Optimizing Location for Starting a Pearl Milk Tea Store in the Bay Area

Author: Austin Tao

Importance

Location is an essential factor!

- Where a store is located can make or break its success

- Think:

    - Rent costs

    - Competitors

    - Geography

    - Demographic

    - Popularity

    - Distance to urban areas

   - etc.

How do we pick a new location?

Data Science!

- Name, address, and rating information of all boba shops in the Bay Area were compiled in bayarea_boba_shops.csv

- In total, 603 rows and 8 columns are in the csv file

- Data was mostly clean as-was, several NaN values were dropped and one redundant column was dropped

- **Data was analyzed after being divided by county**

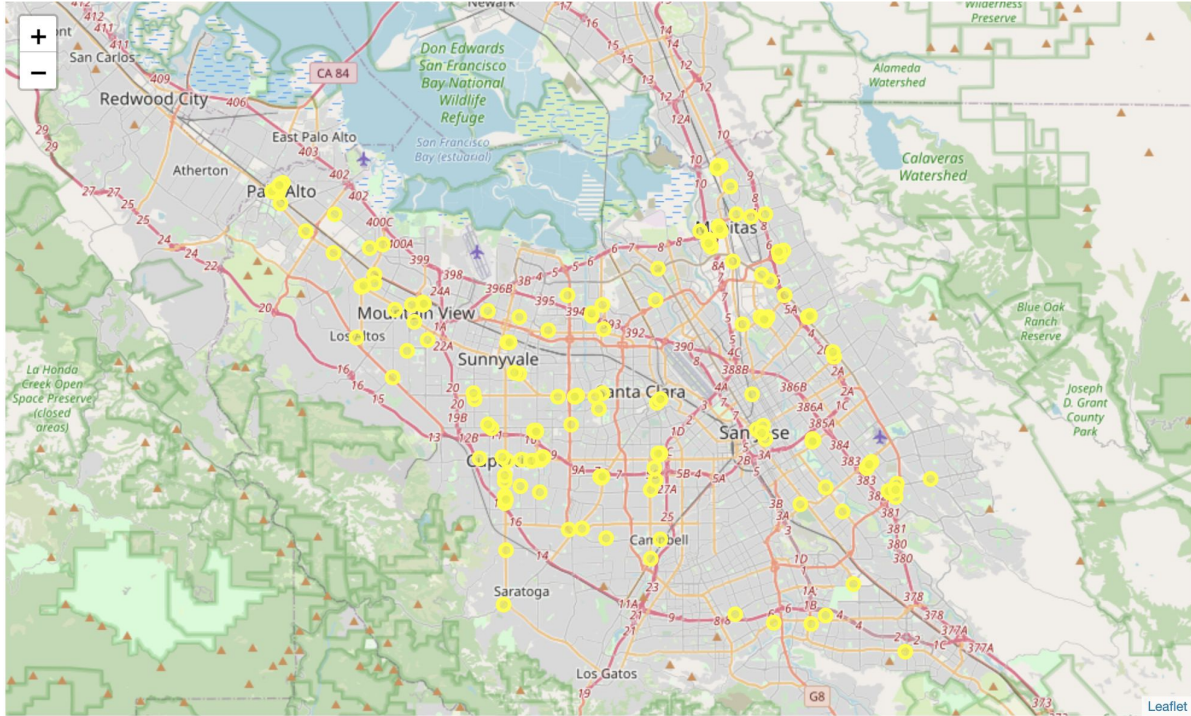- Remaining data was retrieved using the Foursquare API
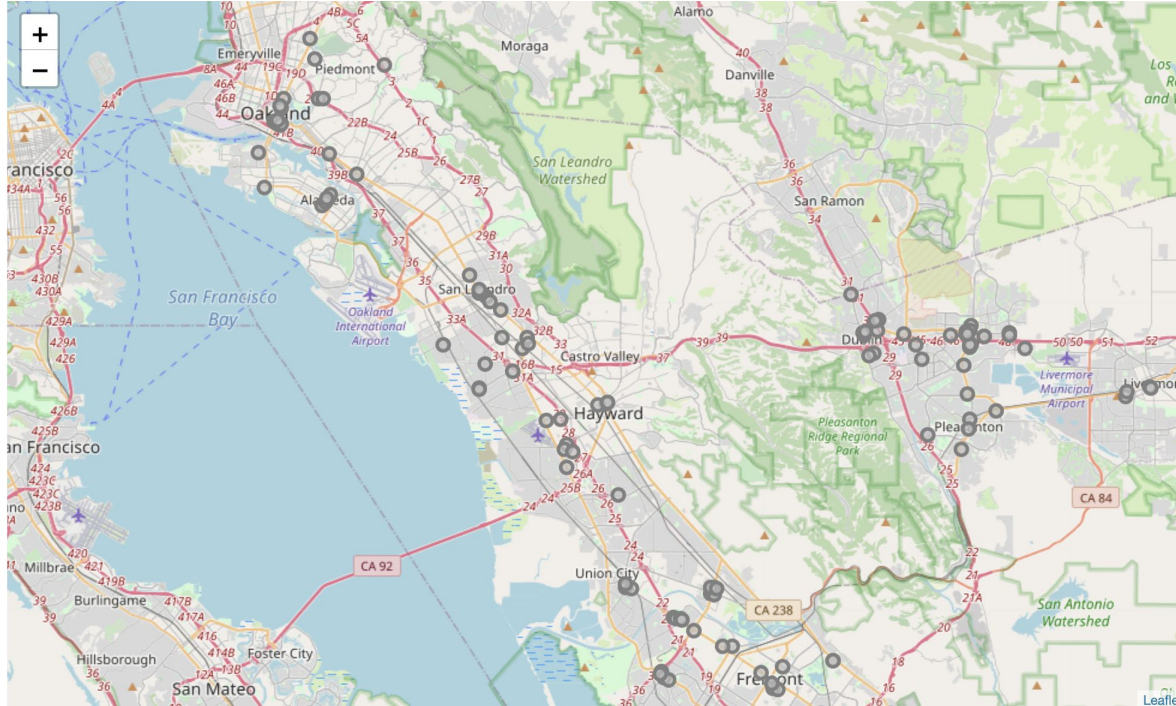
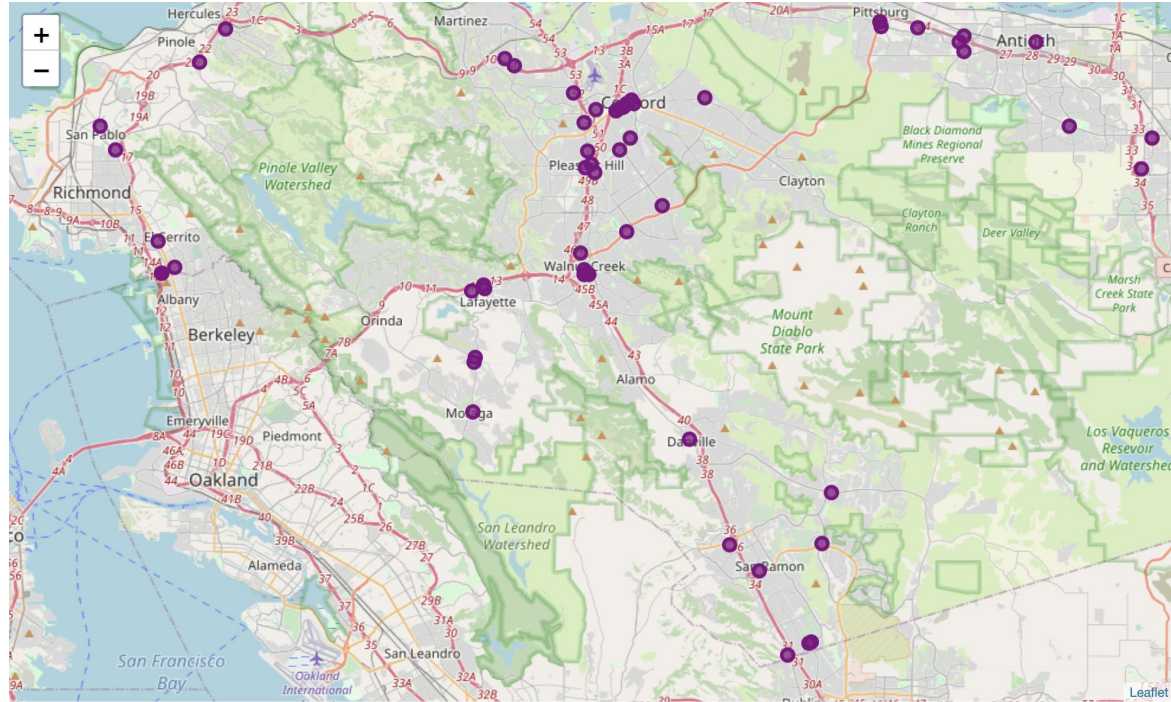# Mapping Data

