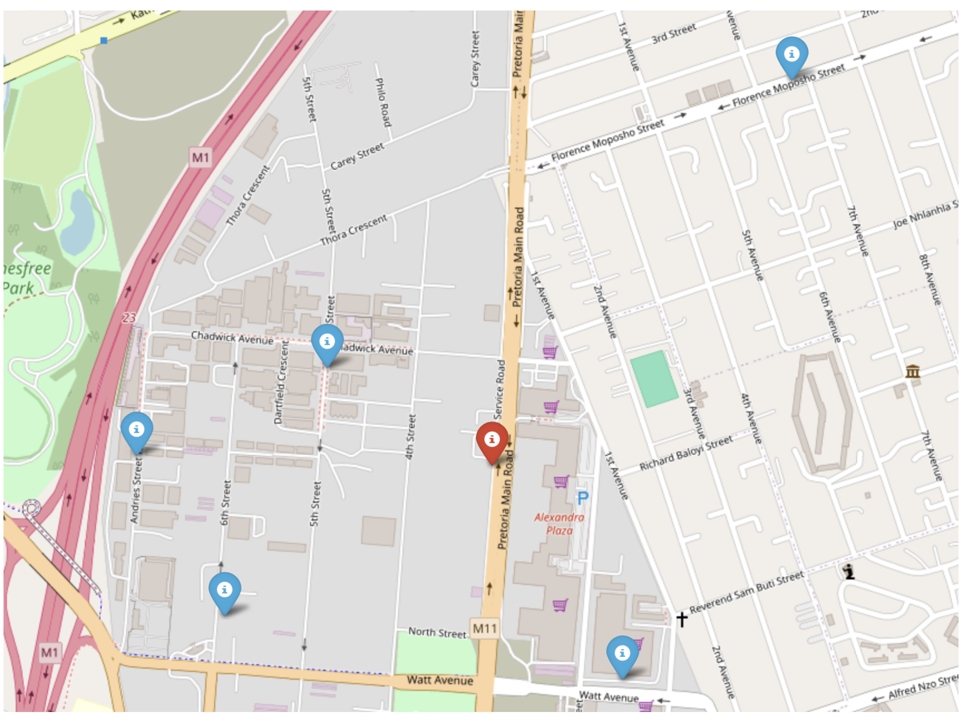
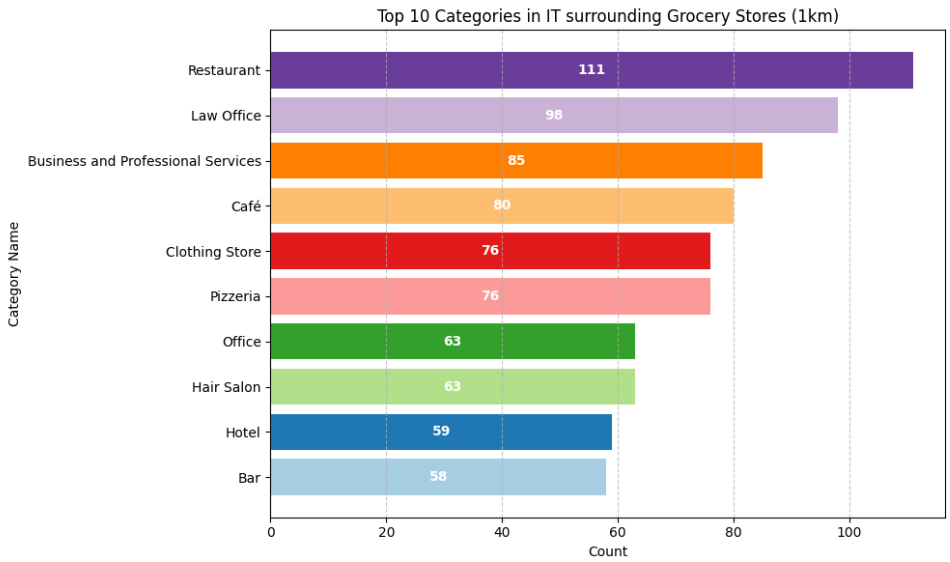
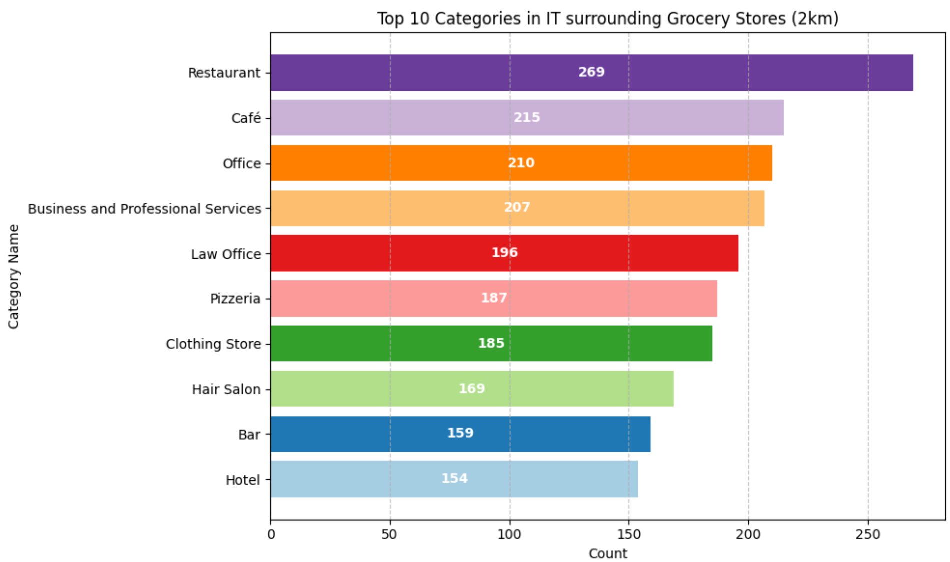


