BATTLE OF NEIGHBORHOODS (MIAMI) - WEEK 2

APPLIED DATA SCIENCE CAPSTONE

This report elaborates on the Battle of Neighborhoods in Miami, which constitutes a Coursera Capstone Project geared towards extracting insights from data using data science tools, such as Data Analysis with Python, Data Visualization, Machine Learning, etc. The project is further discussed in the following sections.

1. Introduction

The Battle of Neighborhoods is a Coursera Capstone project that seeks to use Foursquare location data, in combination with other selected data, to develop a business case around the suitability of a particular neighborhood for opening a new Restaurant in a given City. The city selected for this endeavor was Miami, well-known for being a touristic magnet, and, therefore, a good opportunity for the restaurant industry, which seeks to consistently respond to a high demand of food, while maintaining a steady operation in a safe neighborhood. This goal can be approached by gathering and processing enough information towards decision making. This information is often not given, thus requires structured and/or intensive searches, depending on the availability of the required information. In this regard, the recourse to Foursquare to obtain ample information about the venues in a specific location is a good starting point, but this information should be processed adequately to segregate the signal from the noise. Furthermore, this information can be also be complemented by additional information, when required, to gather enough insights prior to providing a constructive feedback to a client.

The client in this project could be any corporation or individual willing to invest in the restaurant industry.

2. Data used and methodology

2.1 Methodology

According to FSR Magazine in this issue:https://www.fsrmagazine.com/expert-takes/8-factors-choosing-new-restaurant-location, the selection of a new restaurant location should be driven by 8 factors, including but not limited to:

- Visibility,
- Parking,
- Space size,
- Crime rates,
- Surrounding Businesses and Competitors analysis,
- Accessibility,
- Affordability, and
- Safety.

To this end, the following steps will be taken to determine which neighborhood wins the battle of neighborhoods:

1. To use web-scrapping to collect the list of neighborhoods in Miami. The web site to be used to achieve this objective is: https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Miami,
2. To use geolocation data and a Foursquare API to get the list of venues in these neighborhoods,
3. To filter the list of restaurants in each Neighborhood from the Foursquare result and group them in a dataframe using Pandas library of python,
4. To plot the number of restaurants in each neighborhood using matplotlib,
5. To plot the population of each neighborhood using matplotlib,
6. To identify from the plots (bar charts) of step 4 and 5 which neighborhood may lack restaurants (visibility, surrounding businesses and competitor analysis),
7. To use a linear regression to determine whether a correlation exists between population and number of restaurants,
8. To cluster each neighborhood of Miami and display them in a leaflet using K-Means Clustering and Folium (Parking),
9. To obtain the data pertaining to crime rate data in neighborhood of Miami (https://www.areavibes.com/miami-fl/most-dangerous-neighborhoods/) and plot it (Crime rates and Safety),
10. To obtain data relevant to Median Household income in each neighborhood of Miami, (https://www.homesnacks.net/richest-neighborhoods-in-miami-128997/) ,and plot it (Targeted clientele and Affordability), and
11. To suggest which neighborhood(s) is suitable for opening a new restaurant in Miami.

2.2 Data used

An essential tool for the execution of this project is Foursquare, which, through an API, enables the exploration of venues around a provided location. However, essential geolocation data must be compiled prior to requesting results from the exploration of venues using Foursquare. In this project, we use a Wikipedia page listing the neighborhood names of Miami, as well as their corresponding population and coordinates. Additionally, in order to get well acquainted to other essential parameters driving the selection of a suitable location for the creation of a new restaurant, other information relevant to the safety of neighborhoods was collected from this link: https://www.areavibes.com/miami-fl/most-dangerous-neighborhoods/; while the median household income in each neighborhood was collected from this link: https://www.homesnacks.net/richest-neighborhoods-in-miami-128997/. Thus, the data provided from these different sources will be analyzed with the sole objective of making the best selection.

