1. **Introduction**
   1. **Background**

Bratislava is currently the house of the back office of some of the biggest IT and professional service companies in the world. Companies like IBM, Amazon, Lenovo or Dell have located European centers in this city. Internet and the location of the city in the center of Europe are without any doubt two of the most important factors that have facilitated the centralization of their internal departments in Bratislava. To be able to absorb all the demand of business centers from these companies, several of them have been built in the neighborhoods around the city center in the last decade.

Workplace environment is one of the factors that affects the most to the employees’ happiness and productivity. Therefore, these companies would find added value if they look for the best locations for their employees, in terms of facilities and venues that would create a good working environment for them.

* 1. **Problem**

From the workers’ perspective, the variety of restaurants, parks and coffee shops venues around the office can clearly increase the willingness to move to a new environment. The existence of bus and tram stops in the surroundings is also a good point. As one of the requirements for this report is the use of Foursquare location data, the report will be focused on comparing the different Bratislava’s neighborhoods in terms of the amount of venues related to the categories mentioned above.

The goal of the study is to make a ranking of the neighborhoods in Bratislava based on the frequency of bus and tram stops, outdoor, food and shop & service venues to see which are the most suitable ones to locate an office from this perspective. It will be focused on those neighborhoods where there are business center locations based on the Foursquare data.

* 1. **Interest**

These companies would be interested on the details about the different neighborhoods, as picking the best location for their offices would increase their workforce productivity and the willingness of potential future employees to move to their company.

1. **Data cleaning**

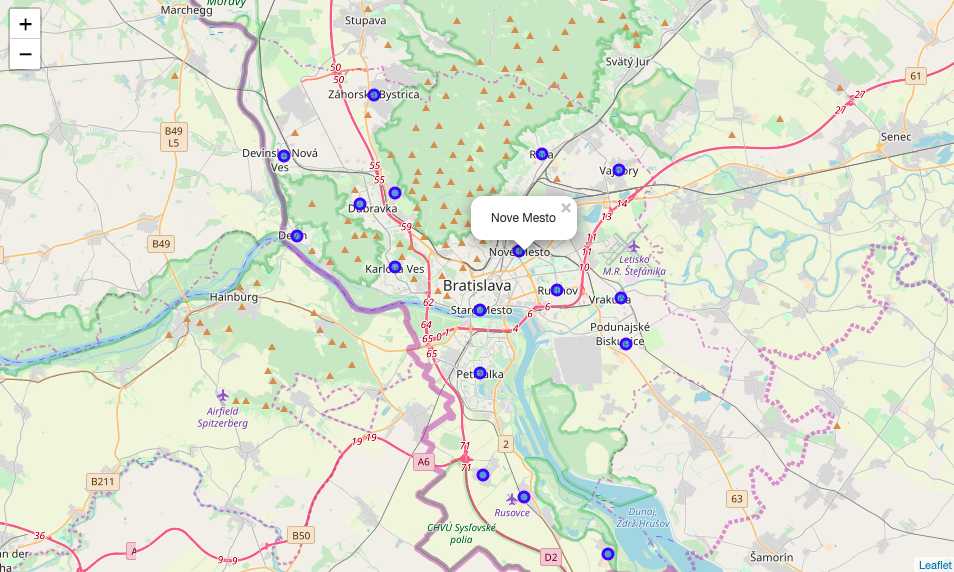
**2.1 Data sources**

For this analysis we need the coordinates from each neighborhood, which were taken manually from Wikipedia. They were taken one by one from their respective webpages and copied them into a CSV file. Apart from this source, location and further information on venues is taken from the Foursquare API.

