

## **Data**

### **Section 1 - Sources of data**

The following is list of data which will used to solve the problem - - Which financial services is highly concentrated in Chicago ?

- 1) List of Chicago neighborhoods had been sourced from Wikipedia site. The website contained a table of different neighborhoods along with community areas . There are 246 Chicago neighborhoods.
- 2) I have used the Geopy package to generate latitude and longitude values for Chicago neighborhoods.
- 3) I have used Foursquare API to find the financial institutions/banks around the neighborhoods of Chicago within 800 meter radius.

### **Section 2 - Data Cleaning and Analysis**

There was one neighborhoods which had missing values. Hence the row was dropped, making number of neighborhoods total as 245. From Geopy package, latitude, longitude, altitude, point and location were generated. Columns such as point, altitude and location had been dropped. Latitude and longitude columns were renamed by making their first letter uppercase.

With help of Foursquare API, I used search query to generate JSON file containing raw list of financial institutions and other institutions within radius of 500 meter. I assigned relevant part of JSON file to obtain results containing details of financial institutions like location address, latitude, longitude , category and etc. After normalizing JSON file, I obtained unfiltered data frame of financial institution. I only took necessary columns like name of the institution, latitude , longitude and category which institution belongs to and stored in different data frame.

This data frame contained 30 venues located in Chicago within radius of 800 meters.

### **Section 3- Analysis of Data**

With help of this data, I was able to generate a map showing the location of thirty financial institutions located in Chicago.

Based on the data frame containing 30 financial institutions, I got the distinct count of categories. There were 14 distinct categories in which I removed the categories that didn't belong to financial services criteria. Finally I had 8 distinct categories. From this information, I was able to generate Bar Graph to represent which was highly concentrated financial service in Chicago.