




Capstone Project

Battle of the Neighborhoods - Boston



Which neighborhood to move to in Boston if you are new?

- When moving to a new big city such as Boston, there are numerous neighborhoods to choose from
 - Factors such as the population density, home values, and the diversity of services in a selected neighborhood could be important
 - It would be helpful if neighborhoods with similar characteristics are grouped together for consideration
- 

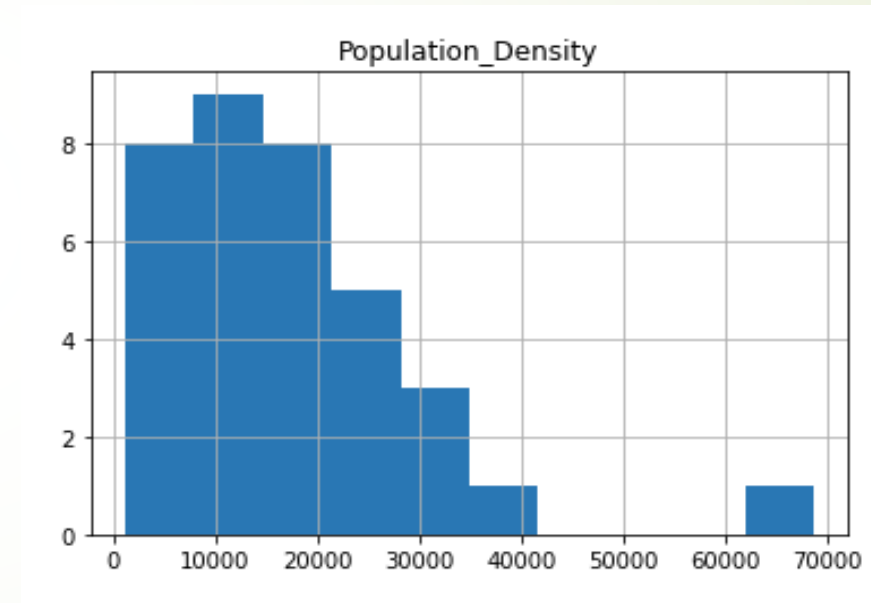
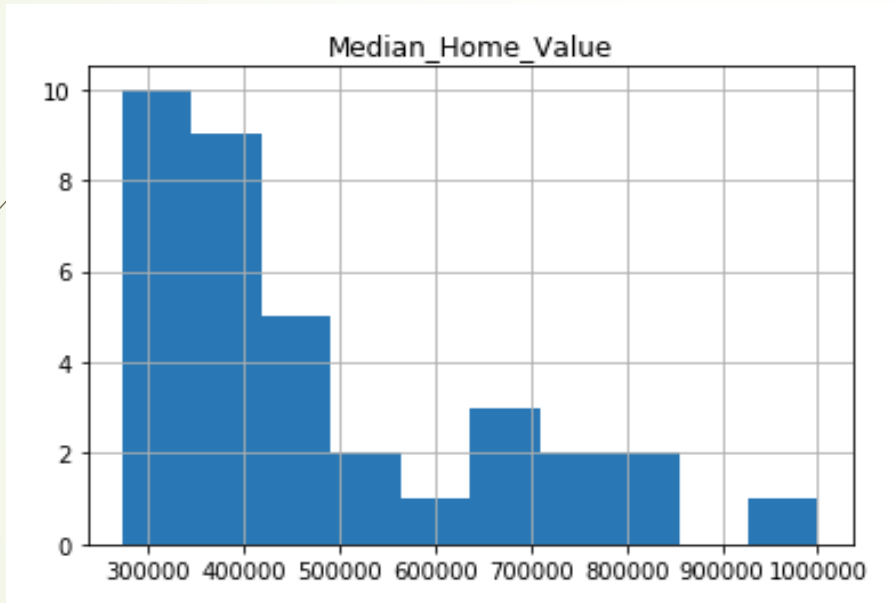


