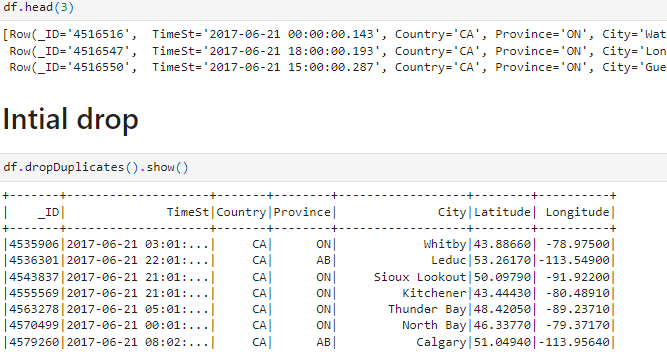
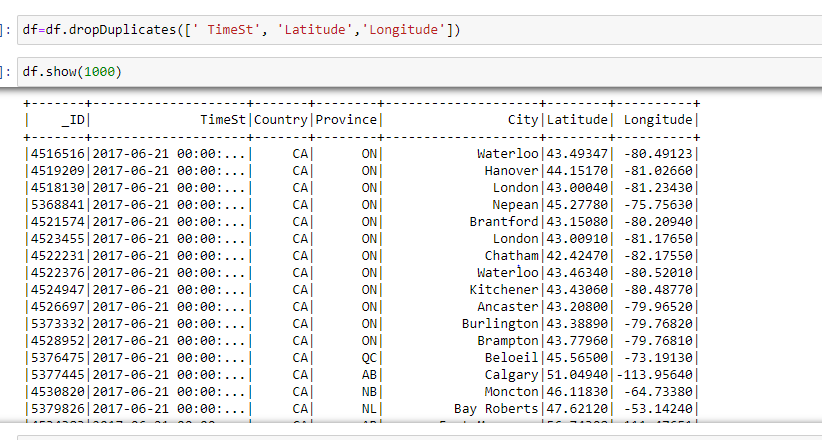
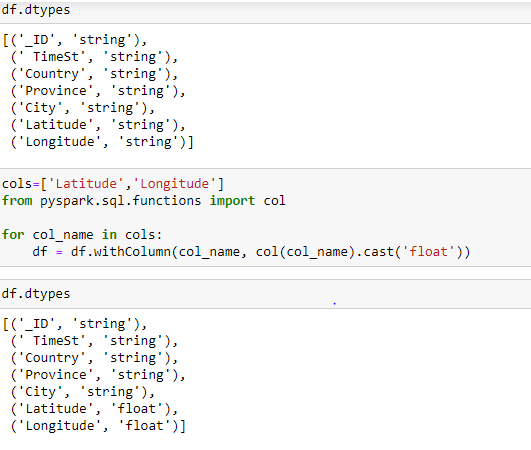
In this file results and discussions related to project is done.



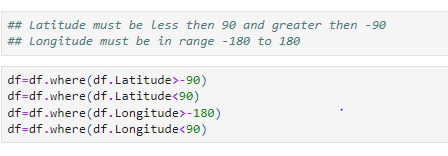
In the above image. Initial duplicate values are dropped. It can be seen that. No duplicate value overall is now available in the data.



In the step above the similar time, latitude and longitude are dropped. If any of the values are similar to each other, then they are dropped regardless of different Ids.