Figure 17: ZA Grocery shop's example surrounding places (1km)

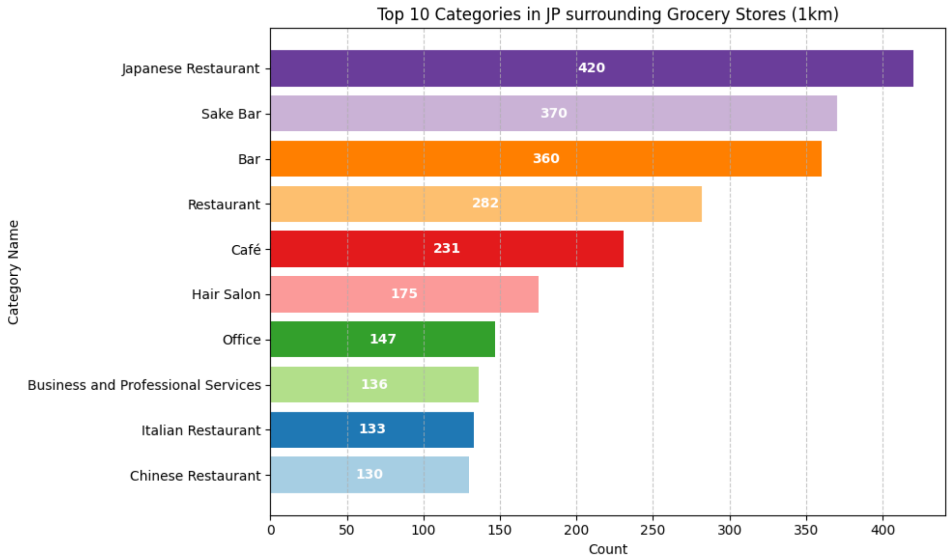
10.2 Appendix B: Additional Plots Distance

