**Capstone Project:**

**The clash of Neighborhoods in Toronto City**

# Description of Data

The data are retrieving from three different sources. The first is a Wikipedia page to retrieve postal code and Borough of neighborhood located within the city of Toronto in the province of Ontario, Canada. (<https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M>). The second source is a csv file that contain all specific location for neighborhoods in Toronto city (<http://cocl.us/Geospatial_data>). With these two databases, a dataframe containing Postal codes, Borough, Neighborhood, and specific location (i.e latitude and longitude) is created (Figure 1). The dataframe has 10 boroughs and 103 neighborhoods (Figure 2).

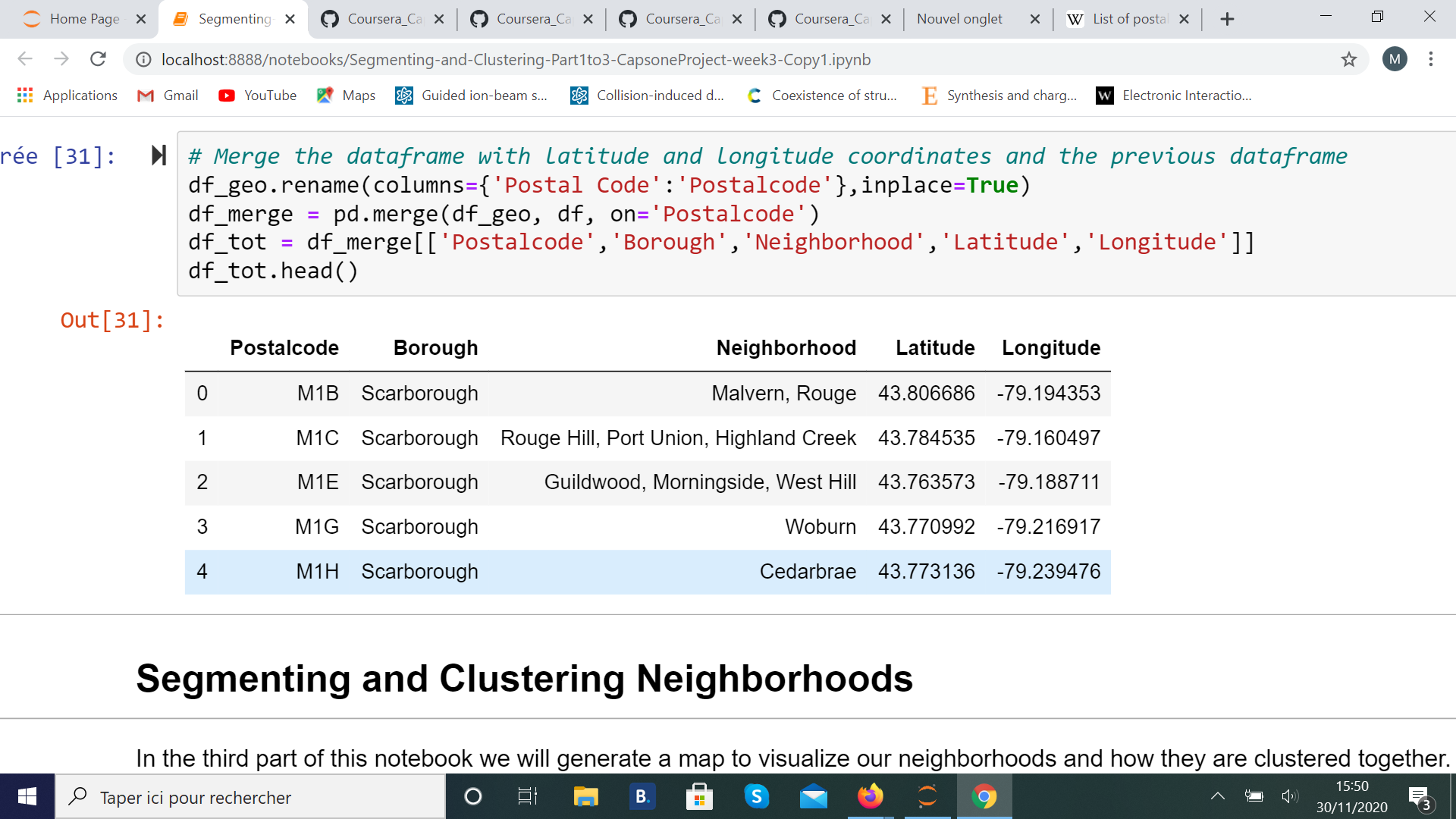


Figure 1 : Dataframe that contains keys information for neighborhood in Toronto city, Canada.