3. Results and Discussion

From the results obtained from Foursquare tabulated as the number of restaurants in each neighborhood, it can be noticed that three neighborhoods of Miami do not have any major restaurant identified by Foursquare. These neighborhoods include Allapattah, Venetian Islands, and Viginia key. It was also noticed that Foursquare returned very few restaurants in neighborhoods such as Grapeland Heights, Coral Way, Coconut Grove, Overtown, and Liberty City. These neighborhoods pass the test of Surrounding Businesses and Competitor analysis, which relates to the chances of restaurant to evolve and to resist competition.

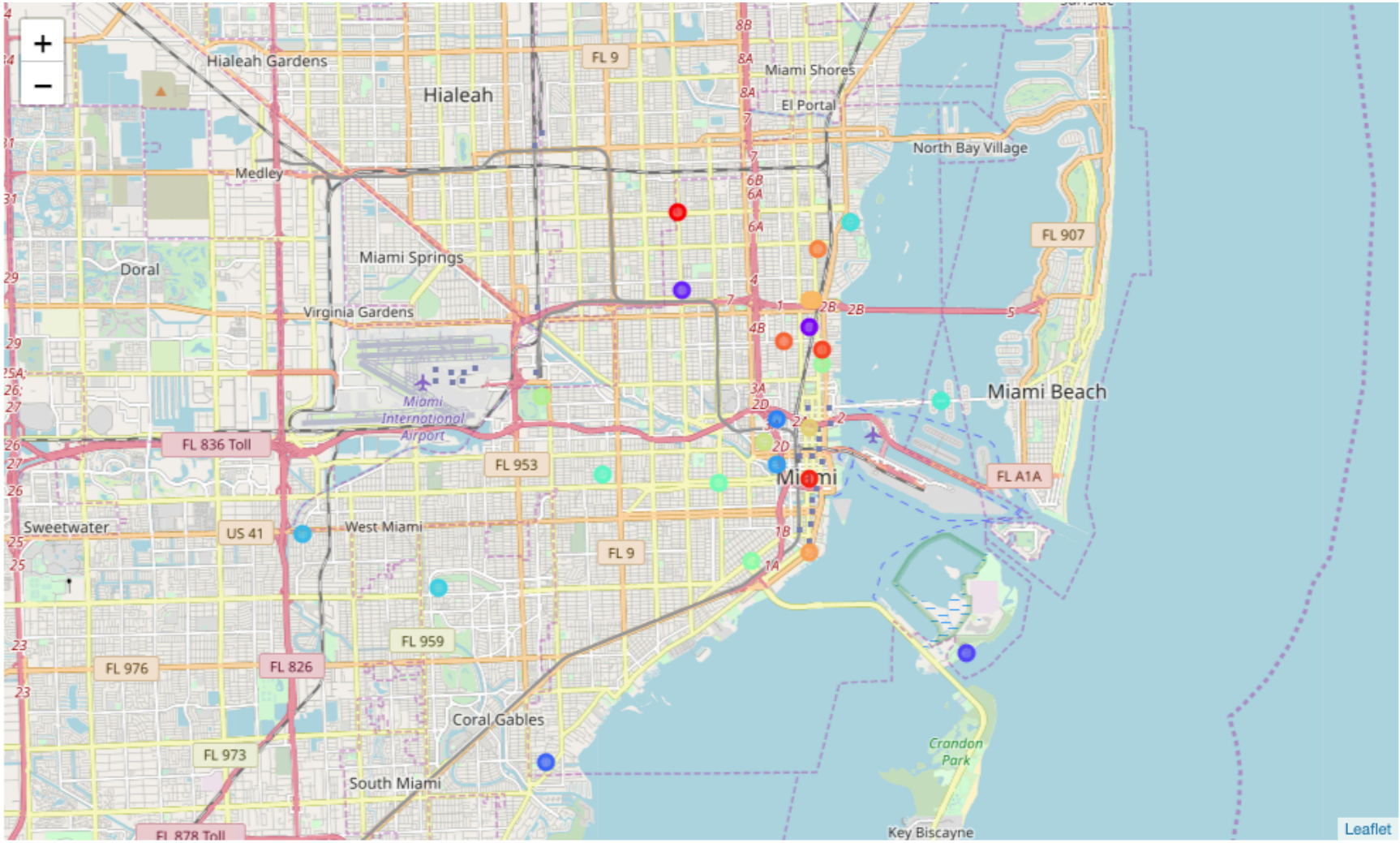

Figure 1: Leaflet of the neighborhoods of Miami

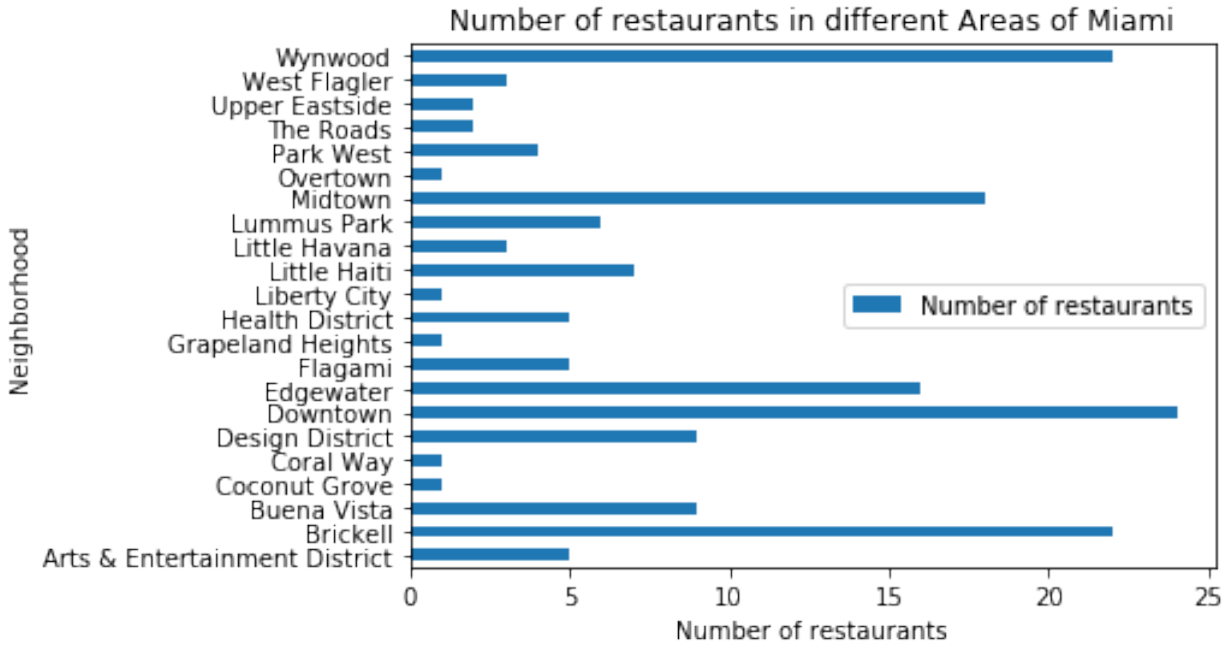

Figure 2: Bar Chart of number of restaurants in different neighborhoods of Miami

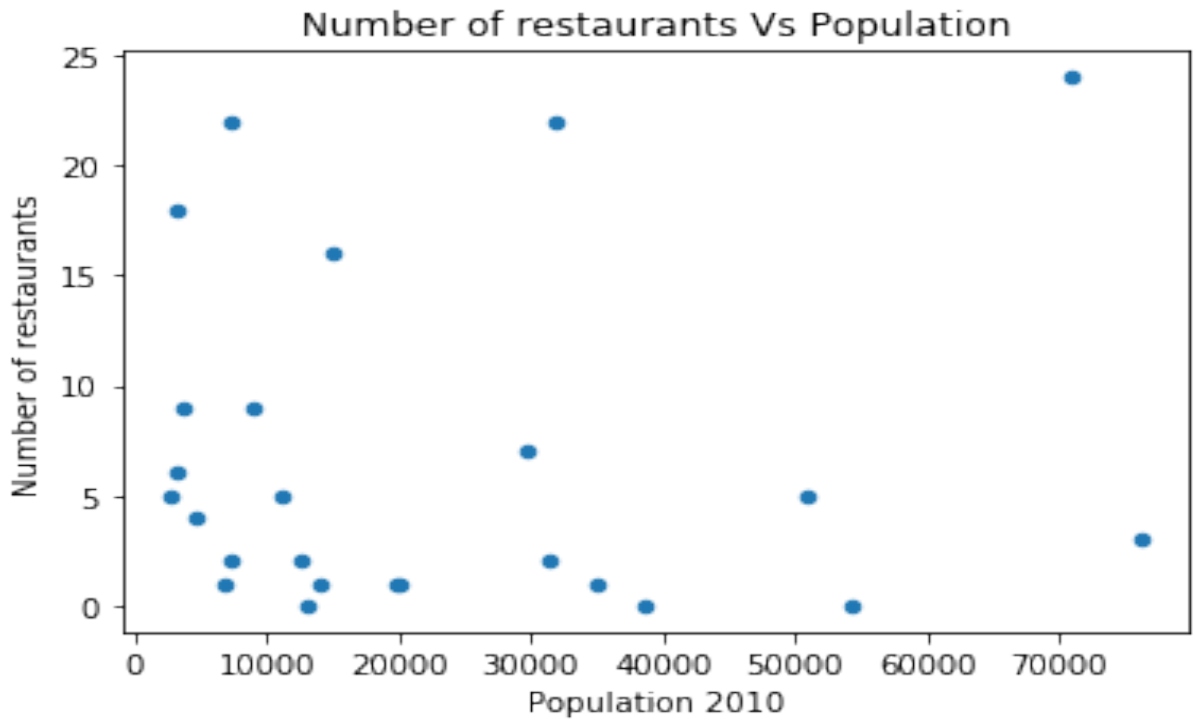

Figure 3: Scatter plot of the population vs the number of restaurants in different neighborhoods of Miami

An analysis of the most common restaurants around each location illustrated that the people living in the neighborhoods listed above could reach restaurants located in other neighborhoods. This was further illustrated while inspecting the proximity of these neighborhoods to one another through neighborhood clustering using K-means Clustering and Folium for graphical representation. This suggested that other analyses were required to get to consolidate this analysis. One approach that was used in this Notebook was a scatter plot, which usually enables to figure out whether there is a correlation between a dependent and an independent parameter. The dependent parameter was the number of restaurants, while the independent parameter was the population in each neighborhood. No correlation was found from this evaluation.