**2.2 Data cleaning**

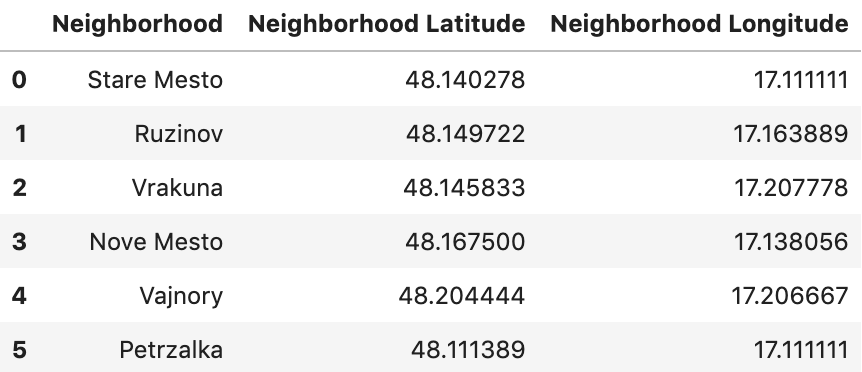
Our first task will be to upload the coordinates for the center of each neighborhood from the CSV file to the notebook. With them and a radius around them, we will create the request to the Foursquare API to get the nearby venues. Those coordinates must be in decimal format to introduce them in a map through the *folium* function. For this purpose, some functions are defined to make the calculations on our behalf and create a dataframe with the neighborhoods and their coordinates on decimal format. Figure 1 shows the coordinates of the center of each neighborhood superimposed over the map of Bratislava.

Figure 1. Map of Bratislava’s neighborhoods



Once the coordinates are in the proper format, we access the Foursquare API. Another function is defined to get the nearby venues. The API request is limited to venues under business centers only, providing a list of neighborhoods attached to them. It is important to highlight that not all business centers in Bratislava were inserted in the Foursquare database. However, only data from the Foursquare API is used for this report. Table 1 shows the list of neighborhoods under the scope of this report.

Table 1. List of neighborhoods



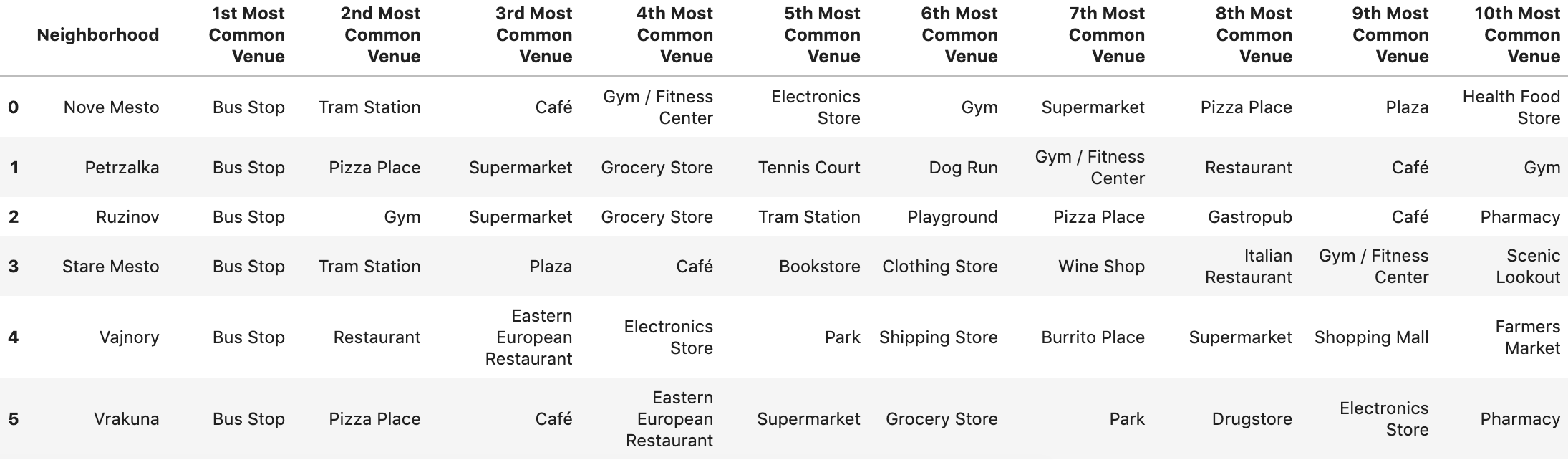
With this list of neighborhoods, a second API request retrieved all venues for those neighborhoods under categories of bus/tram stops, food, outdoor and shop & service. A total of 590 venues were retrieved under these categories only. These venues were grouped by neighborhood (Table 2) and a table of frequencies of each venue category was created.

Table 2. Total number of venues per neighborhood



Table 3 shows the 10 most common venue categories per neighborhood.

Table 3. Top 10 venues per neighborhood



From this view, we can see that bus stops and tram stations are among the most common venues in Nove Mesto, Ruzinov, Stare Mesto, meaning that there are good means of transport in those neighborhoods.

