California Exodus:

Neighborhood Venue Analysis of Austin & San Francisco

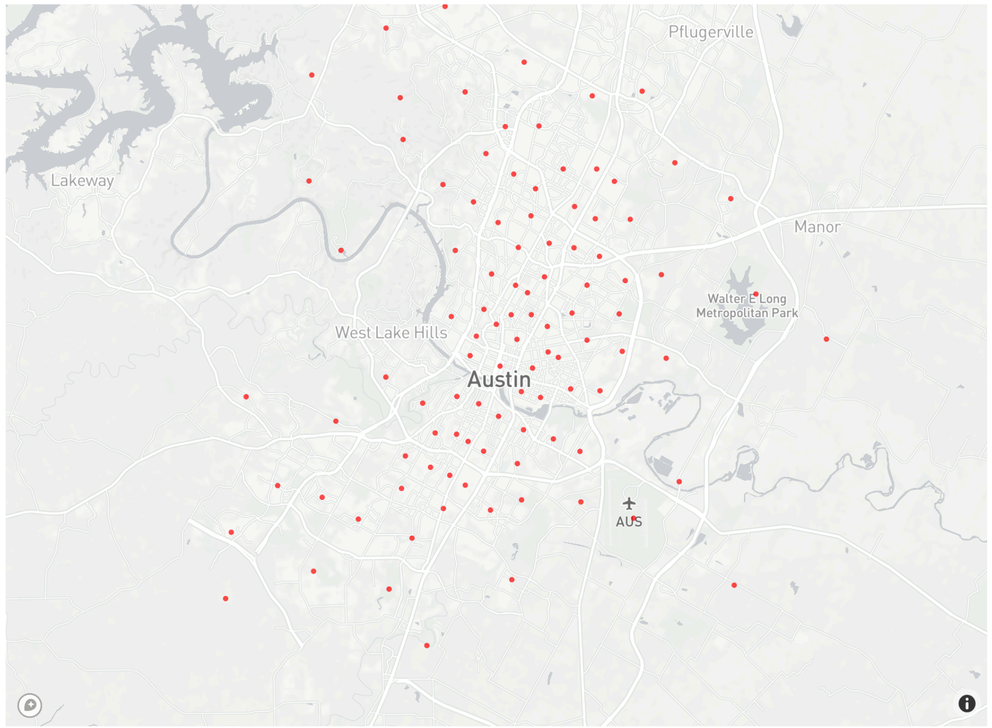
Introduction

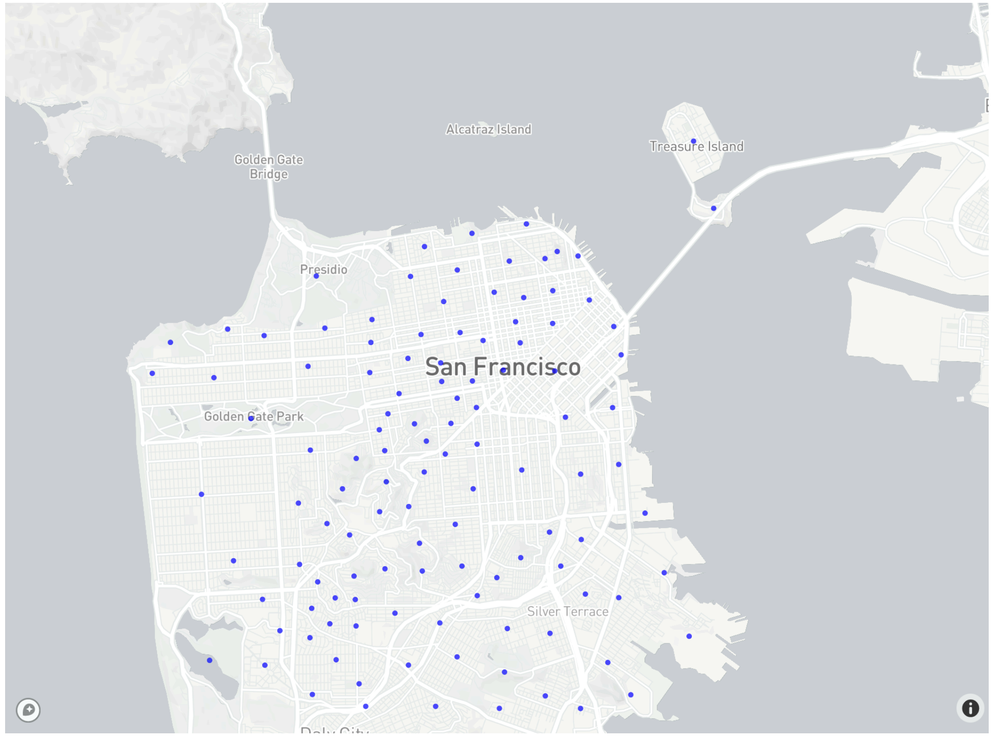
- Record numbers of residents moving out of California, especially from the Bay Area.
- Austin, Texas is a popular destination.
- Identifying similar neighborhoods in San Francisco and Austin can help the migration process easier.
  - Many people value the types of venues available near the neighborhood when choosing their new residence.

