Coursera Data science Final Project

Author: Oliver Ignetik

Contents

[Introduction 1](#_Toc9862017)

[Data 1](#_Toc9862018)

[Methodology 1](#_Toc9862019)

[Results 1](#_Toc9862020)

[Discussion 1](#_Toc9862021)

[Conclusion 1](#_Toc9862022)

# Introduction

Business problem: Opening a new Italian restaurant in Toronto, Ontario.

This project will investigate the best location to open a new restaurant in the city of Toronto, Ontario. This project will have value to investors who would like to open a new restaurant in Toronto and need to know which locations would lead to a better profit margin.

# Data

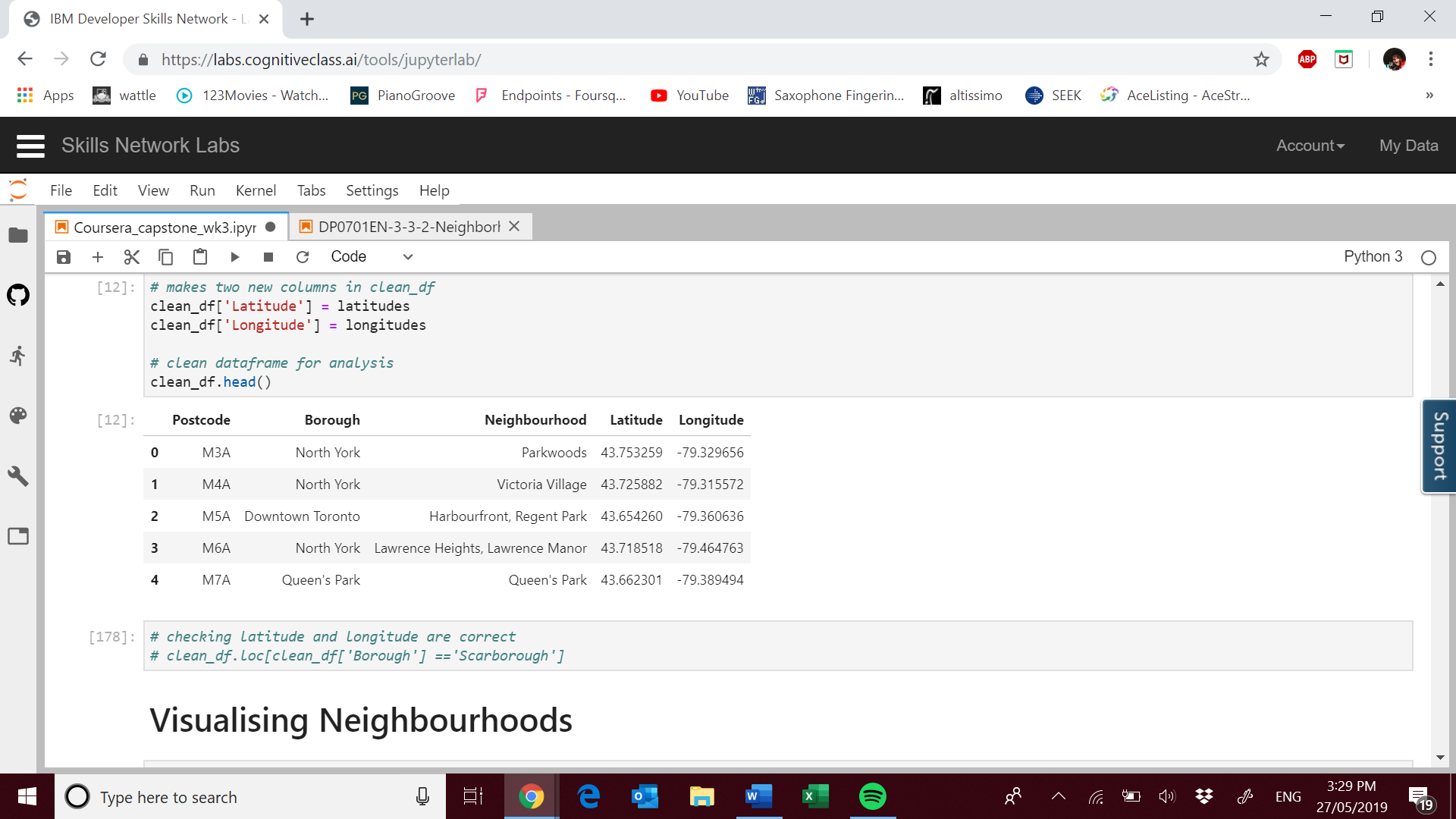
The borough and postcode of each neighbourhood was retrieved from: <https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M>

|  |  |
| --- | --- |
| Features | Description |
| Postal code | Unique postcode in Ontario, Toronto |
| Neighbourhood | District within Toronto |
| Borough | Administrative district within Toronto |