The following list shows the top venues by category, leaving aside bus stops and tram stations:

* **Nove Mesto**: 2 food categories, 3 shop and service categories, 3 outdoor and leisure categories.
* **Petrzalka**: 3 food categories, 2 shop and service categories, 4 outdoor and leisure categories.
* **Ruzinov**: 3 food categories, 3 shop and service categories, 2 outdoor and leisure categories.
* **Stare Mesto**: 2 food categories, 3 shop and service categories, 3 outdoor categories.
* **Vajnory**: 3 food categories, 5 shop and service categories, 1 outdoor category.
* **Vrakuna**: 3 food categories, 5 shop and service categories, 1 outdoor category.

Although the nature of the venues inside each category differs, the split of the top 10 venues is more or less the same in all neighborhoods, and Vajnory and Vrakuna show a different trend. Furthermore, the variety inside each category has to be higher in Nove Mesto, Petrzalka, Ruzinov and Stare Mesto, which more than double the amount of total venues of the other two locations.

1. **Methodology**

In this project we will focus now on finding those neighborhoods in Bratislava that are more suitable to open an office from the point of view of the kind of venues that can be found in them.

During the data gathering step, we retrieved all venues under the categories relevant for our case from the Foursquare API and created a frequencies table of venue categories per neighborhood. It is important to mention that this frequency is calculated as the amount of venues per category in each neighborhood over the total number of venues in that neighborhood. This means that it is not taking into consideration the differences in total venues between neighborhoods. However, the results shown in the next section also consider this fact in their reasoning.

In the analysis phase, two different clustering methods are used. These methods split our list of neighborhoods into smaller groups that have similar characteristics, which in this case means having venues in a similar frequency.

We do not want to pre-set the number of clusters, so we have used the hierarchical clustering method. It builds a hierarchy of nodes based on the distance between clusters, calculated through the Euclidean distance between points and visualized in a dendrogram.

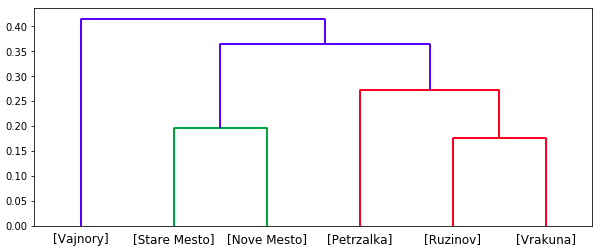
The second method will be the k-means clustering method, which needs the number of clusters as input. We will apply the number of clusters that was chosen as optimal in the hierarchical clustering method.

1. **Analysis**

The hierarchical clustering method was developed using the *scipy* package. The complete method was used to calculate the distance between clusters. It has provided an optimal number of clusters of 3 (Figure 2):

* Vajnory, considered as part of the countryside of Bratislava, is alone in the first cluster. It is the neighborhood with the lowest amount of venues, with only 23 in total.
* Stare Mesto and Nove Mesto, which form the city center of Bratislava in the northern side of the Danube, are in the second cluster.
* Petrzalka on southern side of the Danube and Ruzinov and Vrakuna on the Eastern part of the city, the bedroom communities of Bratislava, and are in the third cluster.

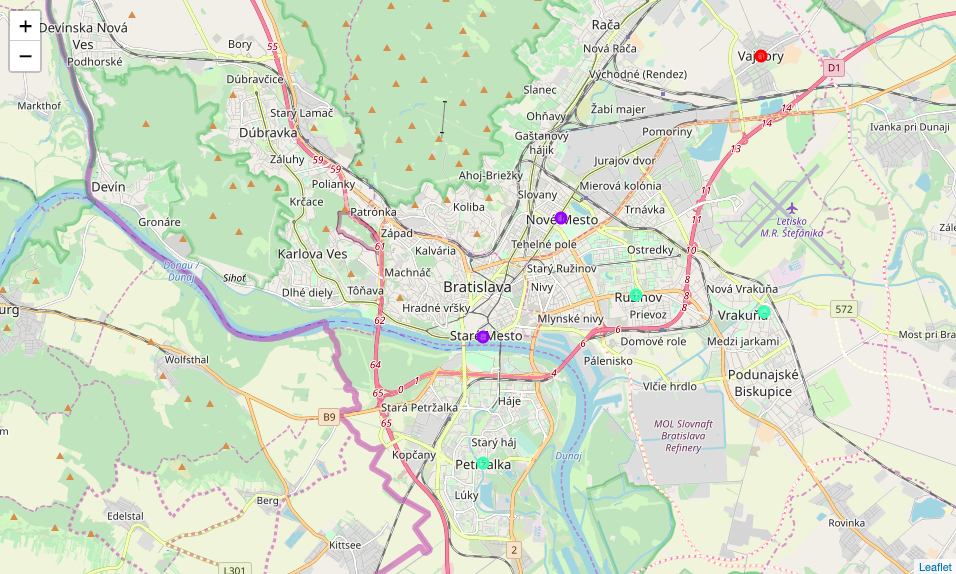
Figure 2. Dendrogram from the hierarchical clustering method



