The code will run on Anaconda Navigator

The important libraries that may need to pip install if not available

Pip install pyspark

Pip install findspark

Pip install geopy

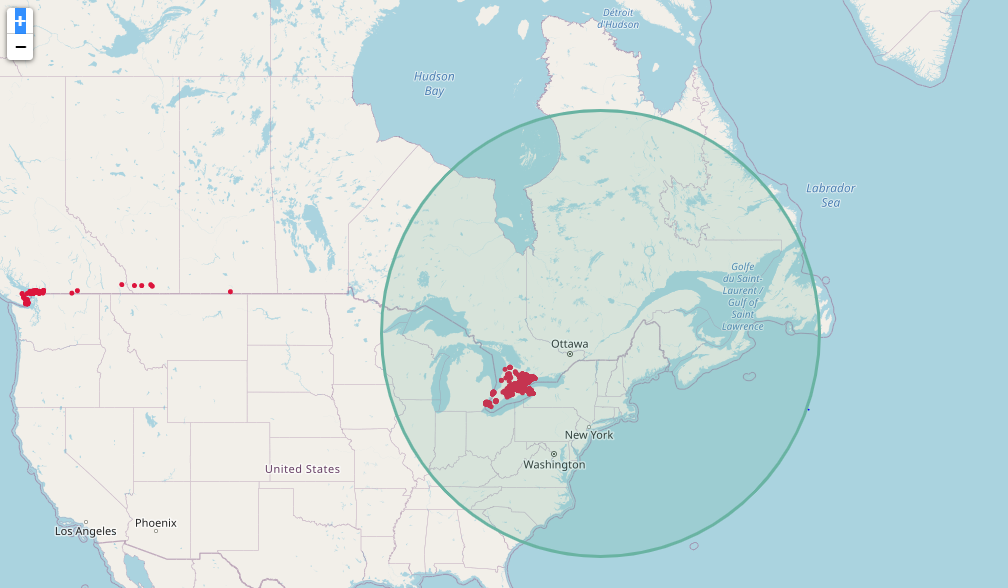
Pip install folium

Once code is ready to run. You can run code step by step. The cell above must run first and in time to run the next cell. Skipping cells will results in errors. Move from cell 1 to last and keep running.

Another method to check if code runs without error is to restart and run all. See which error come. If it is related to pyspark, geopy, folim then use pip statement as above.

The important part to consider in this project is google map which will be is used to answer question 3 and 4.

To get complete information from the map. One must be able to use + and – signs to zoom in and zoom out from the data.



For example zooming out several times will results in better show the circle. It will also show the outliers and other details.