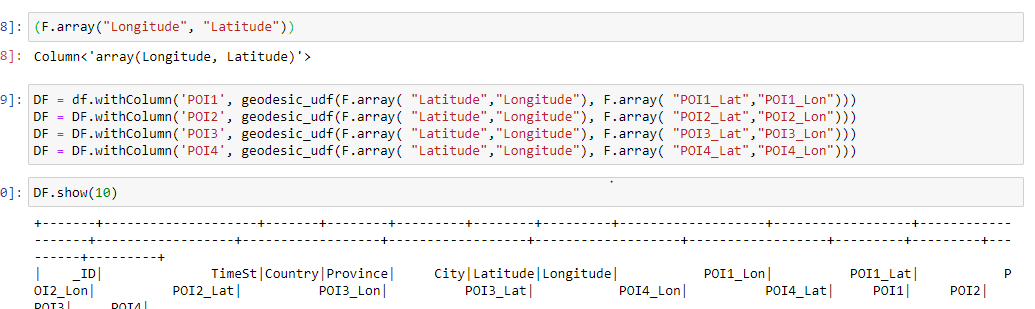
The above work is done for conversion of the data types of the specific locations. The string data is converted into the float data.



The next step is dropping the wrong entries. It is necessary to stay within the range. Otherwise the map would not be built.



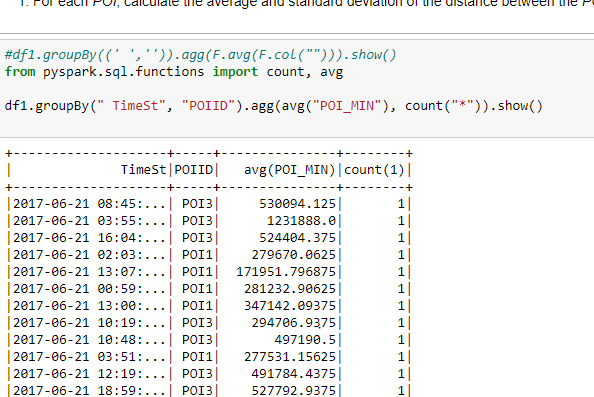
The next step is to read POI data. The POI1 and POI2 are same The POI3 is little bit different from POI4.



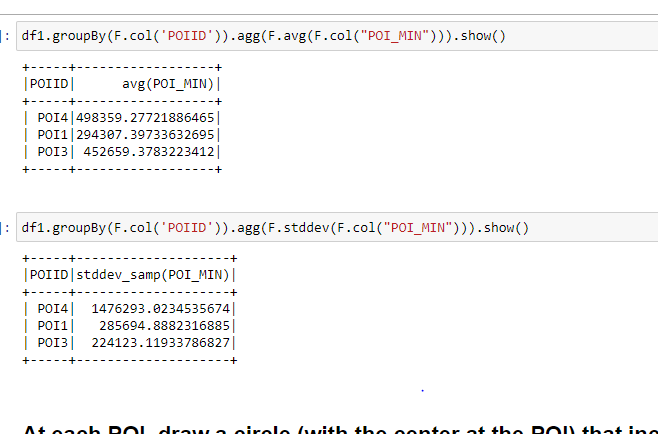
The geo distance is used to calculate the distance between POI longs and lats. And now our data frame has many columns of all POIs and their distances.



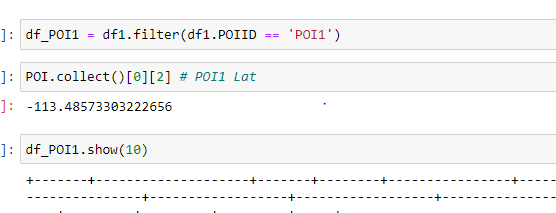
Moving on further we need to select minimum POI value and also need to select which POI is minimum. Through which we have POIID and POI\_MiN columns containing this information.



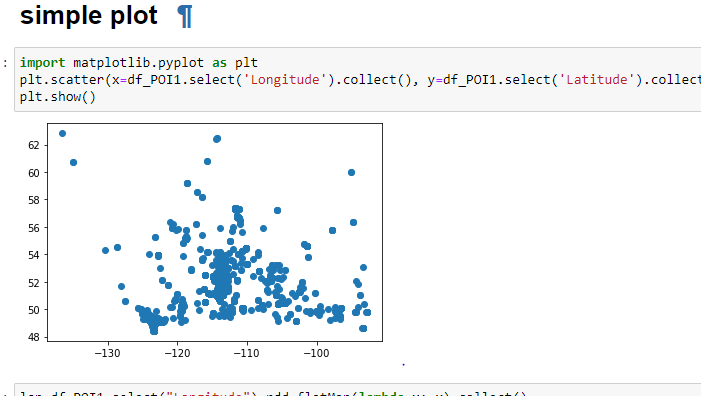
If we consider POIs and timelines, it can be seen that there is no similar time set available. So, we need to move further to see average.