Now we will do the same analysis with the k-means clustering method, giving as input the optimal number of clusters used in the hierarchical clustering method, 3. The target is to confirm whether it provides the same split between neighborhoods than the hierarchical method.

Figure 3 in the next page shows the map of Bratislava with the coordinates of the center of each neighborhood, colored with different colors which represent the cluster to which they belong. Applying the same number of clusters, both methods provide the same split between clusters.

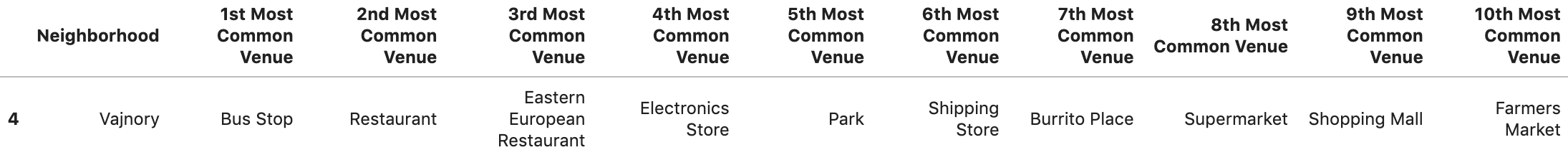
Figure 3. Map of clusters through k-means clustering method



* 1. **Examining clusters**
     1. **Vajnory: countryside neighborhood**

Vajnory only contains 23 venues in Foursquare (the smallest amount among all neighborhoods) in a radius of 1.6 km from its center. The lack of variety of venues and their nature explain why it was included in a separate cluster. A clear example of this is that it is the only location which contains farmers markets among its most common venues, making clear that this is a countryside neighborhood (Table 4).

Table 4. Vajnory cluster



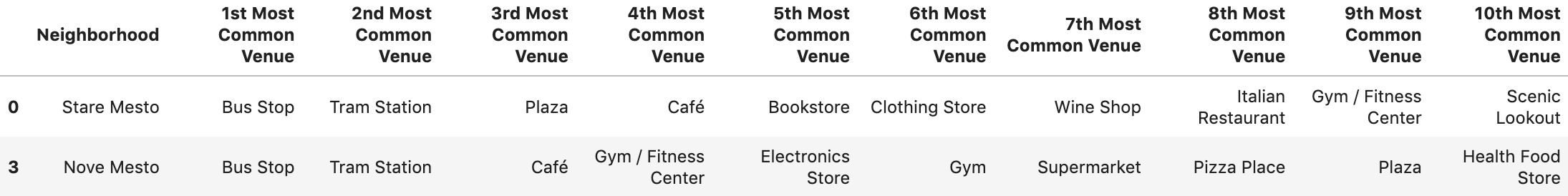
* + 1. **Cluster of city center neighborhoods**

This second cluster includes the two neighborhoods with the highest amount of venues Stare Mesto and Nove Mesto, or “old town” and “new town” in English.

They have in common that bus and tram stops are their most common venues. Cafés and plazas are between their most common venues. Stare Mesto also shows scenic lookout venues among its top 10 venues, demonstrating that it is the tourist part of the city. This also means more people on the streets, more noise and higher prices in all shops and restaurants in the surroundings, which is not the best combination for an office location.

Italian food is also among their most common venues. In Stare Mesto we can find Italian restaurants, while in Nove Mesto it is more common to find fast-food Italian restaurants, called pizza places in the Foursquare API. Both show a broad offer of gyms and other stores that could be visited by workers in their spare time.

Table 5. City Center cluster



* + 1. **Cluster of dormitory suburbs**

These three neighborhoods are dormitory suburbs of Bratislava. Therefore, all of them share the same kind of venues in their surroundings. That could explain why they are together in one cluster.

