Introduction/Business problem

We want to the build a segmentation and clustering of the neighborhoods of the city of California.

The objective of the segmentation/Clustering is to provide guidance to investors or privates who want to buy a house in this city. Hence our aim is to identify neighborhood clusters in the city based on several features of a given neighborhood. Some include

* Proximity to the ocean
* Median income
* Median house value
* Median age of houses
* Location (longitude, latitude)
* Number of households

The primary group that might be interested in the clustering of houses are mortgage companies. The clustering will enable to quickly find a choice that fits a client.

A second group for which this job is of interest is made up of businesses such as Theater/Movies, Restaurants, Malls. Cluster will determine which type of ressources or products will be allocate to the neighborhood.

The clustering of neighborhood is also important for the city council. Depending on the factors defining the clusters, some neighborhoods might be rebuild or other plans developed.