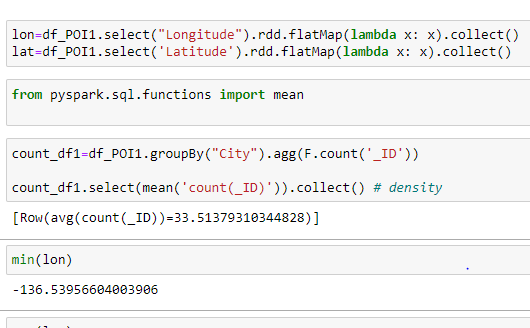
The image above calculates the average and standard deviation of the POIs data. It can be seen that the average of all 3 is different and far from each other.



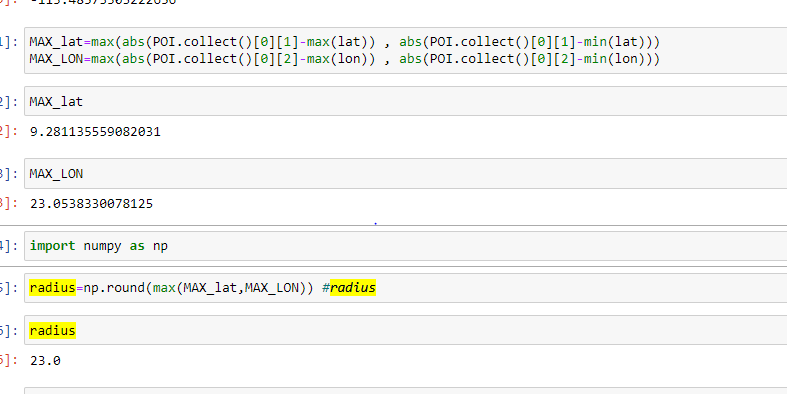
First, POI is selected to move further.



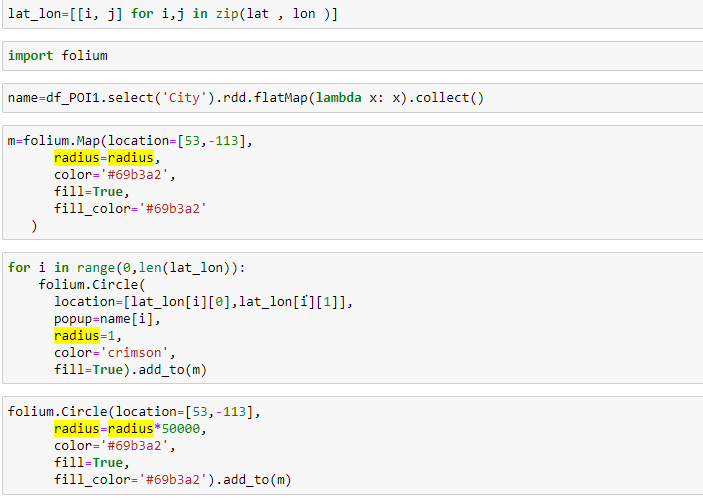
A sample plot is drawn to check the spread of the map.