Data collection and cleaning

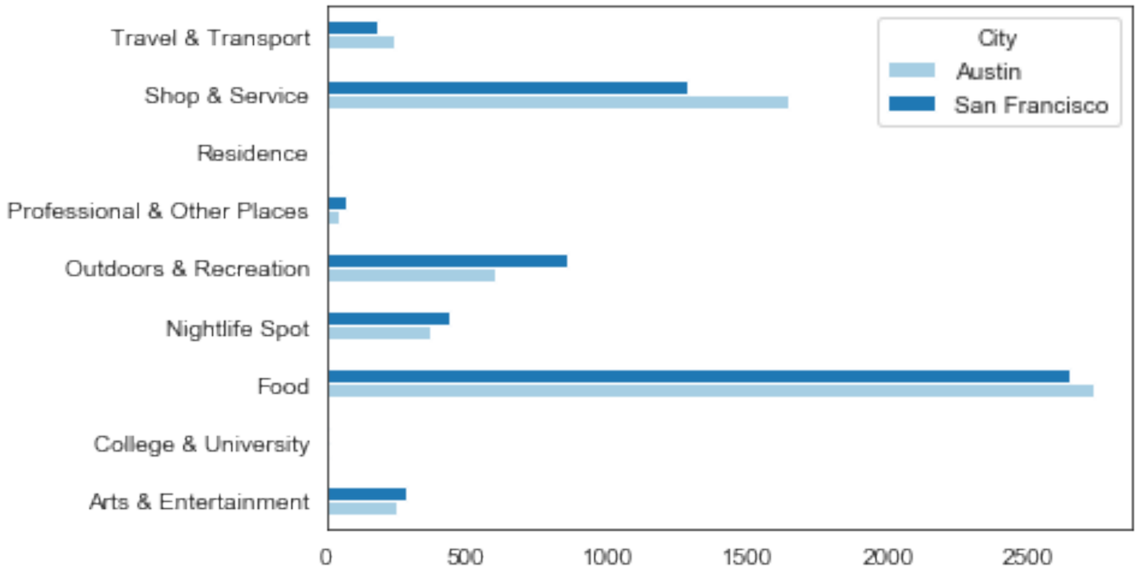
- Neighborhood geospatial data downloaded from SF Open Data Portal and Austin Open Data Portal.
- 117 neighborhoods in San Francisco and 103 neighborhoods in Austin.
- Shapely package to calculate centroids and extract coordinate data.
- Foursquare API to collect venues data. (11701 total venues)
- Subcategories assigned to 10 top-level venue categories
- Cleaned data contains 13 features.

