



OPENING A NEW SHOPPING MALL IN ROME, ITALY

**Coursera Capstone
IBM Applied Data Science Capstone**

By: Nour Shosharah



INTRODUCTION :

Business problem :

- The importance of a shopping mall is mainly for people to get out of the house for a while and do something entertaining. Shopping malls can provide the best shopping experiences such as social gatherings, entertainment, performances, product launches, promotions and festivals. The events list at shopping malls goes on and on for any, particular, person to be entertained for a number of hours
- Location of the shopping mall is one of the most important decisions that will determine whether the mall will be a success or failure
- Business question : in the city of Rome , Italy if a property developer is looking to open a new shopping mall, where would you recommend that they open it?
- Objective : to analyze and select the best location in the city Rome , Italy to open a new shopping mall



DATA :

Data needed and sourced :

Data required :

- List of neighborhoods in Rome , Italy
- Latitude and longitude coordinates of those neighborhoods
- Venue data, particularly data related to shopping malls.

Data sourced:

- This Wikipedia page (https://en.wikipedia.org/wiki/Category:Subdivisions_of_Rome) contains a list of neighborhoods
- Python Geocoder package for latitude and longitude coordinates of the neighborhoods.
- Foursquare API (<https://foursquare.com/>) to get the venue data for those neighborhoods



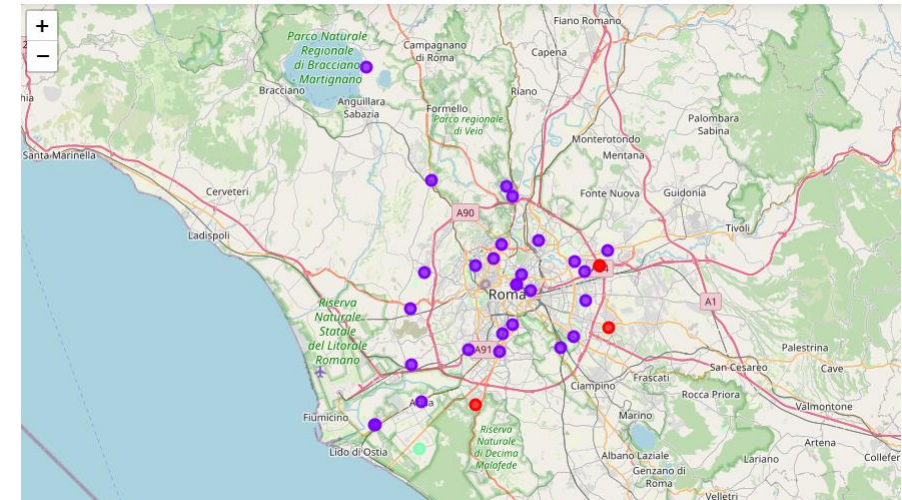
METHODOLOGY

- web scraping using Python requests and beautiful soup packages to extract the list of neighborhoods
- get the geographical coordinates in the form of latitude and longitude using Python Geocoder package
- Using Foursquare API (<https://foursquare.com/>) to get the venue data
- Analyze each neighborhood by grouping the rows by neighborhood and taking the mean of the frequency of occurrence of each venue category.
- Filter venue category by Shopping Mall
- perform clustering on the data by using k-means clustering

RESULTS

categorize the neighborhoods into 3 clusters based on the frequency of occurrence for “Shopping Mall”:

- Cluster 0:
Neighborhoods with moderate number of shopping malls
- Cluster 1:
Neighborhoods with low number to no existence of shopping malls
- Cluster 2:
Neighborhoods with high concentration of shopping malls





DISCUSSION



- The highest number of shopping mall in cluster 2 and moderate number in cluster 0 . On the other hand, cluster 1 has very low number to no shopping mall in the neighborhoods.
- **Recommendations**
- The neighborhoods in cluster 1 are the most preferred locations to open a new shopping mall.



CONCLUSION :



- **To answer the business question :**
- The neighbourhoods in cluster 1 are the most preferred locations to open a new shopping mall. The findings of this project will help the relevant stakeholders to capitalize on the opportunities on high potential locations while avoiding overcrowded areas in their decisions to open a new shopping mall.



RECOURCES :



- **References**
- Category: Subdivisions of Rome in *Wikipedia*. Retrieved from https://en.wikipedia.org/wiki/Category:Subdivisions_of_Rome
- Foursquare Developers Documentation. *Foursquare*. Retrieved from <https://developer.foursquare.com/docs>