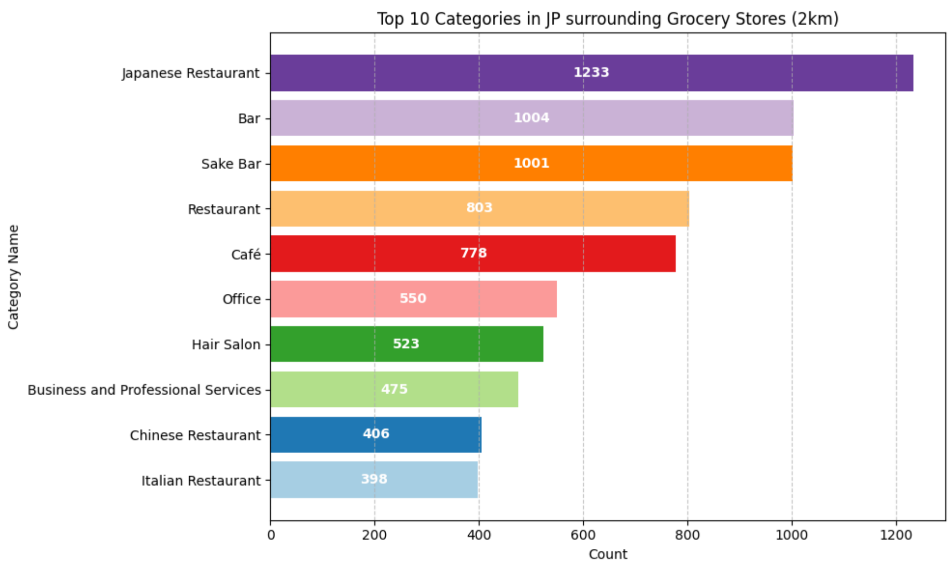
10.2.1 Italy



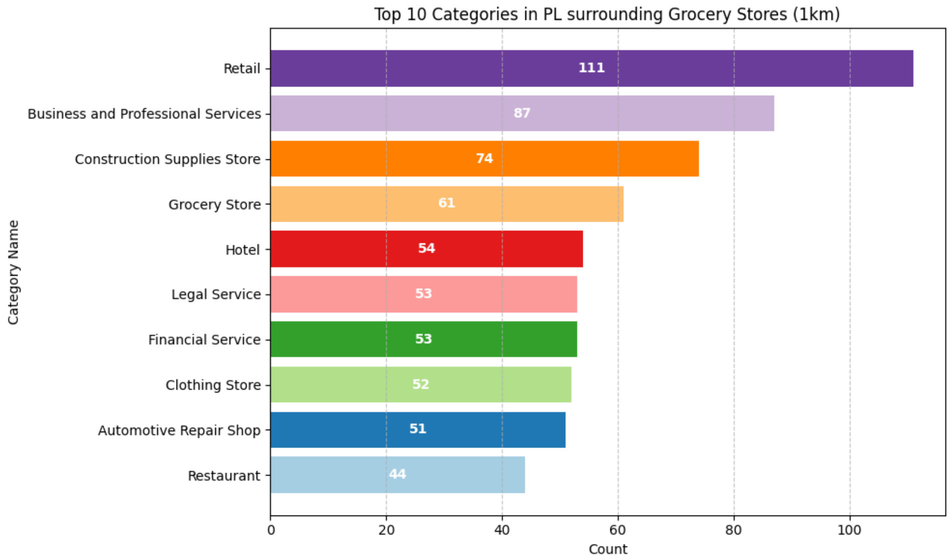


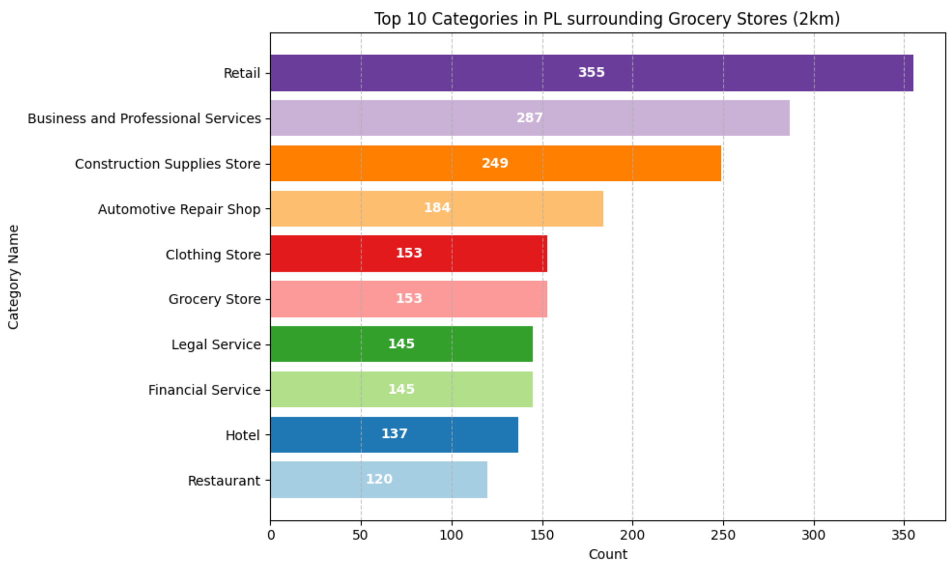
10.2.2 Japan (JP)