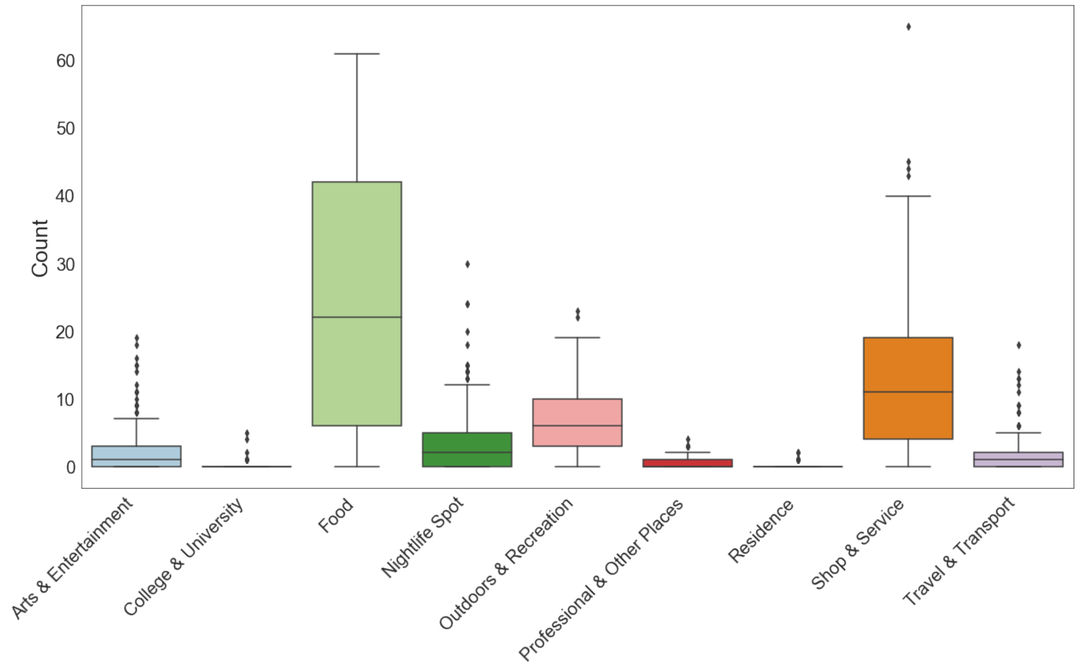
Coordinate data extracted



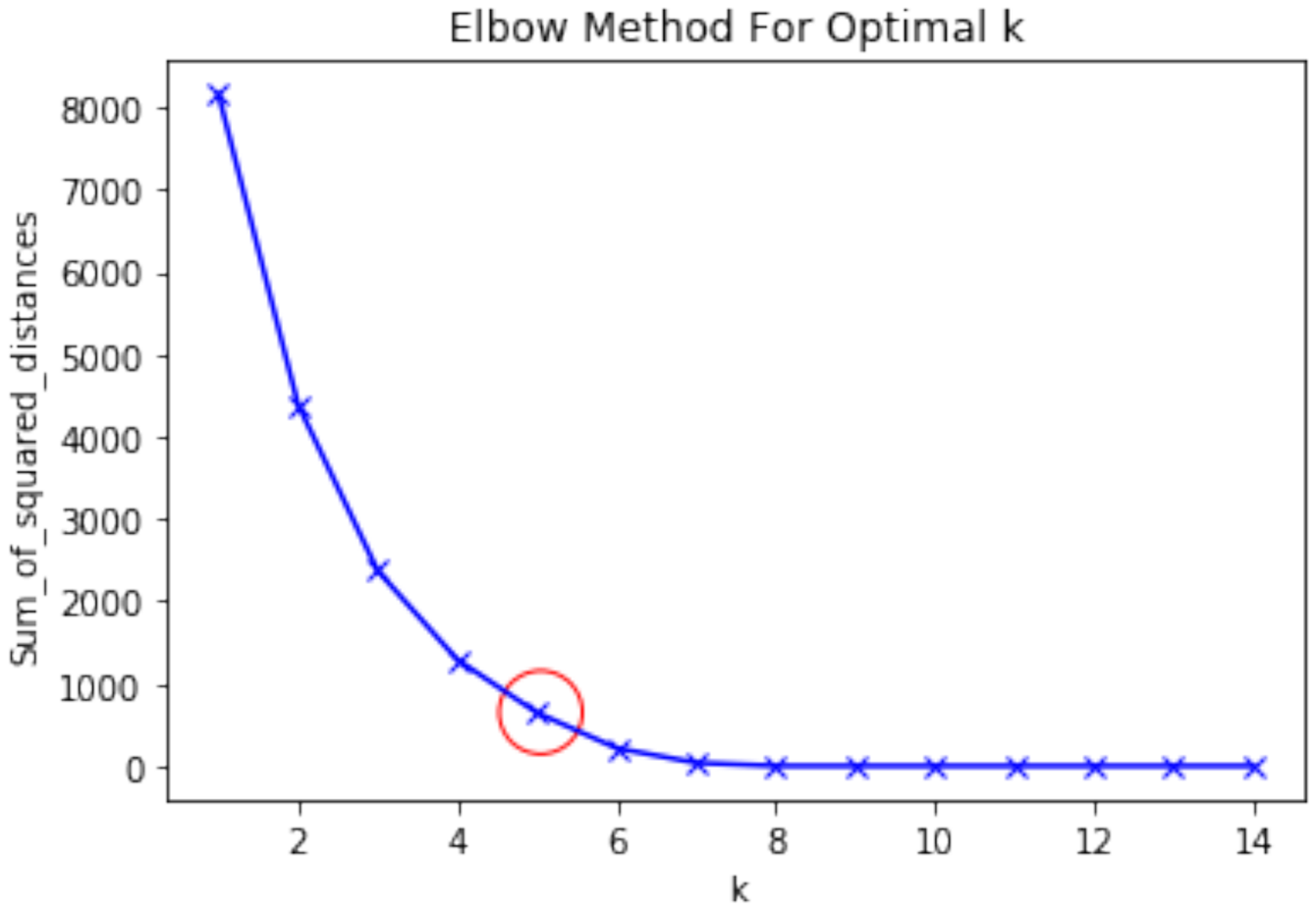


Exploratory Data Analysis



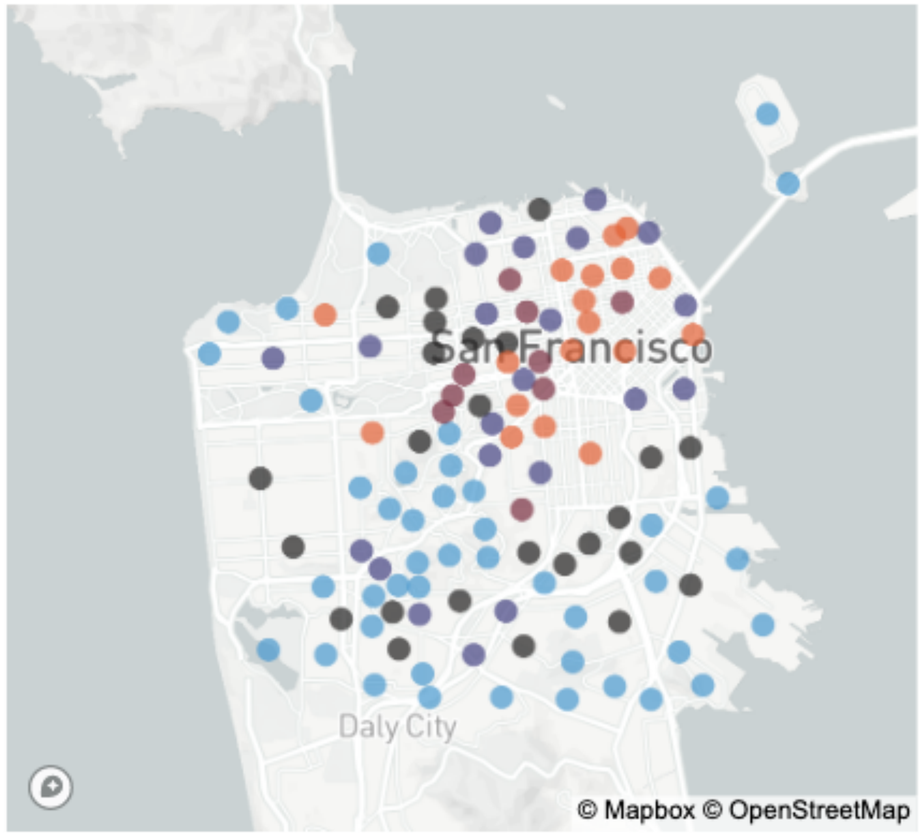


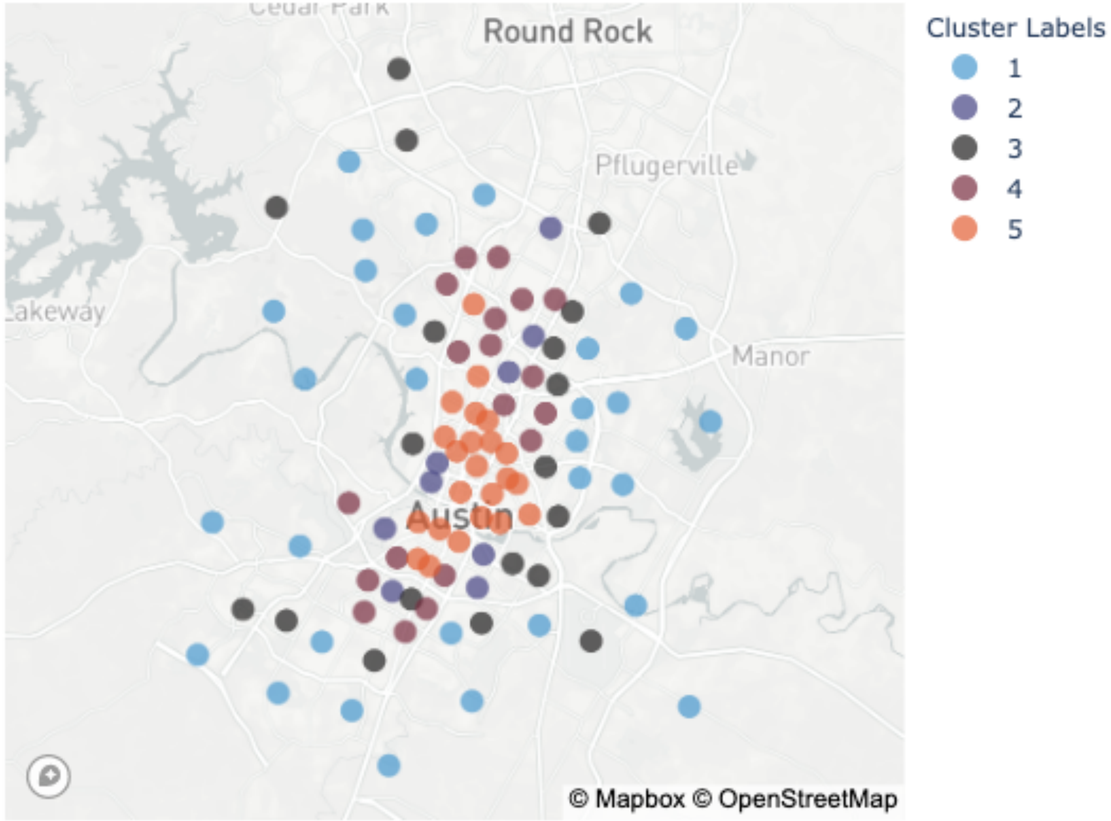
K-means Clustering: Optimal K



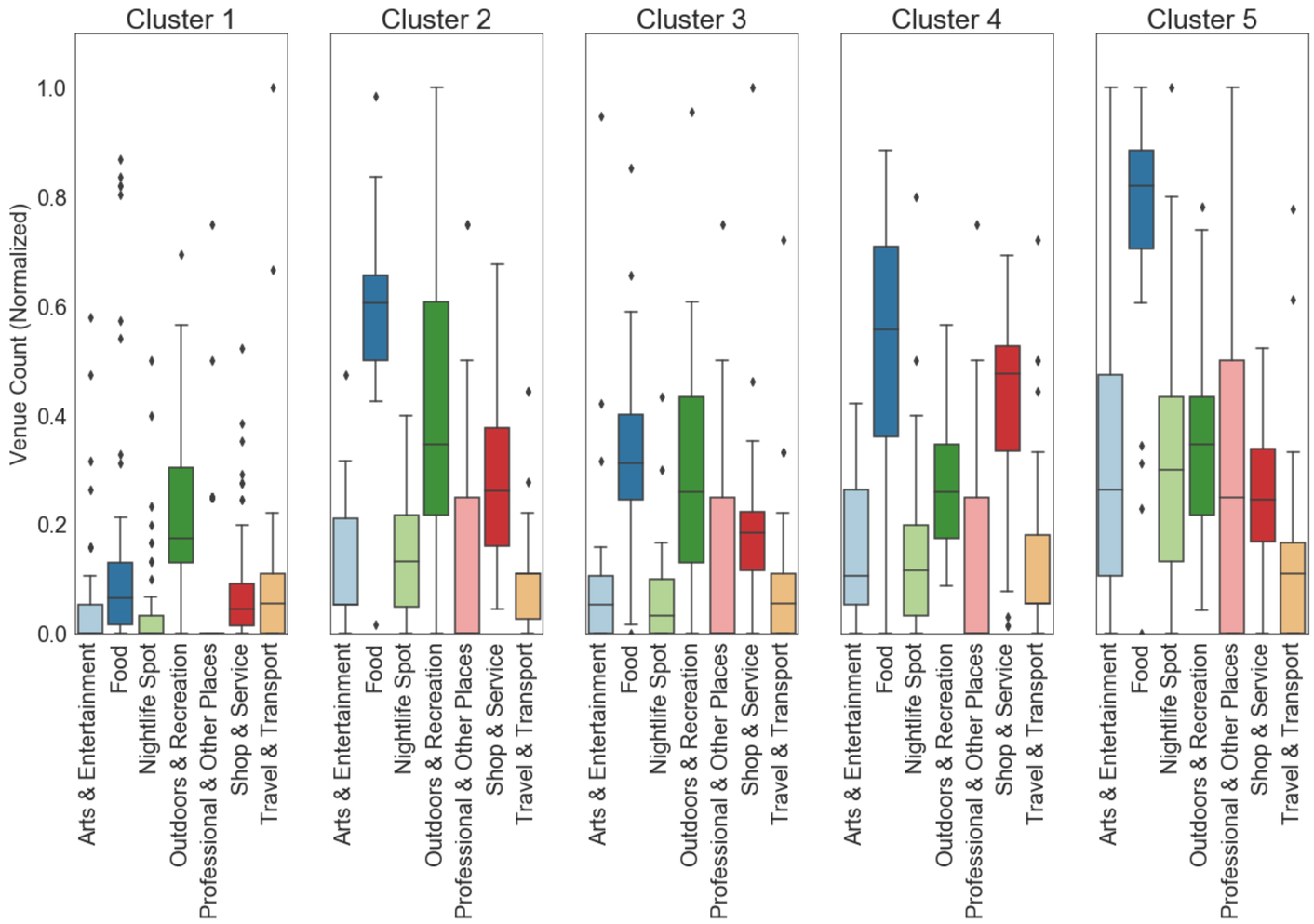
- Using the Elbow method, the optimal k for k-mean clustering algorithm was determined to be 5.

K-means Clustering: Results





K-means Clustering: Optimal K



- **Cluster 1:** Far from city center, Outdoor & Recreation
- **Cluster 2:** Good balance, just outside of city center
- **Cluster 3:** similar to cluster 2 but further away from city
- **Cluster 4:** many shops and services, right outside of city center
- **Cluster 5:** city center, hight venue counts across the board

Conclusion

- Built useful guide to group similar neighborhoods in Austin and San Francisco.
- Accuracy of the models has room for improvement
- Capture more variety using more variables
  - Demographics, socioeconomic data, crime rate, housing types, and political orientation.
  - More cities (New York City, Seattle, and Boston)
  - Recommendation algorithm