Applied Data Science Capstone

by IBM

Report on
The Battle of Neighborhood

Stated Problem: The problem is to find an optimum place for placing a restaurant in the city Toronto, Canada.

Index:

Introduction:

Toronto, the capital of the province of Ontario, is a major Canadian city along Lake Ontario's northwestern shore. It's a dynamic metropolis with a core of soaring skyscrapers, all dwarfed by the iconic, free-standing CN Tower. Toronto also has many green spaces, from the orderly oval of Queen's Park to 400-acre High Park and its trails, sports facilities and zoo.

It is a great place for the entrepreneurs and the businessmen to set up new business. Food business is the most cost friendly and profitable business right now. It could be set up in a smaller place with a low investment. So for the food business, setting up a restaurant in the right place is the most important issue.

Problem Statement:

As stated the business problem is to find suitable places for an entrepreneur or the chain food shops to place a new restaurant in the city Toronto, Canada.

Method of Solving the Problem and Priorities:

For solving the problem two major issues will be the main factor. How many restaurants are already in the community? How many buyers the community holds? For the first one, the question is solved by finding all types of restaurants in different locations of Toronto. Then the second question is solved by finding the places which have more commercial places that are more densely populated with customers.

Data Collection:

The data is collected from the wikipedia source (https://en.wikipedia.org/w/index.php?title=List_of_postal_codes_of_Canada:_M) and

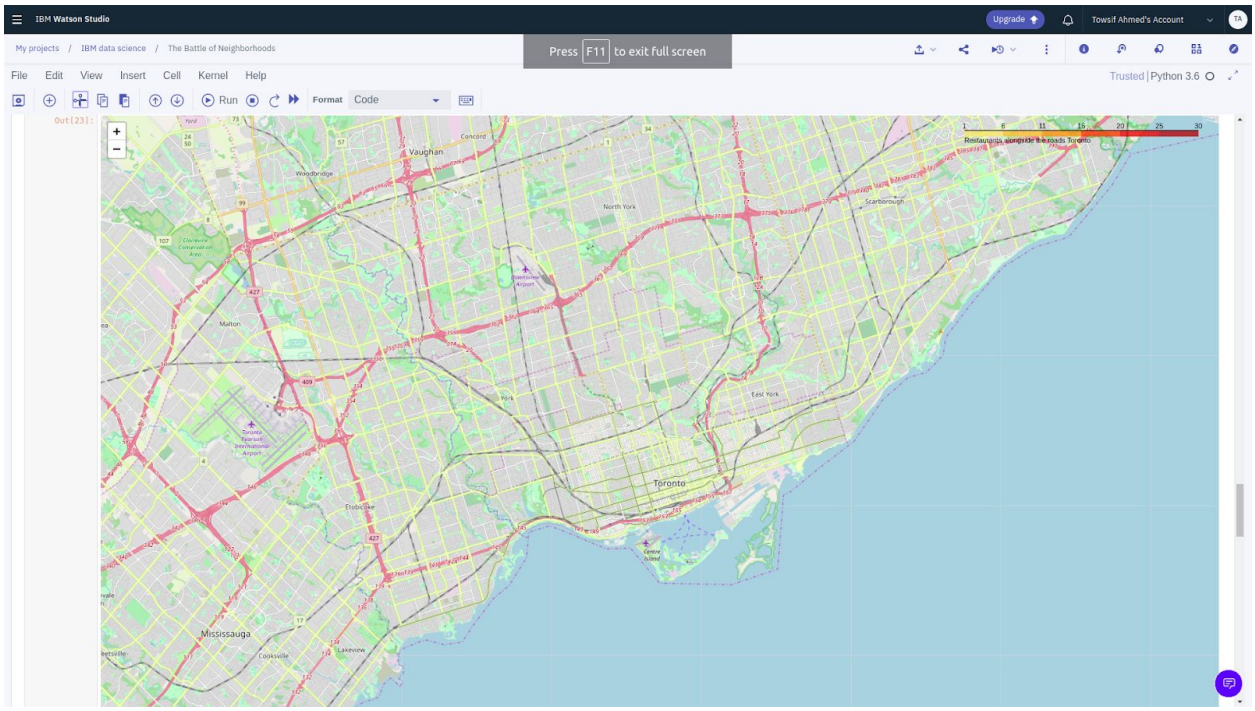
then the geo spatial data was got by the csv file given by the coursera(https://cocl.us/Geospatial_data) , It has given the locations of the neighborhood which were used for mapping. Foursquare was used for the data of finding restaurants and commercial places in the neighborhoods of Toronto.

