

# Opening an ice cream shop in Rome

Chiara Picardi, December 2020.

# Business problem

- **A client famous for his high quality ice cream (gelato) wants to open a new ice cream shop in Rome**
- **He is not sure in which neighborhood is better to open the shop due to the high competition**
- **In this project the neighborhoods of Rome will be examined in order to determine which are the neighborhoods with a low presence of ice cream shops and then less competition**
- **This project is useful for investors interested in opening an ice cream shops in Rome.**

# Data

- **Data are retrieved from wikipedia:**  
[https://en.wikipedia.org/wiki/Category:Quarters\\_of\\_Rome](https://en.wikipedia.org/wiki/Category:Quarters_of_Rome)
- **Data are preprocessed using web scraping techniques, Python requests and BeautifulSoup packages.**

| Neighborhood |  |
|--------------|--|
| 0            | ▶ Rome Q. XXIII Alessandrino (7 P)     |
| 1            | ▶ Rome Q. XXV Appio Claudio (1 C, 4 P) |
| 2            | ▶ Rome Q. IX Appio-Latino (18 P)       |
| 3            | ▶ Rome Q. XXVI Appio-Pignatelli (2 P)  |
| 4            | ▶ Rome Q. XX Ardeatino (13 P)          |

# Creating Dataset(1)

- The Python Geocoder package is used to obtain the longitude and the latitude for each neighborhood

|   | Neighborhood                           | Latitude | Longitude |
|---|--|----------|-----------|
| 0 | ► Rome Q. XXIII Alessandrino (7 P)     | 41.87139 | 12.57974  |
| 1 | ► Rome Q. XXV Appio Claudio (1 C, 4 P) | 41.98425 | 12.71421  |
| 2 | ► Rome Q. IX Appio-Latino (18 P)       | 41.87461 | 12.51333  |
| 3 | ► Rome Q. XXVI Appio-Pignatelli (2 P)  | 41.84326 | 12.54058  |
| 4 | ► Rome Q. XX Ardeatino (13 P)          | 41.53654 | 12.56257  |

# Creating Dataset(2)

- **Foursquare API** is used to retrieve information of the frequency of ice cream shops for each neighborhood

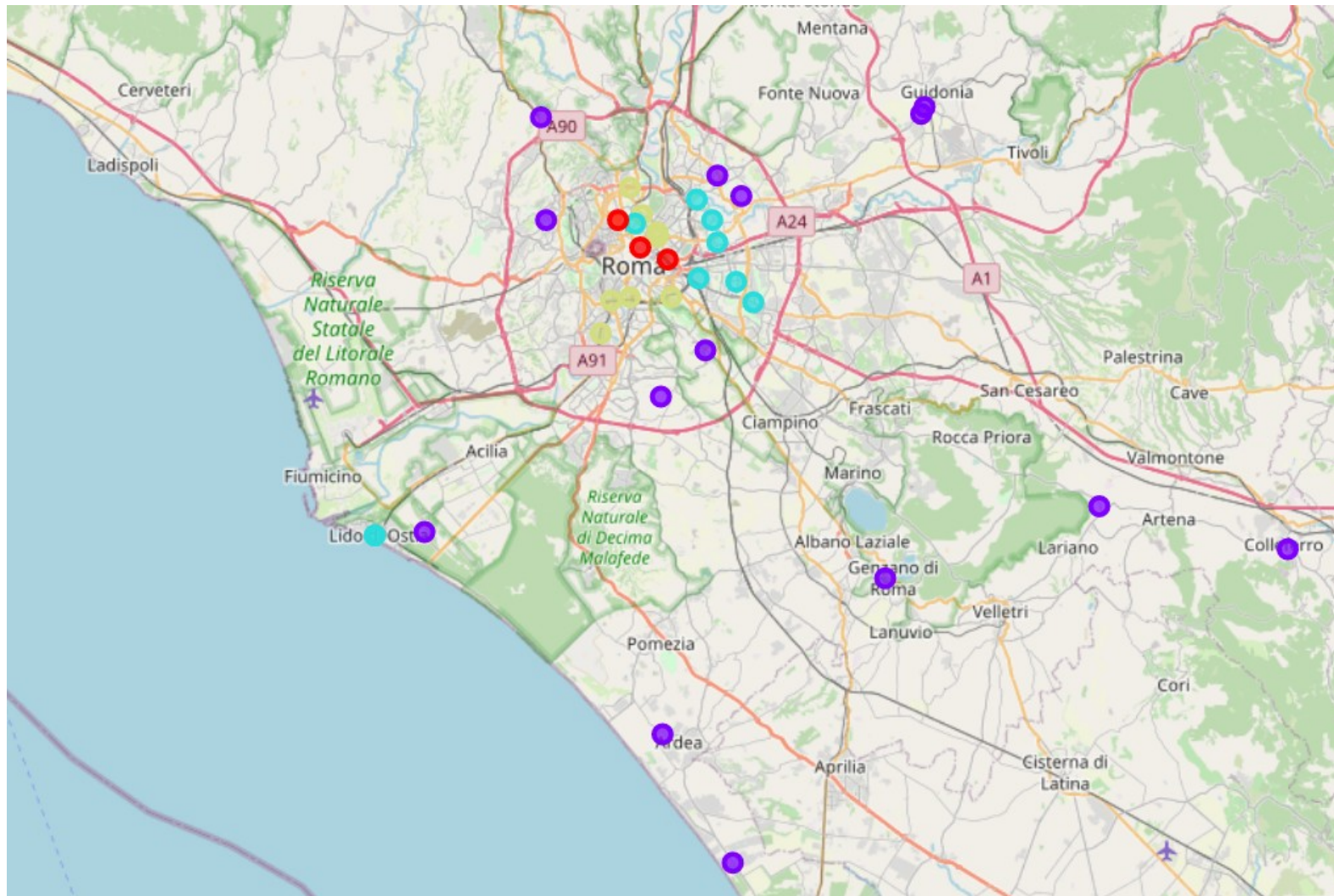
| Neighborhoods |                                    | Ice Cream Shop |
|---------------|------------------------------------|----------------|
| 0             | ► Rome Q. I Flaminio (18 P)        | 0.10           |
| 1             | ► Rome Q. II Parioli (16 P)        | 0.05           |
| 2             | ► Rome Q. III Pinciano (3 C, 20 P) | 0.07           |
| 3             | ► Rome Q. IV Salaria (5 P)         | 0.08           |
| 4             | ► Rome Q. IX Appio-Latino (18 P)   | 0.09           |