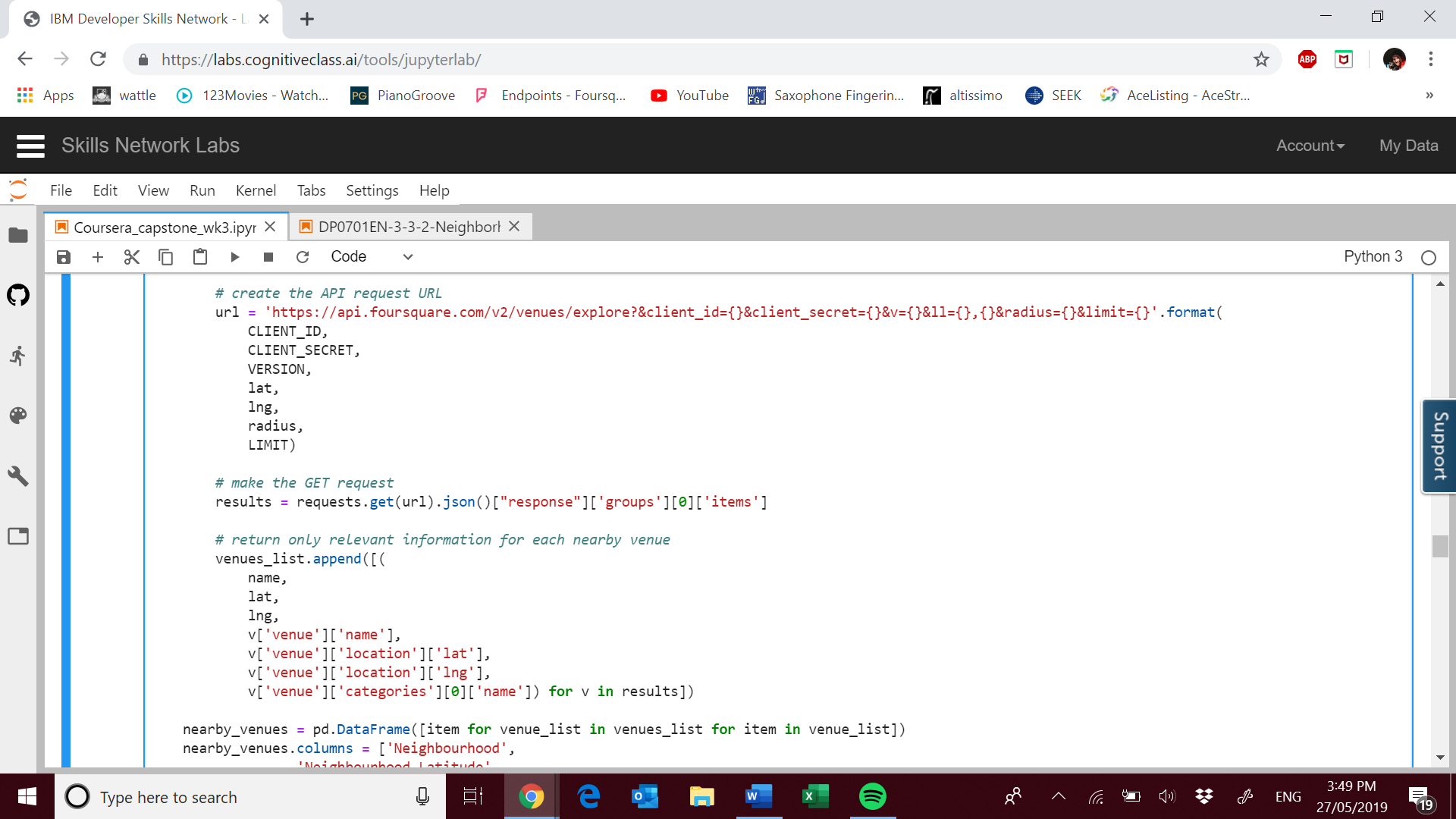
The latitude and longitude of each postcode was retrieved from: <https://cocl.us/Geospatial_data>. The fields of this csv file are discussed in the table below.

|  |  |
| --- | --- |
| Features | Description |
| Postal code | Unique postcode in Ontario, Toronto |
| Latitude | Angular distance from equator |
| Longitude | Angular distance from Greenwich median |

The data from both these sources was filtered and stored in the data frame named clean\_df.



The Foursquare API was used to retrieve information about the venues nearby each postcode. Foursquare requests are in the form of a JSON file. The information of value was extracted from the JSON file by accessing the relevant fields as shown below:



Once the data had been retrieved from the JSON field the rest of the data analysis was completed entirely within the Jupyter Notebook environment.

# Methodology

# Results

# Discussion

# Conclusion