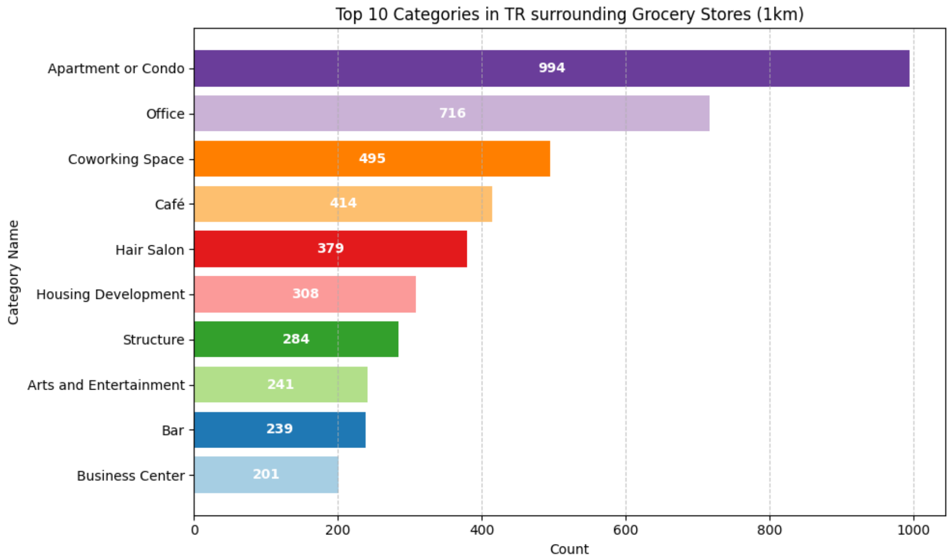


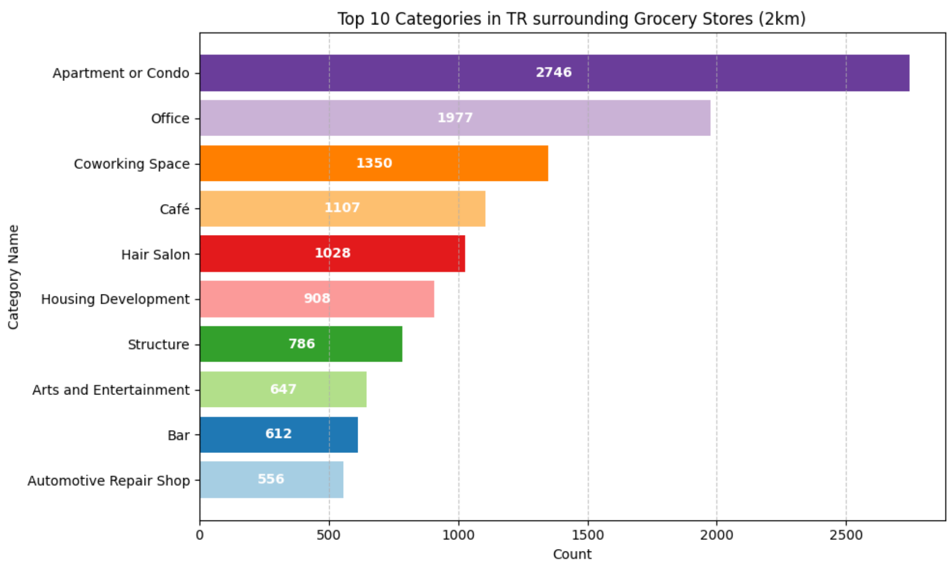
10.2.3 Poland (PL)