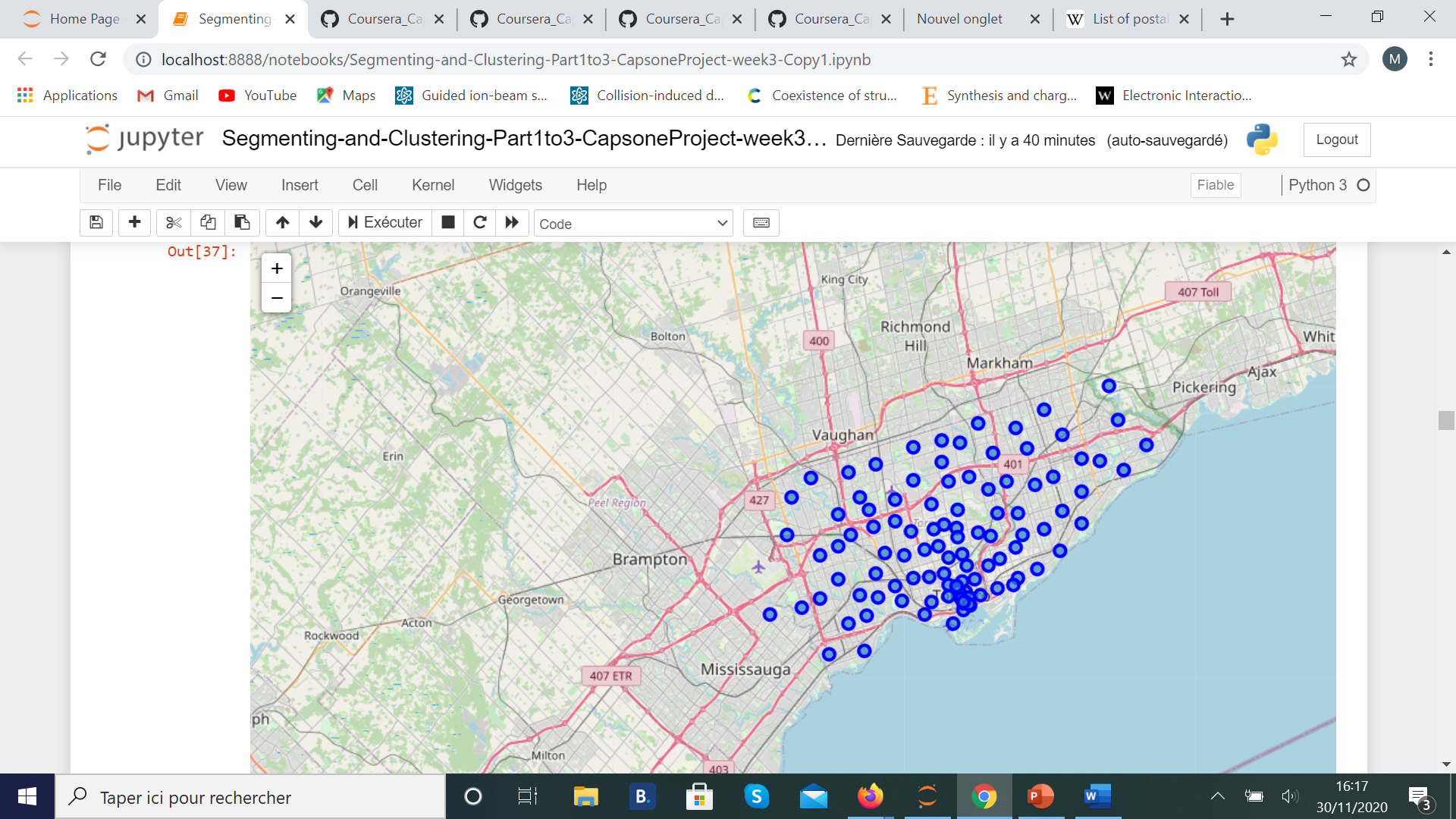


Figure 2 : Map of all neighborhoods retrieve from database for Toronto city, Canada.

The third source is the Foursquare website to retrieve database for the venue in neighborhoods in Toronto. The features available in the dataset are for example: venues, store reviews, longitude and latitude, names of the neighborhood, categories of venues, types of stores. ­­From this data we can create several dataframe, one example is given is the figure above (Figure 3).



Figure 3 : Example of dataframe that can be created with data from the Foursquare website.