Problem Solving:

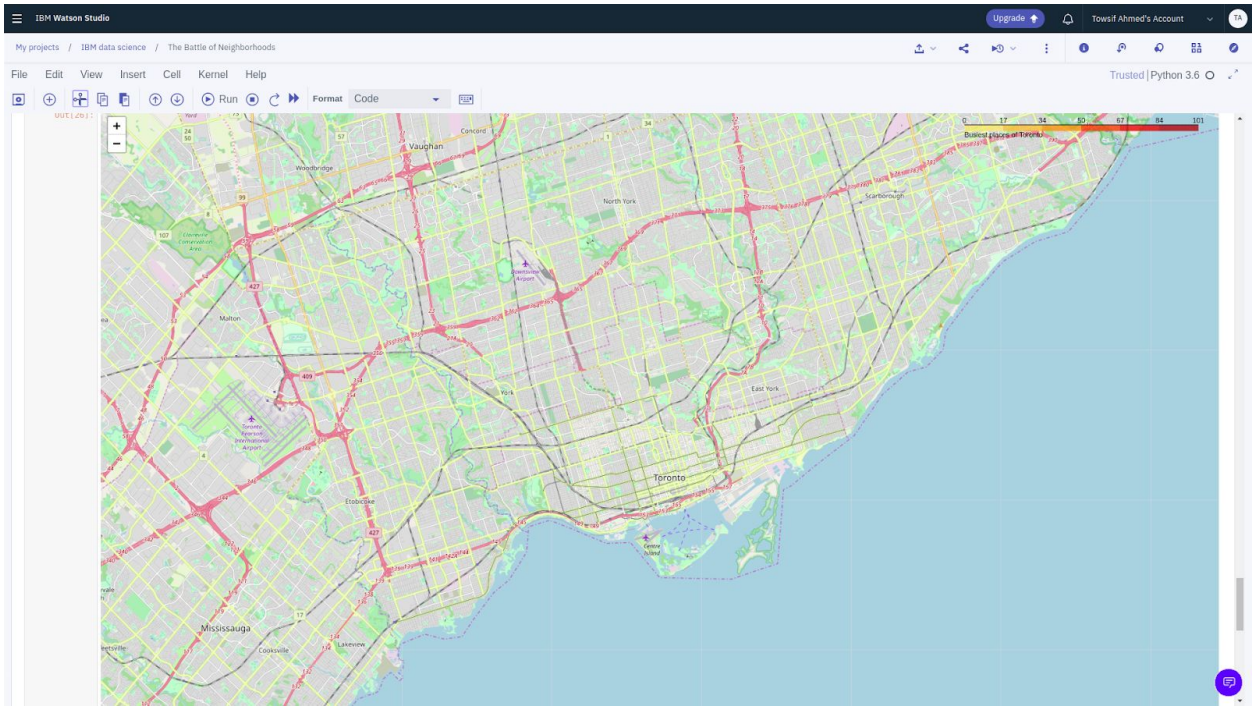
From the source of wikipedia, the data was taken using the methods of BeautifulSoup. The data was then stored in the DataFrame. Then the DataFrame was cleaned by removing missing values and duplicate rows. Then it was optimised by adding up the data into similar chunks.

Then a map was created to see the neighborhoods across the city and then Foursquare was loaded to get the details of the shops in the city based on different areas.

Firstly, the data of restaurants based on the locations was derived using Foursquare from the Dataframe. It has been used for counting down the restaurants. Then a folium choropleth map has been set to see the busiest street roads with the number of restaurants on them.



Then the busiest places with them most number of commercial places were determined from the data and also using the choropleth the figure has been shown.



So then from the comparisons the decisions were taken and the solve was coming through.

Discussion:

Harbourfront East, Toronto Islands, Union Station are the busiest place with the least amount of restaurants. There are 100 shops in those locations but only 15 of the shops are related to the food business. So a restaurant on the streets of those places will surely bring success.

Conclusion:

The analysis of the problem along with the solve and operations on the data sets were completed sincerely ensuring the right places to set up a restaurant in Toronto. Thanks for your time and hope you have a great day.