



# Coursera Capstone Project

## IBM Applied Data Science Course

---

### ***Opening a new Shopping Mall in Sydney***

By: Akash M

# Business Problem

---

- Location of the shopping mall is one of the most important decisions that will determine whether the mall will be a success or a failure
- Objective: To analyse and select the best locations in the city of Sydney to open a new shopping mall
- Business question
  - **If a property developer is looking to open a new shopping mall in the busy city of Sydney, where would you recommend that they open it?**

# Data

---

- Data required
  - List of neighbourhoods in Sydney
  - 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/List\\_of\\_Sydney\\_suburbs](https://en.wikipedia.org/wiki/List_of_Sydney_suburbs)
  - Geocoder package for latitude and longitude coordinates or csv file
  - Foursquare API for venue data

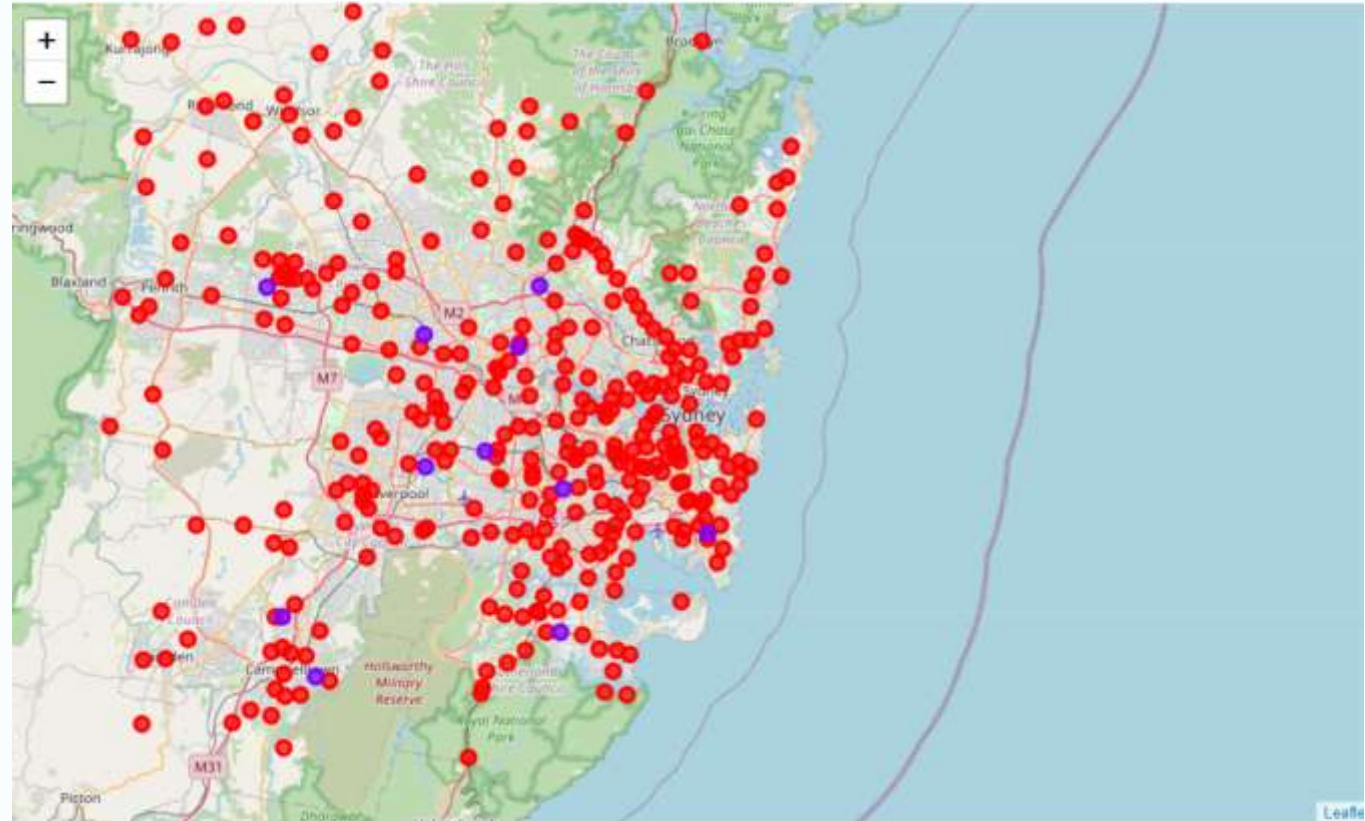
# Methodology

---

- Web scraping Wikipedia page for neighbourhoods list
- Get latitude and longitude coordinates using Geocoder or csv file if available
- 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 Shopping Mall
- Perform clustering on the data by using k-means clustering
- Visualize the clusters in a map using Folium

# Results

- Categorized the neighbourhoods into 2 clusters :
  - Cluster 0: Neighbourhoods with no shopping malls
  - Cluster 1: Neighbourhoods with shopping malls



# Observation

---

- All shopping malls in cluster 1
- Cluster 0 has no shopping mall in the neighbourhoods

# Recommendations

---

- Open new shopping malls in neighbourhoods in cluster 0 with no competition
- Even though the shopping malls are all concentrated in cluster 1 the total number of them in the cluster still provides an opportunity for opening malls.

# Conclusion

---

- Answer to business question: The neighbourhoods in cluster 0 are the most preferred locations to open a new shopping mall
- Findings of this project will help the relevant stakeholders to capitalize on the opportunities on high potential locations in their decisions to open a new shopping mall





---

THANK YOU