When we analyze data in this database, we see that there is 2161 venues listed in Toronto city with 276 uniques categories of venue. If we look at the 20 first neighborhoods in terms of the number of total venues (Figure 4), we see that there the 6th first neighborhoods count 100 venues (all categories included) and the 20th neighborhood, Runnymede, Swansea, count only 30 different venues. This is an important feature because independently of the list of requirements and therefore the category of venue desired, the Robert’s family as everyone seek for a large number of services available near their future home and a to ensure a dynamic neighborhood lifestyle.

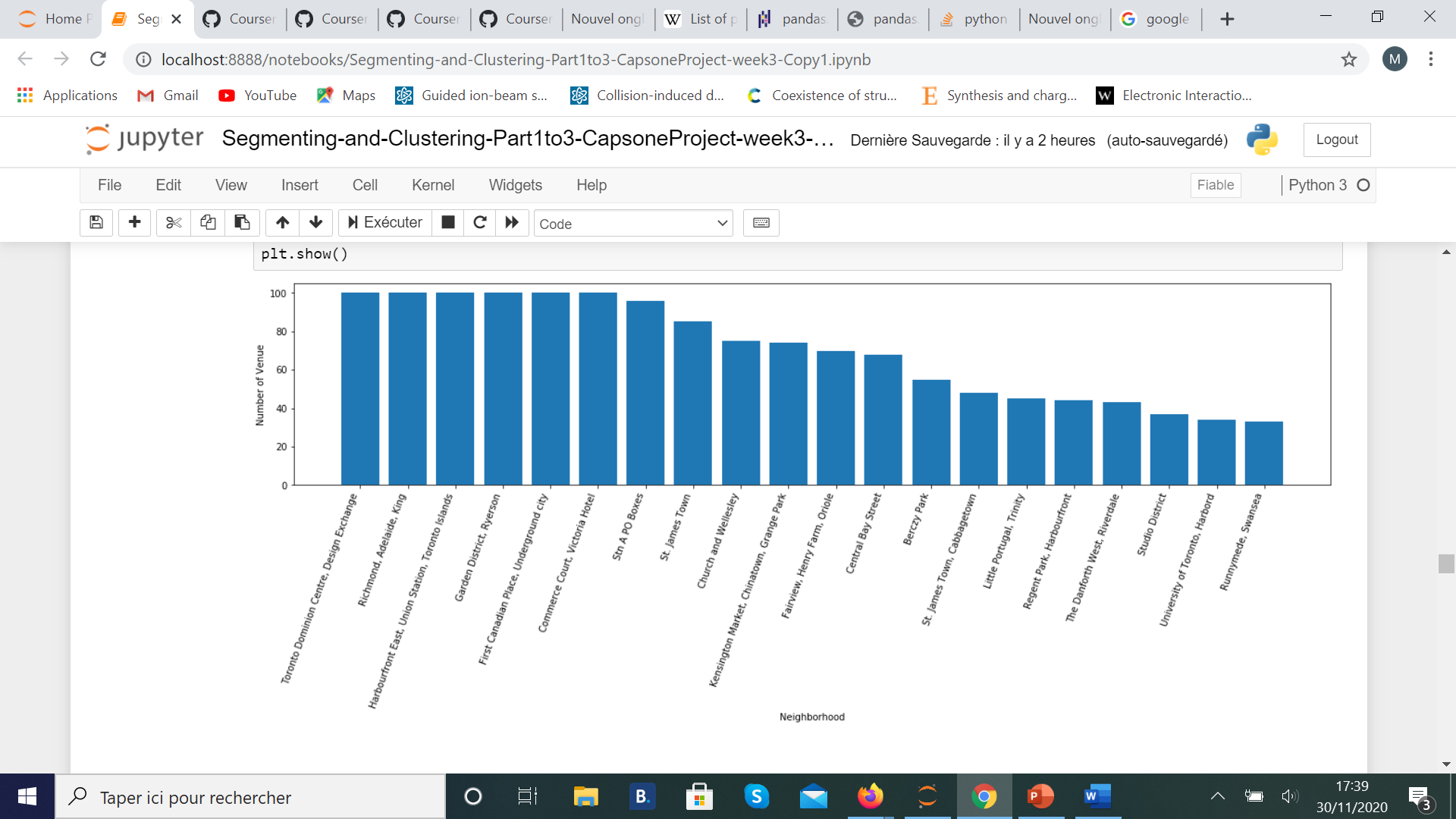


Figure 4 : Count of venues within the 20 first neighborhoods in terms of the higher number of venues.