Opening a Shopping Mall in Los Angeles

Author: Shiyun Lu

Introduction and Business Problem

There are plenty of malls in Los Angeles. However, there will never be enough of them. If someone with a lot of money wants to open a new shopping mall in L.A., data science should come in to help.

Let's assume we are helping some major stakeholders who want to pick an address in L.A. to oopen up a new shopping mall. We are trying to find out locations which are the most profitable. We will use methodologies like data science and machine learning to solve this problem based on some models we can build (e.g. clustering models).

So what is an ideal location for Los Angeles to open a new shopping mall? This will be the problem we stick to in the following process of exploring.

Data

We need below data for this project to be done:

1. Neighbourhood information of the greater Los Angeles area.
2. 2.Geographical information (latitude and longitude) of the greater Los Angeles area.
3. 3.Venue data of the greater Los Angeles area.

The Los Angeles government offered a list of zip codes of L.A.:
(http://file.lacounty.gov/SDSInter/lac/1031552_MasterZipCodes.pdf) from where we are able to extract L.A. neighbourhood information. For Geographical information (latitude and longitude) we can get it from Python package and for venue data we are able to get it from Foursquare API according to the methods

Conclusion

I have clustered Los Angeles into clusters using and the result shows cluster 2 has higher density of shopping malls while cluster 1 has lower one. Thus I might advise stakeholders to open a new shopping mall in cluster 1 area to get potentially high customer inflow.