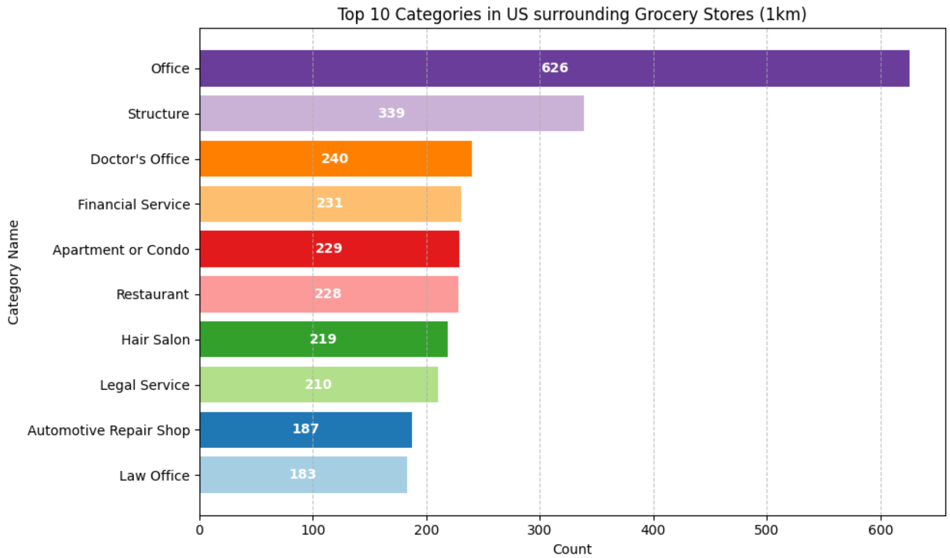


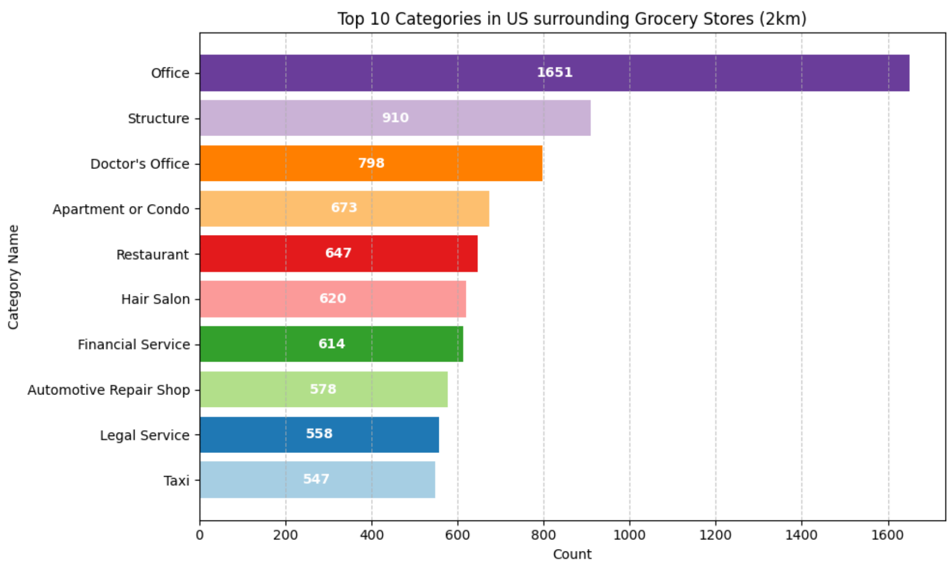
10.2.4 Turkey (TR)