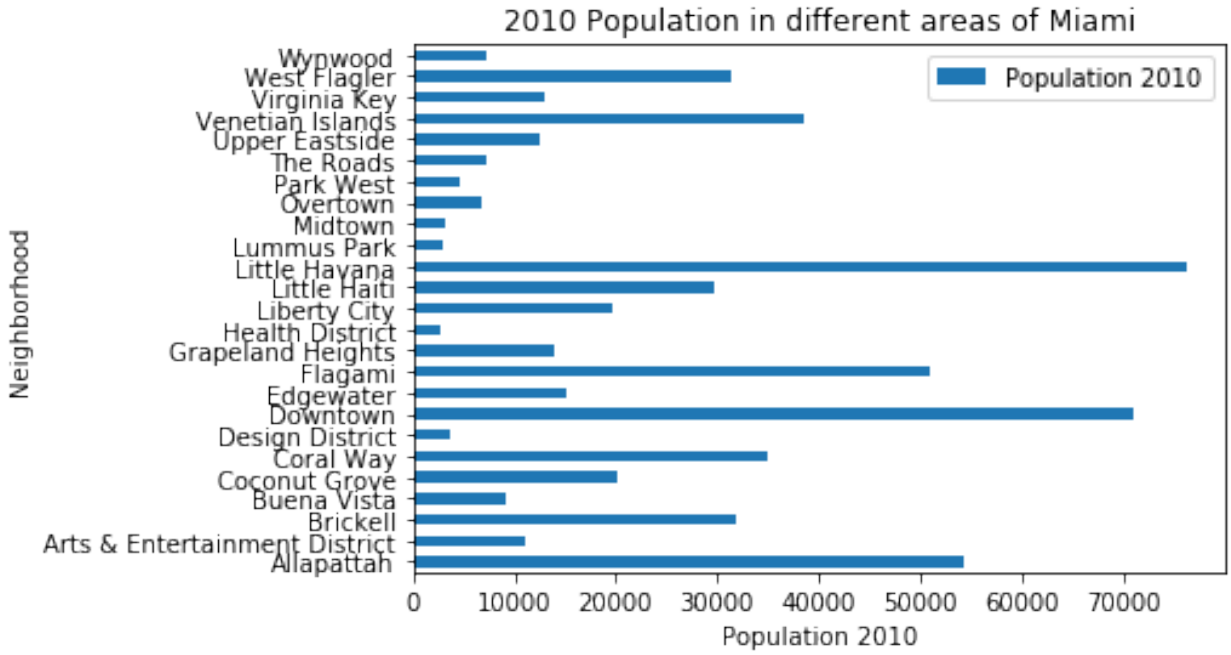

Figure 4: Population in different neighborhoods of Miami

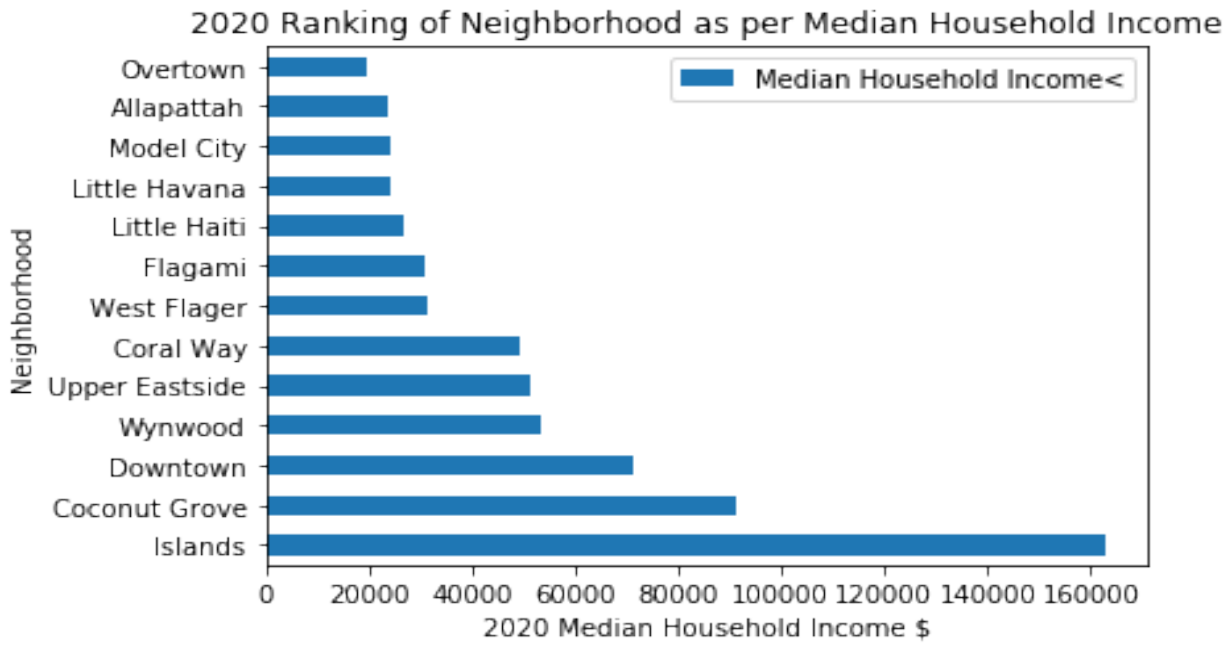

Figure 5: Median Household Income in different neighborhoods of Miami

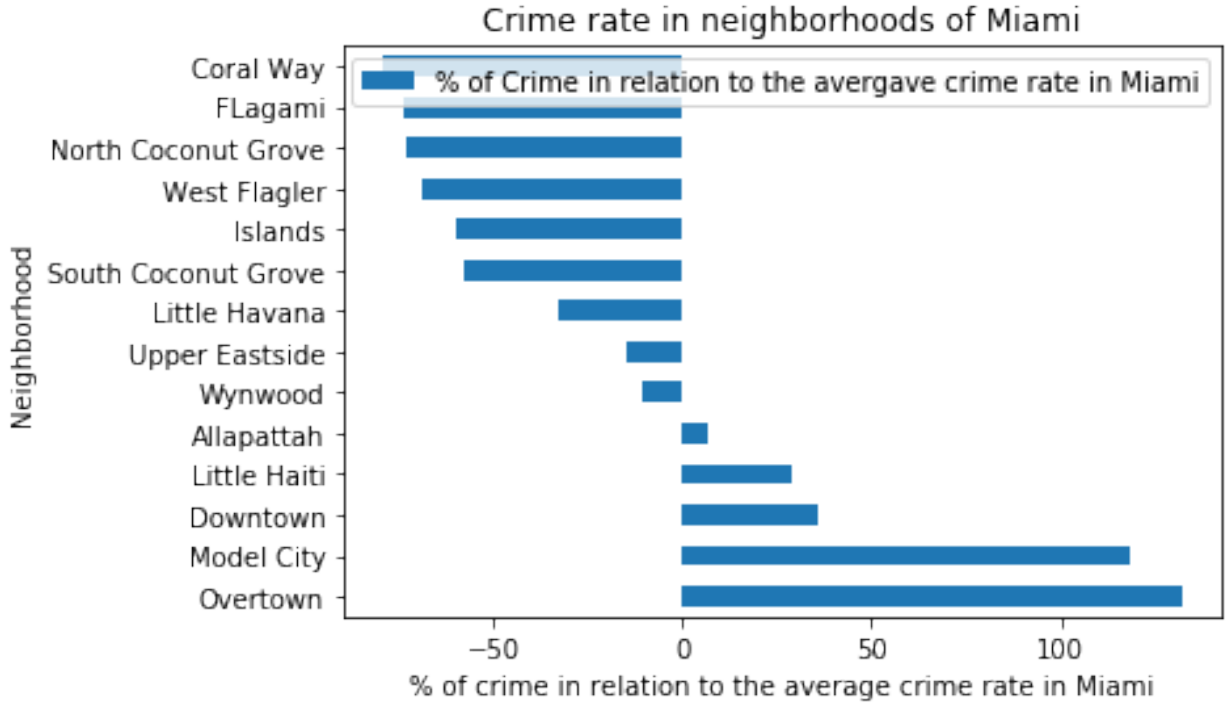

Figure 6: Crime rate in different neighborhoods of Miami

Subsequently, the next step in the analysis was to look at the number of people, the median income salary, and the crime rates in each neighborhood. From this analysis, it appears that Allapattah has a crime rate slightly above the average crime rate in Miami and higher than Venetian Islands, has a

big population with a median household income significantly lower than Venetian Islands. Viginia key's median household income or crime rate was not unveiled in these studies, but has a population significantly lower than Venetian Islands. Although the population of Venetian Islands is lower than the one of Allapatah, this neighborhood appears as the most attractive option for the development of a new restaurant in Miami.

4. Conclusion

In fine, from the analysis of the number of restaurants, the population, the location, the most common restaurants, the median household income, and the crime rate in each neighborhood, Venetian Islands appears as the most suitable neighborhood for the opening of a new restaurant in the city of Miami.