

The background of the slide is a dense field of three-dimensional, light blue numbers (0-9) of various sizes and orientations, creating a sense of depth and data. A solid black rectangular box is positioned on the right side of the slide, containing the title and author information in white text.

# IBM Applied Data Science Capstone

Essam M. Albishi

# Help you to determine your new business location in Riyadh, Saudi Arabia

## ◆ Introduction

- ◆ Riyadh city is the capital of Saudi Arabia

## ◆ Problem description

- ◆ Businessmen who would like to open a new business place
- ◆ Similar business exist around
- ◆ Avoid duplicate



# Help you to determine your new business location in Riyadh, Saudi Arabia

## ◆ **Data**

- ◆ List of neighborhoods of Riyadh
- ◆ Latitude and longitude coordinates
- ◆ Venue data

# Help you to determine your new business location in Riyadh, Saudi Arabia

## ◆ Methodology

- ◆ Provide list of Riyadh neighborhood from (<https://en.wikipedia.org/wiki/Riyadh>) using beautiful soup package in Python.
- ◆ Provide list of geographical coordinates using Geocoder package in Python.
- ◆ Use Foursquare API to parse all venue in Riyadh by providing their latitude and longitude.
- ◆ Using Python Panda to generate a dataframe merging all the above.
- ◆ Applying Machine learning method (K-mean clustering). We will cluster neighborhoods into 6 clusters.

# Help you to determine your new business location in Riyadh, Saudi Arabia

## ◆ Result

- ◆ Cluster 0: high concentration of restaurants.
- ◆ Cluster 1: low concentration of restaurants.
- ◆ Cluster 2: High concentration of restaurants.
- ◆ Cluster 3: low concentration of restaurants.
- ◆ Cluster 4: low concentration of restaurants.



# Help you to determine your new business location in Riyadh, Saudi Arabia

## ◆ Discussion

- ◆ Crowded neighborhood in our consideration

## ◆ Conclusion

- ◆ Highest number in cluster 0 then cluster 2
- ◆ Cluster 1,3,4 has very low number