

# Coursera : Capstone project IBM datascience professional certificate

**Opening a big chain grocery store in new neighbourhoods**

# Business problem

- Location of supermarkets is one of the most important decisions that will determine whether the project will be a success or a failure
- Objective: To analyse and select the best locations in the city of Casablanca, Morocco to open a new supermarket
- This project is timely as the city is currently suffering from oversupply of supermarkets
- Business question
  - In the city of Casablanca, Morocco, if a property developer is looking to open a new supermarket, where would you recommend that they open it?

# Data

## ❖ Data required

- List of neighbourhoods in Casablanca
- Latitude and longitude coordinates of the neighbourhoods
- Venue data, particularly data related to shopping malls

## ❖ Sources of data

- Wikipedia page for neighbourhoods (https://en.wikipedia.org/wiki/Category:Neighbourhoods\_of\_Casablanca)
- Geocoder package for latitude and longitude coordinates
- Foursquare API for venue data

# Methodology

- Web scraping Wikipedia page for neighbourhoods list
- Get latitude and longitude coordinates using Geocoder
- Use Foursquare API to get venue data
- Group data by neighbourhood and taking the mean of the frequency of occurrence of each venue category
- Filter venue category by Supermarket
- Perform clustering on the data by using k-means clustering
- Visualize the clusters in a map using Folium

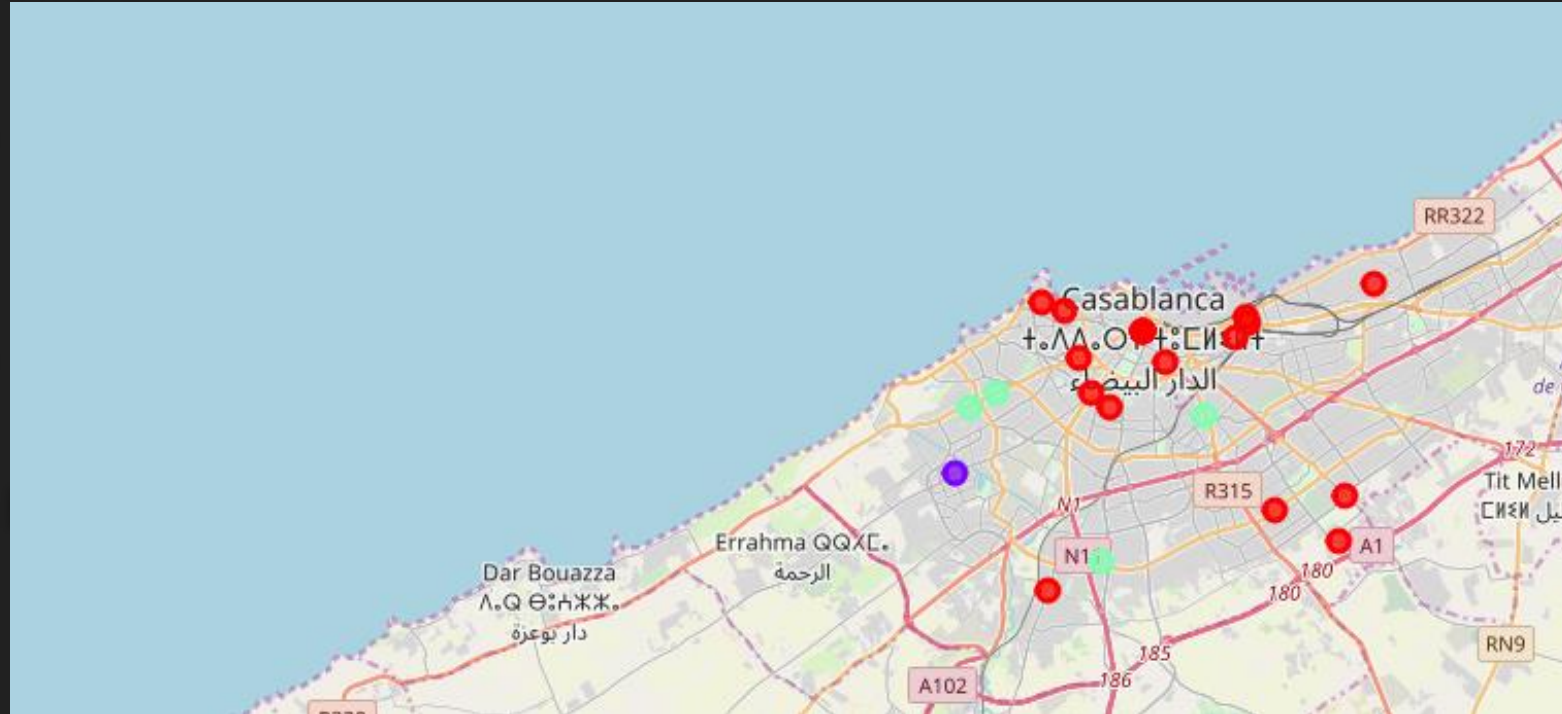
# Results

- Categorized the neighbourhoods into 3 clusters :

- Cluster 0: Neighbourhoods with low to no existing number of supermarkets

- Cluster 1: Neighbourhoods with moderate number of supermarkets

- Cluster 2: Neighbourhoods with concentration of supermarkets



# Discussion

- Most of the supermarkets are concentrated in the outer area of the city
- Highest number in cluster 2 and moderate number in cluster 1
- Cluster 0 has very low number to no supermarkets in the neighbourhoods
- Oversupply of supermarkets mostly happened in the outer area of the city, with the central area still have very few supermarkets

# Recommendations

- ❖ Open new supermarkets in neighbourhoods in cluster 0 with little to no competition
- ❖ Can also open in neighbourhoods in cluster 1 with moderate competition if have unique selling propositions to stand out from the competition
- ❖ Avoid neighbourhoods in cluster 2, already high concentration of supermarkets and intense competition



# Conclusion

- ❖ Answer to business question: The neighbourhoods in cluster 0 are the most preferred locations to open a new supermarket
- ❖ 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 supermarket



Thank you!