Introduction

Are a foodie? and if are, have u ever wondered which localities have the highest variety of food choices available at a walking distance in your city? this project tries to solve just this.

Problem

which neighborhoods have the highest diversity of foods choices available at a walking distance in New York city?

Data Source

The data is acquired from about varies neighborhoods in New York city.

Data Processing and cleaning

The data consisted of 48895 entries. And contained 4 columns namely area, neighborhood, latitude and longitude.

This is how the data looks like after cleaning

<https://github.com/Akshay-s-raut/Coursera_Capstone/blob/master/new_york2.csv>

Choice of neighborhoods

Since the there are 48895 entries in the data.

I selected 200 of these at random and plotted them using the Folium.

200 entries were chosen as there were enough to visualize and process.

Parameter selection

I assumed the walking distance to be 300 meters.

And acquired 30 restaurants from that range.

Assigning variety

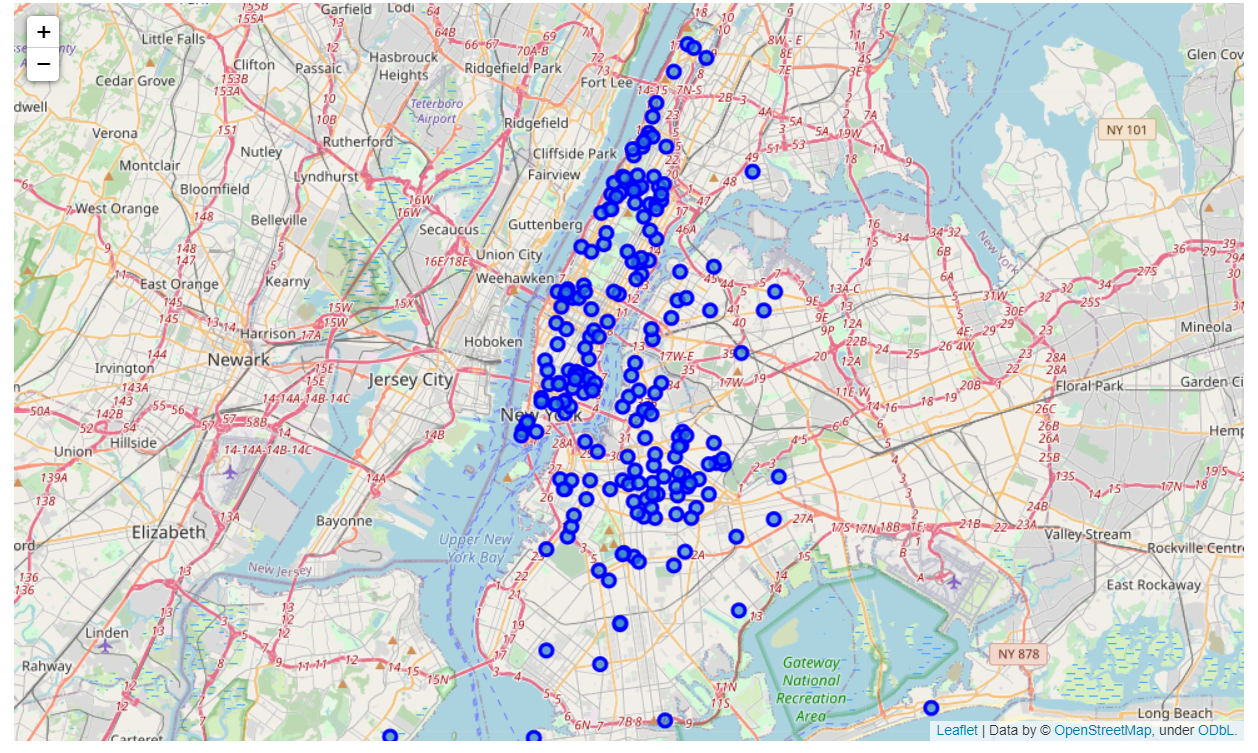
I have used the entropy in selected restaurants to get some idea about the variety of choices available in the neighborhood.

Visualizing tools

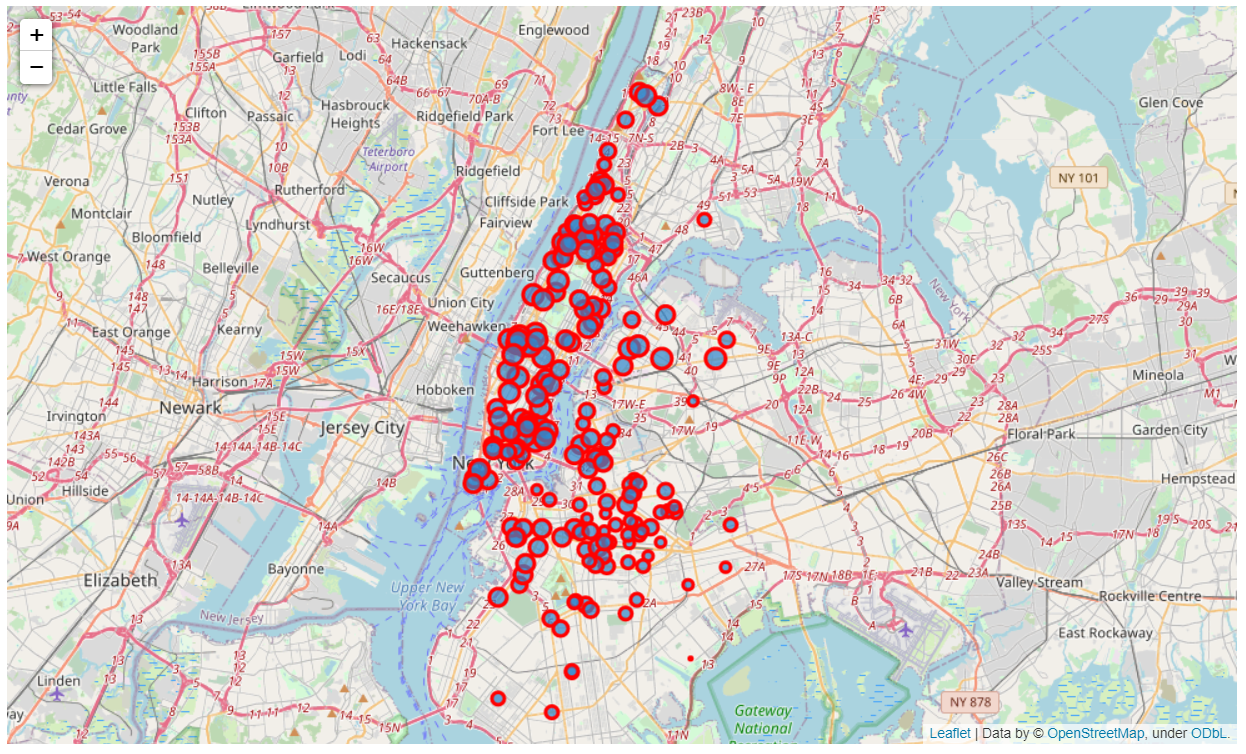
I have used namely Folium to plot and visualize the neighborhoods

Result

For a chosen set of 200 neighborhoods



If found the variety intensity to be



Here larger radius implies high variety.

Conclusion

The model uses simple measure entropy to calculate variety of choices available.

I Have used only 200 entries but the model can do much more if adjusted for processing time and power.

Discussion

In future a model can be trained to also retrieve the ratings of restaurants. And Compensate not only for variety but also for quality using weighted entropy.