

# Drugstore Establishment

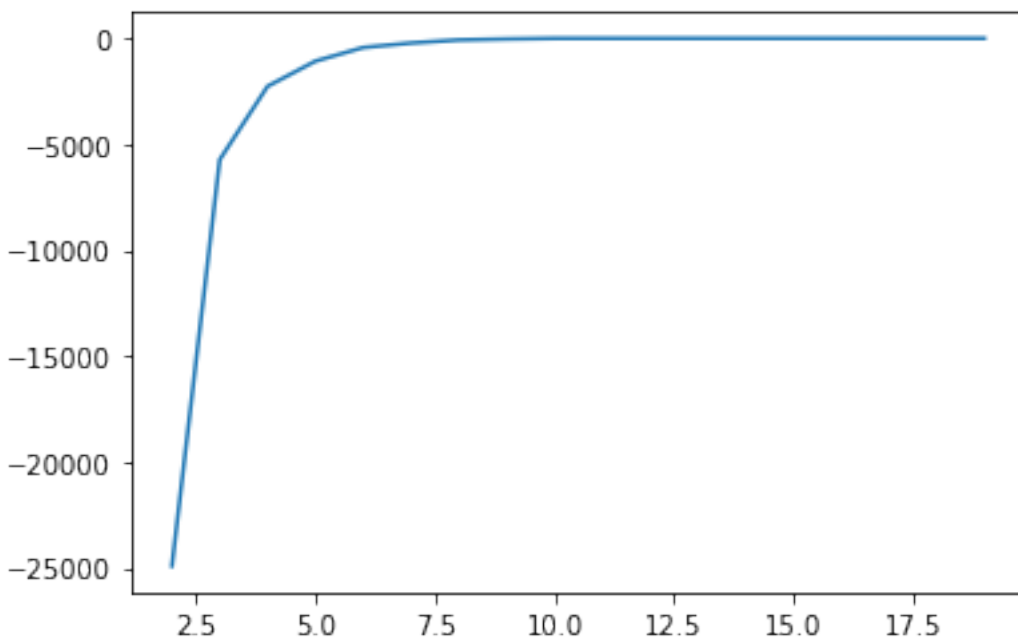
## Introduction

In this project I solved problem of drugstore owner who was interested to establish a new drugstore in US. He/She wants the drugstore to be in the region where most hospitals are located since more hospitals means more patients who need drug.

In order to find the best region I used foursquare api to get list of hospitals in 10 most popular city of US. After collecting the data, I cleaned it and transformed it to a dataframe that has 300 rows and include name of hospital, city, state, lat and lng of the location.

## Methodology

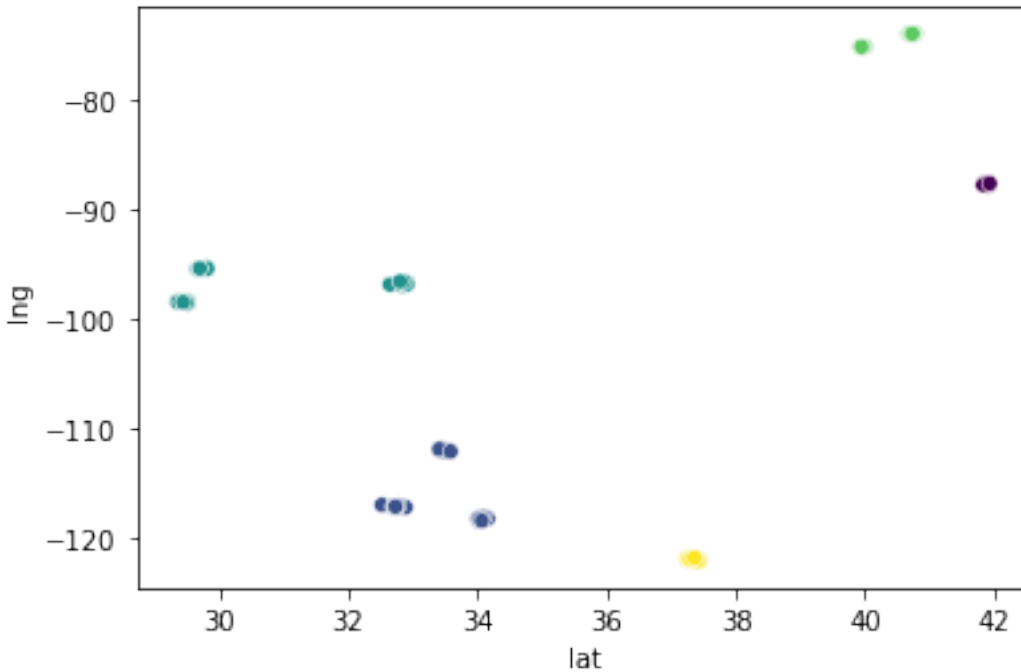
In order to find the region where most hospitals are located I used kmeans clustering to segment the data into 5 region. The reason for k=5 is that with 5 cluster model had the highest accuracy.



## Result

After setting k=5 and training the model I segmented all hospitals into 5 regions and found out that:

*Cluster 1 has 30 data hospitals*  
*Cluster 2 has 90 data hospitals*  
*Cluster 3 has 90 data hospitals*  
*Cluster 4 has 60 data hospitals*  
*Cluster 5 has 30 data hospitals*



## Discussion

Based on the results of previous section I would say the best place for a new drugstore would be within cluster 2 and cluster 3 since most hospitals are located in those area. This project has a lot more that can be improved on, for example maybe cluster 2 and cluster 3 already have many drugstore and a new one won't work.

## Conclusion

So if a drugstore owner is interested in opening up a new one in 10 most popular cities of US, the region in cluster 2 and 3 would be the best choice since it has the highest number of hospitals.