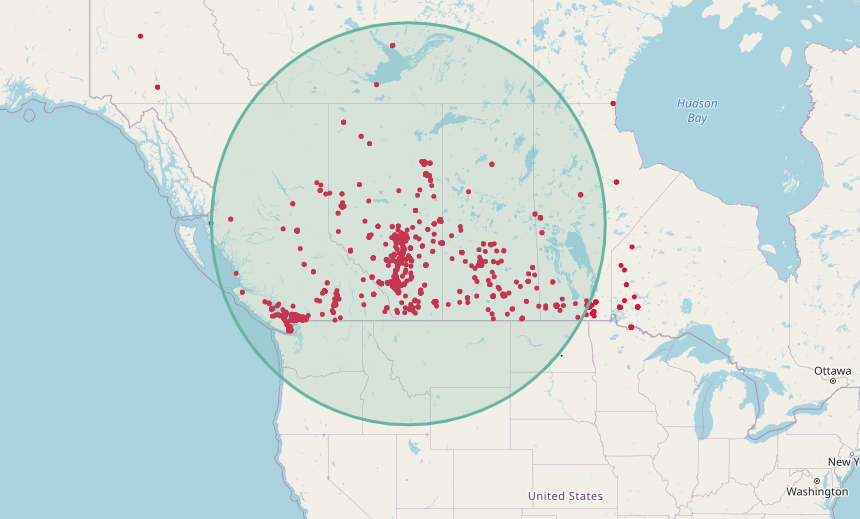
The longitude and latitude values are selected. Now, all the city values are counted and average value is calculated. This value 33 is density per unit area.



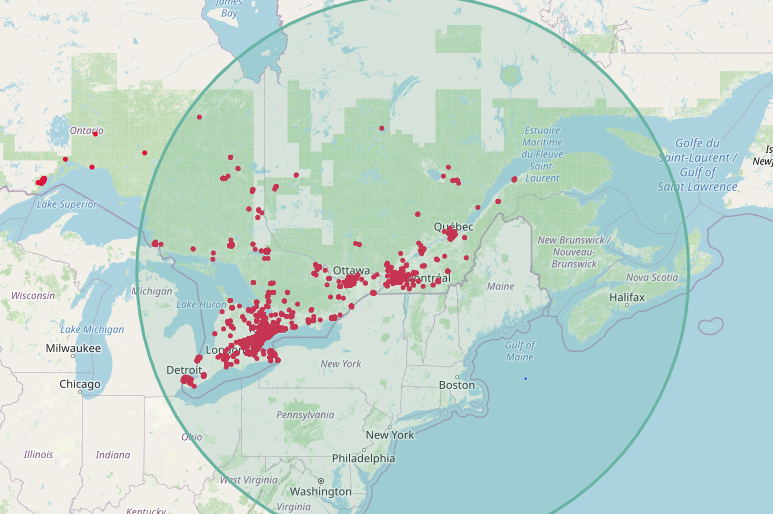
The longitude and latitude different values are calculated. And most far value is selected as radius. So that circle can spread through the whole values.



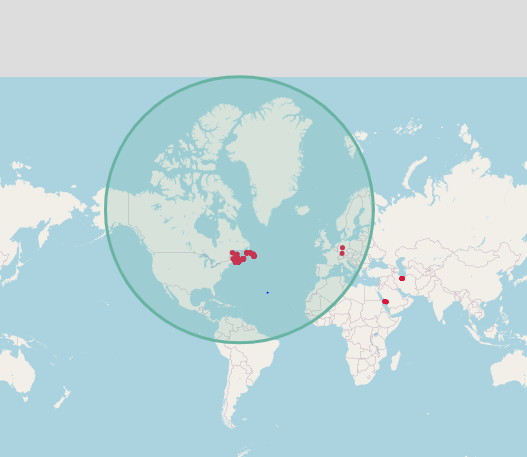
Now, the longitude and latitude values are selected and final folium map is drawn.



