

MY COURSERA CAPSTONE PROJECT
IBM APPLIED DATA SCIENCE PROFESSIONAL

LAUNCHING A MULTIPURPOSE ENTERTAINMENT CENTER IN KUALA LUMPUR, MALAYSIA

REPORT BY: MUHAMMAD MOHSIN JAVAID

AUGUST 2019

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 ANALYZE AND SELECT THE BEST LOCATIONS IN THE CITY OF KUALA LUMPUR, MALAYSIA TO OPEN A NEW SHOPPING MALL
- THIS PROJECT IS TIMELY AS THE CITY IS CURRENTLY SUFFERING FROM OVERSUPPLY OF SHOPPING MALLS
- BUSINESS QUESTION
 - IN THE CITY OF KUALA LUMPUR, MALAYSIA, IF A PROPERTY DEVELOPER IS LOOKING TO OPEN A NEW SHOPPING MALL, WHERE WOULD YOU RECOMMEND THAT THEY OPEN IT?

DATA

- ***DATA REQUIRED***

- LIST OF NEIGHBORHOODS IN KUALA LUMPUR
- LATITUDE AND LONGITUDE COORDINATES OF THE NEIGHBORHOODS
- VENUE DATA, PARTICULARLY DATA RELATED TO SHOPPING MALLS

- ***SOURCES OF DATA***

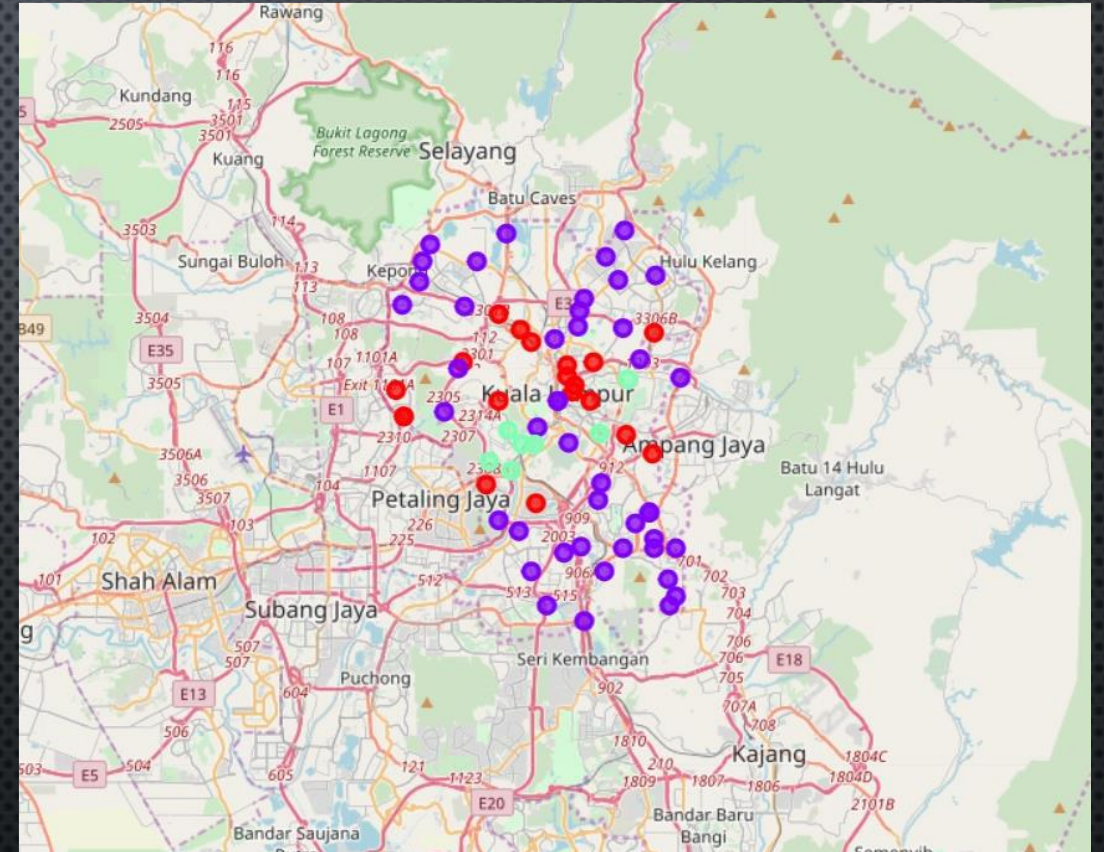
- WIKIPEDIA PAGE FOR NEIGHBORHOODS
([HTTPS://EN.WIKIPEDIA.ORG/WIKI/CATEGORY:SUBURBS_IN_KUALA_LUMPUR](https://en.wikipedia.org/wiki/Category:Suburbs_in_Kuala_Lumpur))
- GEOCODER PACKAGE FOR LATITUDE AND LONGITUDE COORDINATES
- FOURSQUARE API FOR VENUE DATA

METHODOLOGY

- WEB SCRAPING WIKIPEDIA PAGE FOR NEIGHBORHOODS LIST
- GET LATITUDE AND LONGITUDE COORDINATES USING GEOCODER
- USE FOURSQUARE API TO GET VENUE DATA
- GROUP DATA 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
- VISUALIZE THE CLUSTERS IN A MAP USING FOLIUM

RESULTS

- CATEGORIZED THE NEIGHBORHOODS INTO 3 CLUSTERS :
- 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

- MOST OF THE SHOPPING MALLS ARE CONCENTRATED IN THE CENTRAL AREA OF THE CITY
- HIGHEST NUMBER IN CLUSTER 2 AND MODERATE NUMBER IN CLUSTER 0
- CLUSTER 1 HAS VERY LOW NUMBER TO NO SHOPPING MALL IN THE NEIGHBORHOODS
- OVERSUPPLY OF SHOPPING MALLS MOSTLY HAPPENED IN THE CENTRAL AREA OF THE CITY, WITH THE SUBURB AREA STILL HAVE VERY FEW SHOPPING MALLS

RECOMMENDATIONS

- OPEN NEW SHOPPING MALLS IN NEIGHBORHOODS IN CLUSTER 1 WITH LITTLE TO NO COMPETITION
- CAN ALSO OPEN IN NEIGHBORHOODS IN CLUSTER 0 WITH MODERATE COMPETITION IF HAVE UNIQUE SELLING PROPOSITIONS TO STAND OUT FROM THE COMPETITION
- AVOID NEIGHBORHOODS IN CLUSTER 2, ALREADY HIGH CONCENTRATION OF SHOPPING MALLS AND INTENSE COMPETITION

CONCLUSION

- ANSWER TO BUSINESS QUESTION: THE NEIGHBORHOODS IN CLUSTER 1 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 WHILE AVOIDING OVERCROWDED AREAS IN THEIR DECISIONS TO OPEN A NEW MULTIPURPOSE ENTERTAINMENT CENTER.