Considering our visitor’s requirements, we decide to collect data through Wikipedia and make use of foursquare to get further information.

Our raw data from Wikipedia contains the info of Toronto’s postal codes, boroughs and neighborhoods. We then add accurate location of neighborhoods to it. After that we wrangle it and attain the info of venues surrounding those neighborhoods, including which instruments there are on each venue.

Finally we clean the data and drop those information which we aren’t interested in.

In general, there are totally 104 rows and 3 columns in our raw data. We then add two columns to it which describes the location, resembles latitude and longitude. After our processing, there are 16 rows and 11 columns eventually.