Introduction

The project is done for the client “Cuppa-Bliss Inc” since they wanted to open a new branch of their coffee shops in New York city. We were asked to find out the top 5 neighborhoods where they could open a new shop.

Data

The data required is the location of the coffee shops in each neighborhood. To gather the information we used the neighborhoods information from the JSON file provided and we used four squares to find how many coffee shops are there.

Methodology

With the data as described above, we did the data preparation first. Along with the JSON file, we used four squares to find how many coffee shops are there. Then we filtered the neighborhoods with more no. of Coffee shops, and got a list of neighborhood with only one coffee shop available. Also, the client mentioned that they would not want to open the coffee shop where there is a StarBucks in the area. We also excluded the areas with StarBucks on the analysis.

Finally the k means algorithm was applied to get the clusters of the neighborhoods.

Results

We got the following result :

	Borough	Neighborhood	Latitude	Longitude	Cluster Labels
0	Bronx	Kingsbridge	40.881687	-73.902818	3
8	Brooklyn	Bensonhurst	40.611009	-73.995180	4
16	Manhattan	Inwood	40.867684	-73.921210	1
18	Queens	Woodside	40.746349	-73.901842	2
20	Staten Island	New Springville	40.594252	-74.164960	0

Discussion

K-Means algorithm was used because it was a clustering problem.

Conclusion

As per the analysis, we advised the client “Cuppa-Bliss Inc” to open the coffee shop at Kingsbridge neighborhood in Bronx.