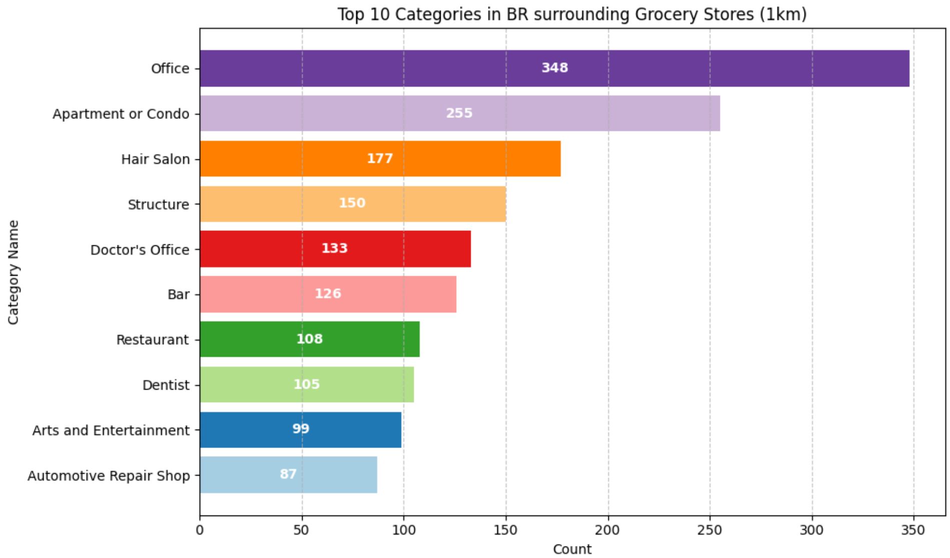


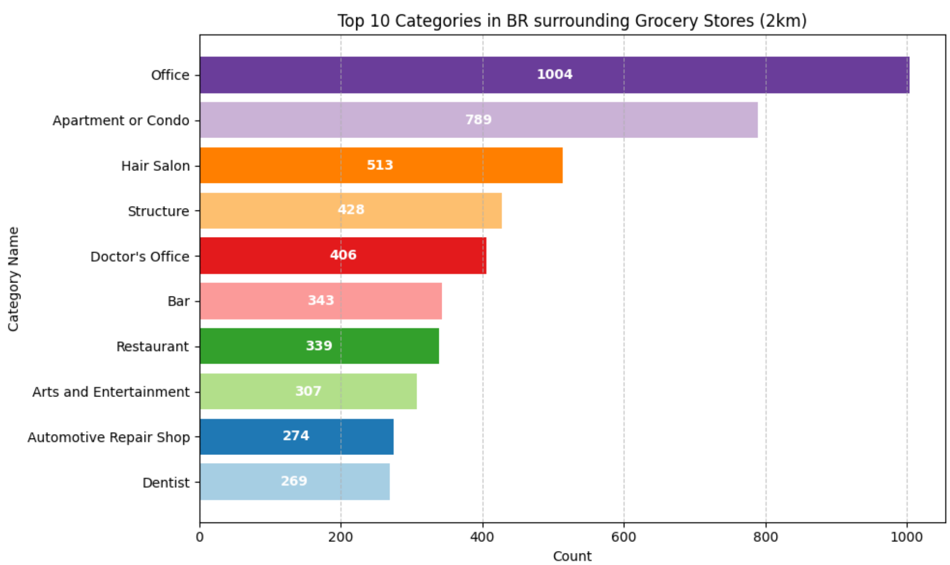
10.2.5 United States (US)