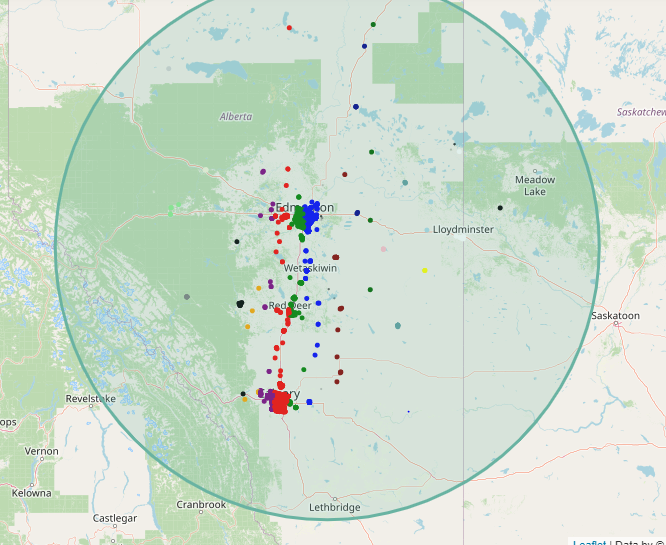
This is the first map and most of the data is inside the circle. The data outside the circle is basically outliers. So, it can be removed from the dataset in next step.



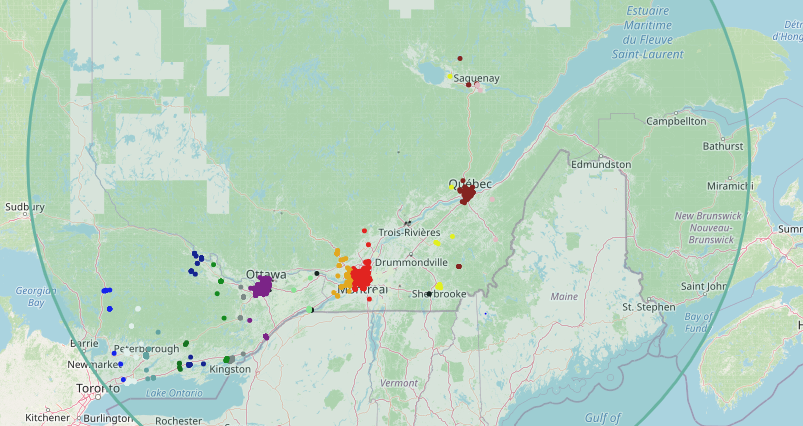
Moving on towards next step the POI3 also has some outliers and they can be removed from the dataset in modelling step.



Drawing the circle for POI4. There are also some outliers that can be removed from the dataset.

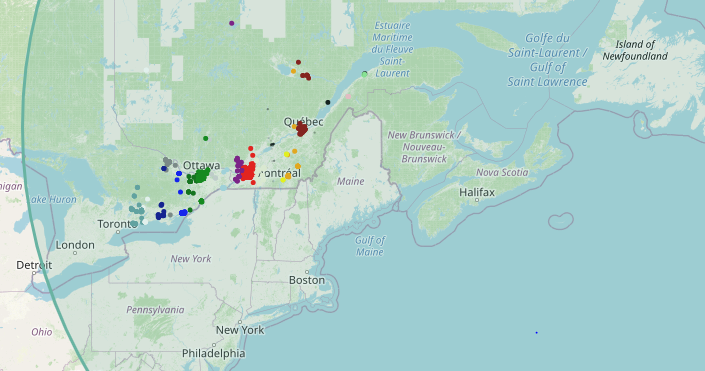


Using the IQR and mean sensitive threshold. Th data is divided into -10 to 10 groups. The different colors in the map shows the different groups. Finally, most of the data is near to the POI1 and circle is drawn correctly and no outlier is available in the data.



Using the IQR and mean sensitive threshold. Th data is divided into -10 to 10 groups. The different colors in the map shows the different groups. Finally, most of the data is near to the POI3 and circle is drawn correctly and no outlier is available in the data.

The map is beautiful and can be considered as perfect distribution of the data as well.



For POI4, the circle center has most of the values towards Ottawa and London and the spread is also done perfectly. The data is grouped perfectly from -10 to 10 and the distribution is successful among all the groups.