Data acquisition and preparation

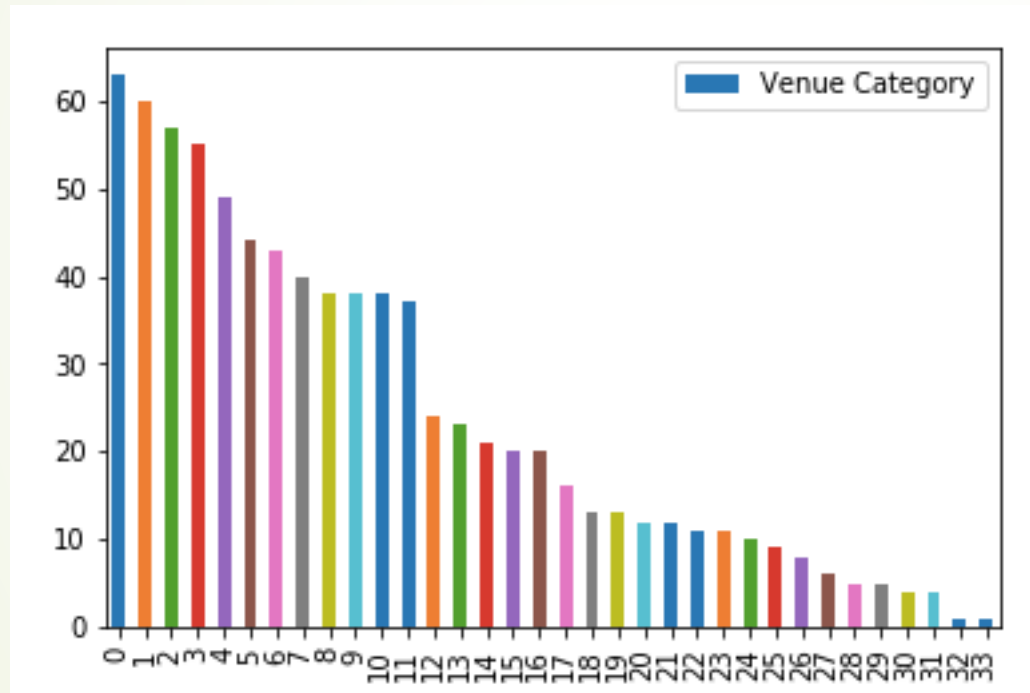


- List of Boston's zip codes obtained from Boston.gov
 - Duplicate zip codes were removed
- Population data and median home value obtained from uszipcode package
 - Some zip codes did not contain data and were removed
 - Some zip codes had similar names
- Venues data obtained from Foursquare via API calls

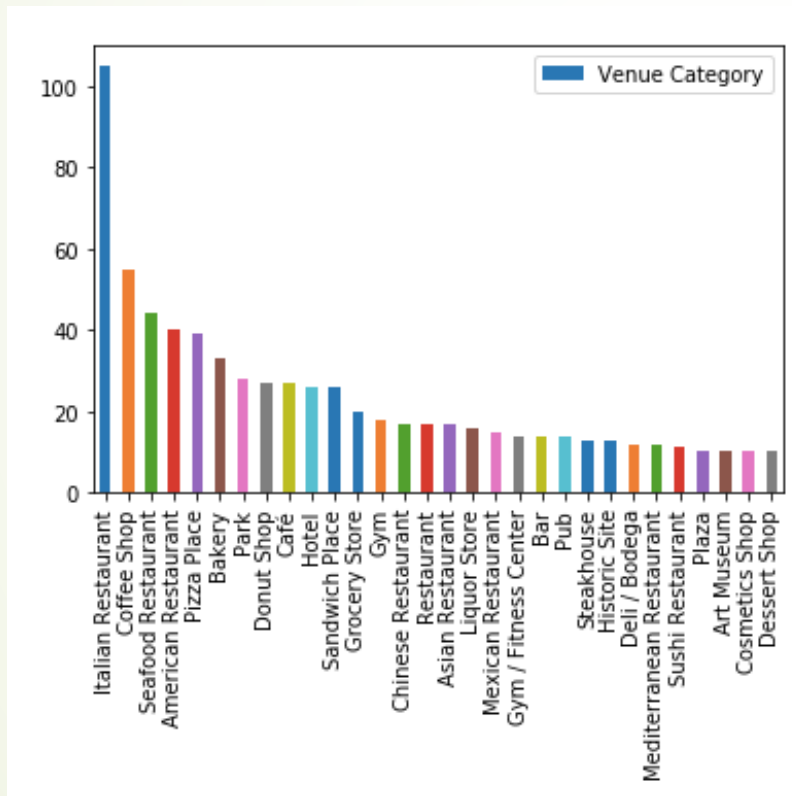
Median home values and population density are clustered on the lower end of the range