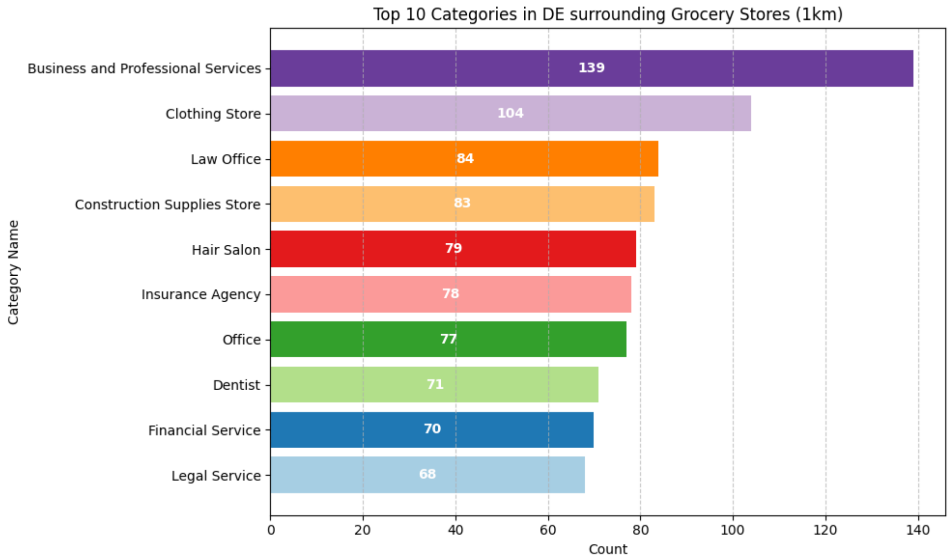


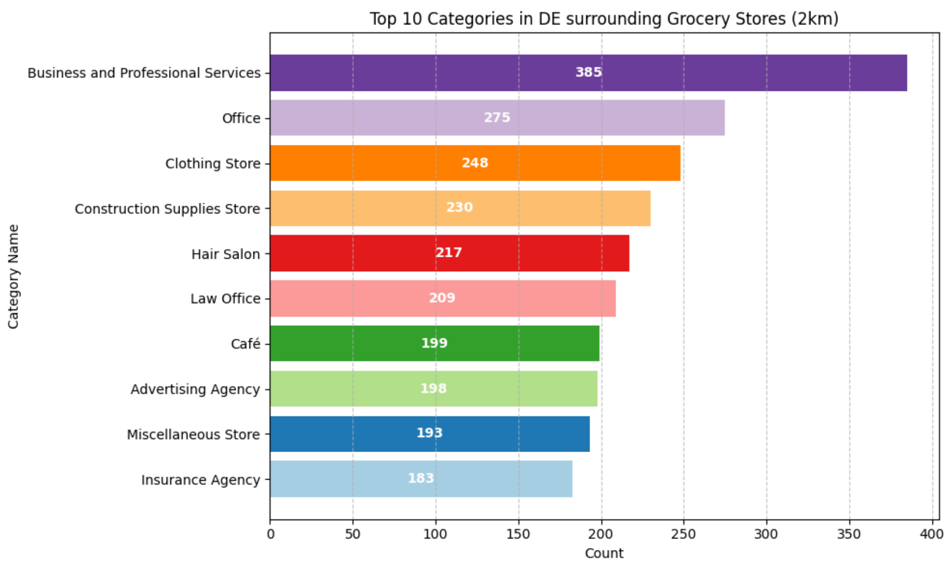
10.2.6 Brazil (BR)



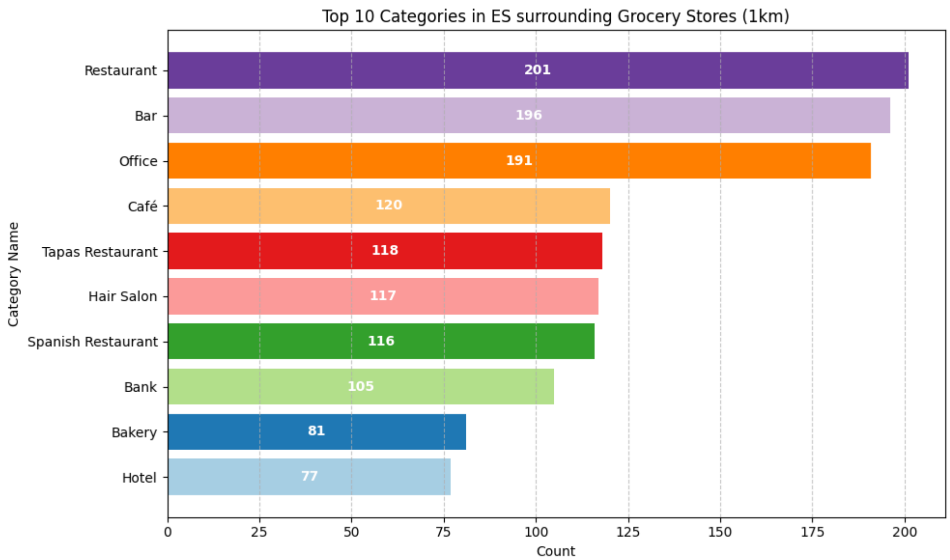


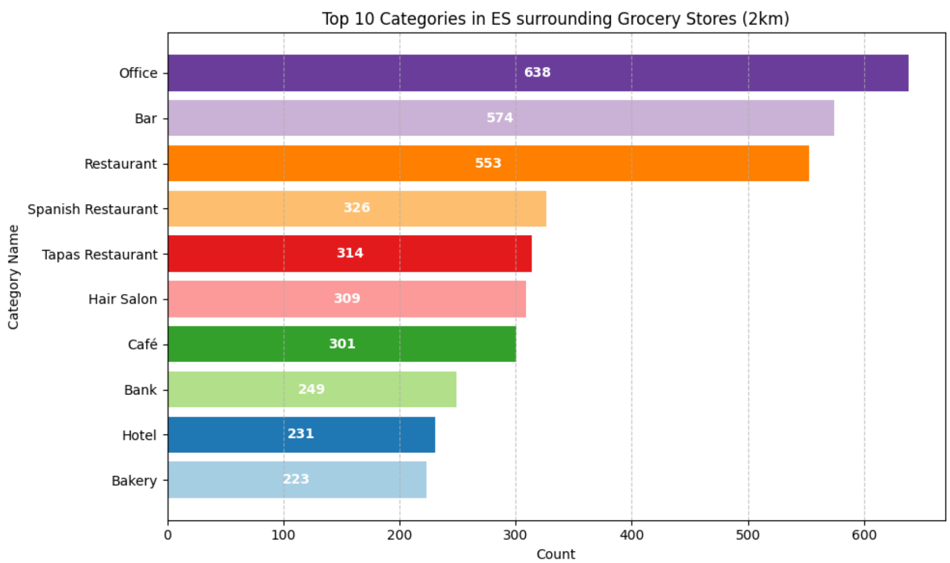
10.2.7 Germany (DE)