It should be highlighted that Vrakuna, with only 61 venues (in the same 1.6 km searching range used for the other neighborhoods), is on the same cluster than Ruzinov and Petrzalka, with twice as much venues. This is explained by the fact mentioned in the methodology, as these clusters were set based on the frequency of venues, not considering the differences in the total amount of venues per neighborhood. They have in common:

* Their first most common venue is bus stops, but Vrakuna and Petrzalka do not have tram stops around their territory yet.
* Fast-food pizza places are the second most common venue in Vrakuna and Petrzalka.
* They all have supermarkets and grocery stores in their top 6 venues.

Table 6. Dormitory suburbs cluster



1. **Results and discussion**

**Vajnory** would be the first neighborhood in being discarded as suitable location. A countryside location is not the most appropriate for an office. Based on the information acquired in this report, the main reasons are that:

* it does not have a proper variety of restaurants and other venues, where workers would spend their spare time, and that
* it is far from the city center by any mean of transportation, which implies long commutes for its workers.

The second neighborhood being excluded is **Vrakuna**. It is part of the dormitory suburbs, but it is more frequent to see single family homes than blocks of flats. It would be complicated to find a block of offices there, and the variety of venues is not enough, with only 61 venues in total.

Due to the lack of tram, meaning worse communication with the city center, **Petrzalka** would also be excluded. It is another dormitory suburb with a lot of potential in the long term, with projects to build new tram lines around this neighborhood. However, these lines are not ready, so this location is not ready for new business centers in the surroundings (excluding the territory located on the Danube riverside, in front of Stare Mesto, where several business centers have been built).

Therefore, from the analysis provided by this report, the top 3 neighborhoods to locate an office would be Ruzinov, Stare Mesto and Nove Mesto.

* Ruzinov is the only dormitory suburb that could be a good location due to its good connections to the city center, and the fact that it is close to residential zones. This could be a good point for workers, who could live closer to the office. Its real estate offer is bigger and prices are lower than in city center. It offers bus but also tram facilities, although not in the same frequency as in the other two locations.
* Stare Mesto is the tourist zone of the city, with several restaurants and varied cafés. It is the neighborhood with the biggest amount of venues, offering also clothing shops, supermarkets, gyms and several plazas with nice terraces for a drink after work. However, in tourist periods, it is always crowded due to the monuments and sightseeing locations located in it.

In overall, from my own experience in this city, this location should not be recommended for an office, although I consider that locations nearby this neighborhood, close to all those venues but enjoying less crowded parks and plazas, are the best locations in Bratislava for an office.

Based on the analysis of the data contained in this report, Nove Mesto would be the best location for the office:

* It has the same variety of venues than Stare Mesto, but at lower prices.
* Together with Stare Mesto, it is the best communicated with bus and tram with all other neighborhoods, which would reduce the commuting time for workers.
* Close to the old town and with good communications with it, which makes it easier workers to visit its shops and cafés.
* Other venues like gyms and supermarkets in the surroundings of the office, highly valued by the average worker.
* Plazas among their most common venues, with outdoor and some green spaces for the spare time.

Nonetheless, recommended zones should therefore be considered only as a starting point for more detailed analysis which could eventually result in a location which fulfills the requirements in terms of variety of venues but also other factors taken into account and all other relevant conditions for the stakeholder. The results above are just a choice based on the analysis provided in the previous section, but it is not the only possible solution, as depending on how the user values each of the items mentioned, the ranking of neighborhoods can differ.

1. **Conclusion**

Purpose of this project was to identify the best locations to open an office in Bratislava, based on the variety of food, shop&service and outdoor venues and the frequency of bus/tram stops existing in the surroundings of those locations.

By extracting all business centers in Bratislava from the Foursquare API, the list of neighborhoods was limited to those which include business centers. Then, all venues from those neighborhoods were retrieved, calculating then the frequency of each venue category in each neighborhood.

This calculation allowed us to make comparisons between them. Clustering of those locations based on those frequencies was then performed in order to create major zones of interest (containing more interesting venues from the point of view of workers and people who would commute to the office on daily basis) for final exploration by stakeholders.

Finally, stakeholders will decide which is the optimal office location, based on the aspects analyzed in this report and other characteristics of neighborhoods not mentioned here, like real state availability or price per square meter.