# Segmenting dataset

- **K-means algorithm is used to segment the dataset into four different classes using the frequency of ice cream shops**

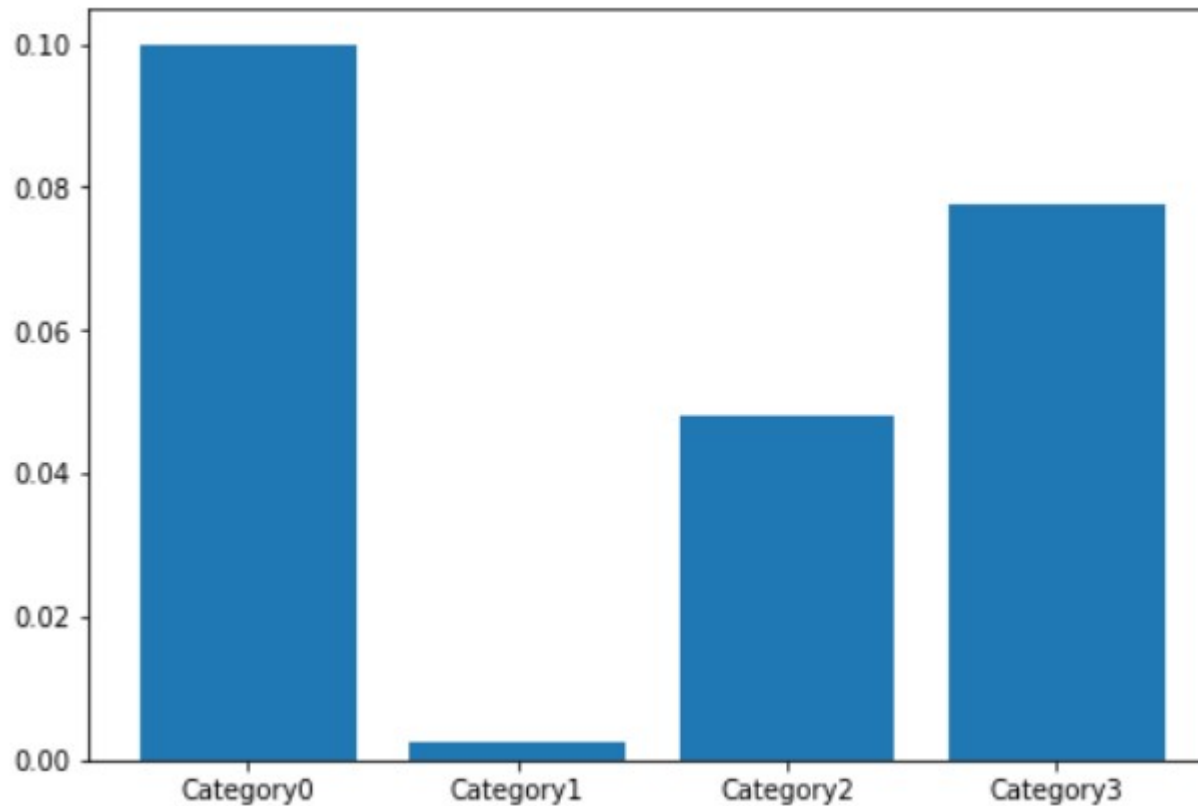
|   | Neighborhood                       | Ice Cream Shop | Category |
|---|------------------------------------|----------------|----------|
| 0 | ► Rome Q. I Flaminio (18 P)        | 0.10           | 0        |
| 1 | ► Rome Q. II Parioli (18 P)        | 0.07           | 3        |
| 2 | ► Rome Q. III Pinciano (3 C, 20 P) | 0.06           | 2        |
| 3 | ► Rome Q. IV Salaria (5 P)         | 0.08           | 3        |
| 4 | ► Rome Q. IX Appio-Latino (20 P)   | 0.08           | 3        |

# Results(1)



# Results(2)

Bar chart reporting mean ice cream shops frequency for each category





# Conclusion

- **Neighborhood belonging to category 1 have low competition but represent a risk because population is not used to this kind of business**
- **The best place to open the ice cream shop will be any neighborhood belonging to category 2**
- **Information about renting and buying prices or important landmarks are not considered**