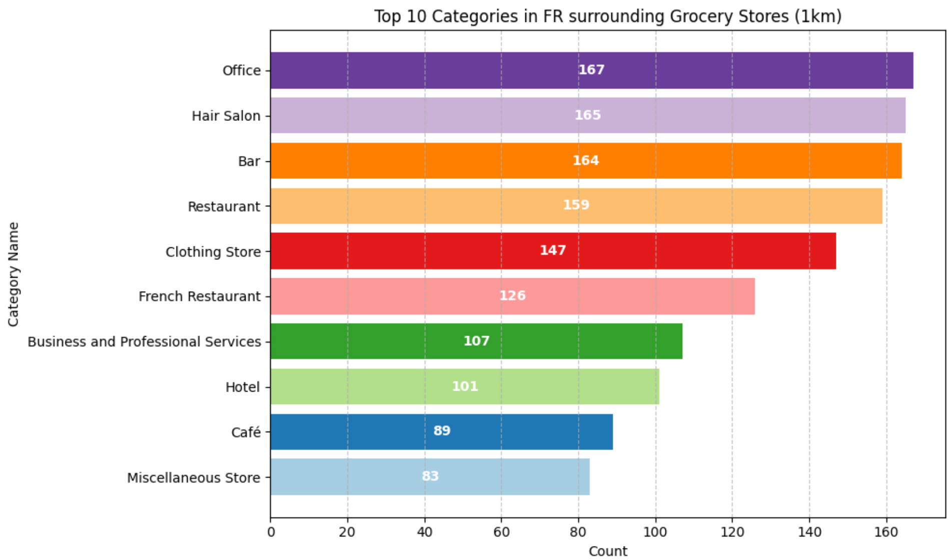


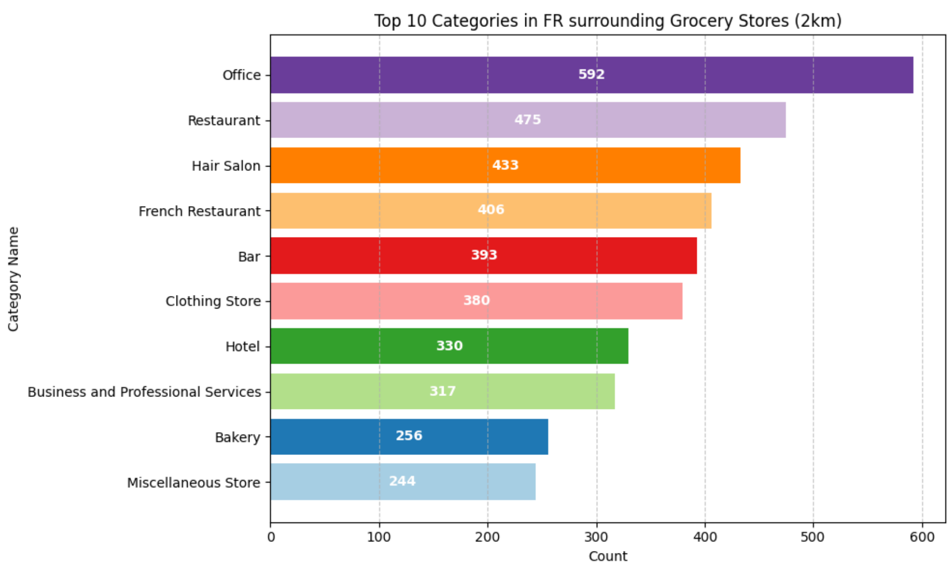
10.2.8 Spain (ES)





10.2.9 France (FR)





10.2.10 United Kingdom (GB)