There is a wide spread in the unique venue categories of each neighborhood

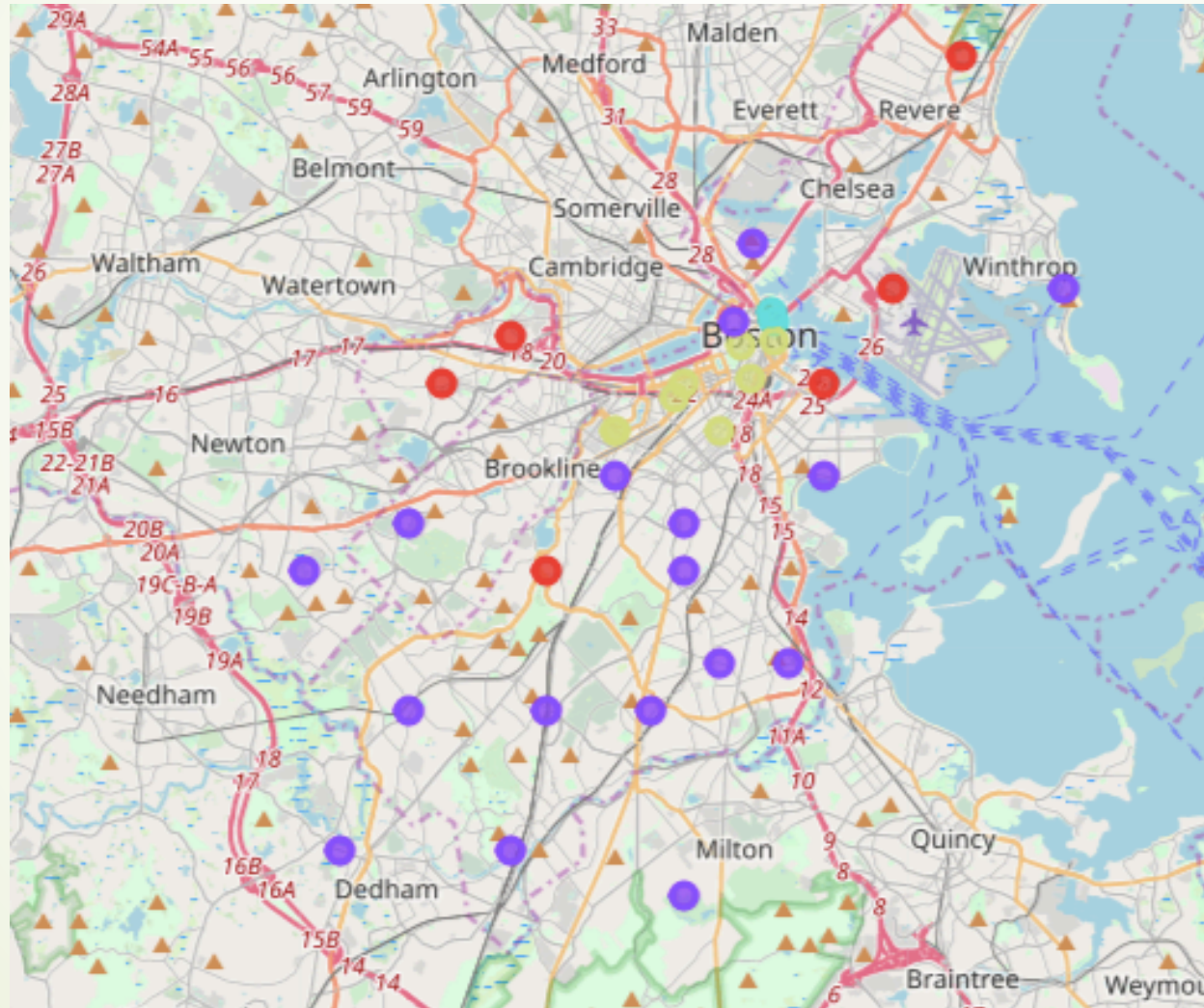


The most popular venue category by a wide margin is 'Italian Restaurant'



Venue Category	
Italian Restaurant	105
Coffee Shop	55
Seafood Restaurant	44
American Restaurant	40
Pizza Place	39
Bakery	33
Park	28
Donut Shop	27
Café	27
Hotel	26

K-means clustering applied with $k=4$; clusters produced appear to be intuitive



- Clusters make sense intuitively by location
- Neighborhoods in the center are clustered together while those on outside are clustered together



Cluster analysis



- Cluster 0 (Red) – Affordable with moderate diversity and availability of popular venues types
- Cluster 1 (Purple) - Generally poor diversity and non-availability of popular venues
- Cluster 2 (Cyan) – High end with high diversity and available popular venues
- Cluster 3 (Yellow) – High end with high diversity and available popular venues

* Cluster 2 contains only 2 neighborhoods and appear to be very similar. Geographically they are very near to each other, hence unsurprising.



Conclusion



- Simple tool used to determine which neighborhoods within Boston were most attractive with just a few parameters
- Intuitive clusters emerged with the use of k-means clustering
- Limitations: lack of a clear pattern within each cluster for population density and median home value – contribution might have been diluted with popular venue categories.