San Francisco
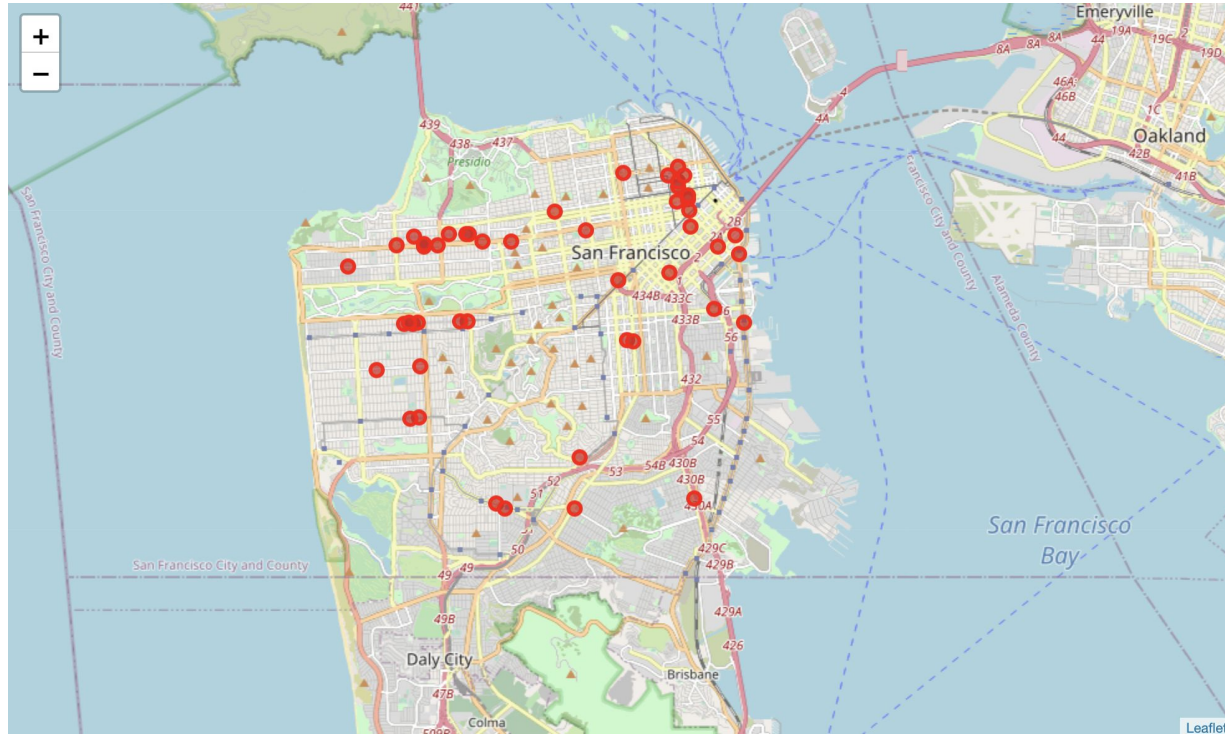
San Mateo
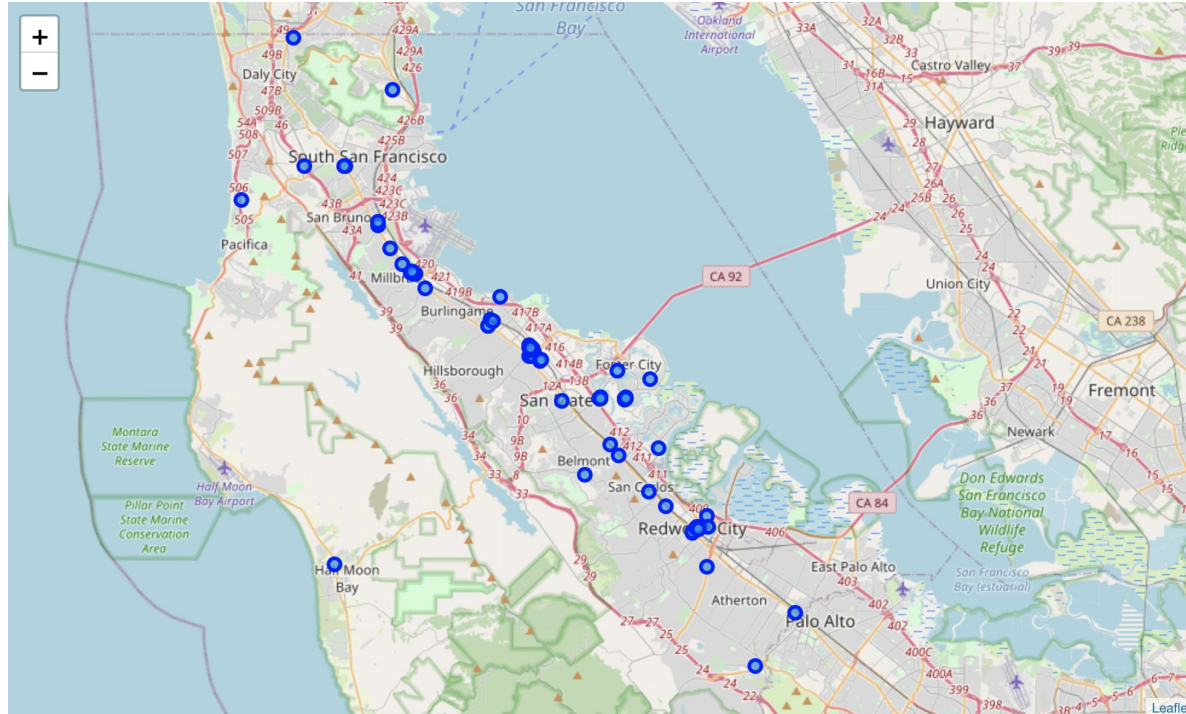
Santa Clara

Alameda

Contra Costa

Analysis

Analysis

1. Mapping (earlier slides)

2. Identifying top Venues:

3. Averaging rating data

4. K-means-clustering

County	Percentage of cities with high competition
S.F.	100% (only one city in county)
San Mateo	64%
Santa Clara	100% (with 15 cities in county)
Alameda	58%
Contra Costa	65%

```
In [27]:    1  findTopVenues(alameda_grouped, num_top_venues)

----Alameda----
               venue  freq
0        Sandwich Place  0.03
1       Bubble Tea Shop  0.03
2                   Bar  0.03
3    Chinese Restaurant  0.03
4           Pizza Place  0.02


----Albany----
               venue  freq
0        Thai Restaurant  0.07
1      Indian Restaurant  0.07
2                    Bar  0.04
3            Pizza Place  0.04
4              Pet Store  0.04


----Berkeley----
               venue  freq
0          Coffee Shop  0.05
1                 Café  0.04
2       Thai Restaurant  0.03
3        Sandwich Place  0.02
4         Hot Dog Joint  0.02


----Dublin----
               venue  freq
0          Coffee Shop  0.05
1        Sandwich Place  0.04
2    Japanese Restaurant  0.04
3               Bakery  0.03
4         Grocery Store  0.03
```
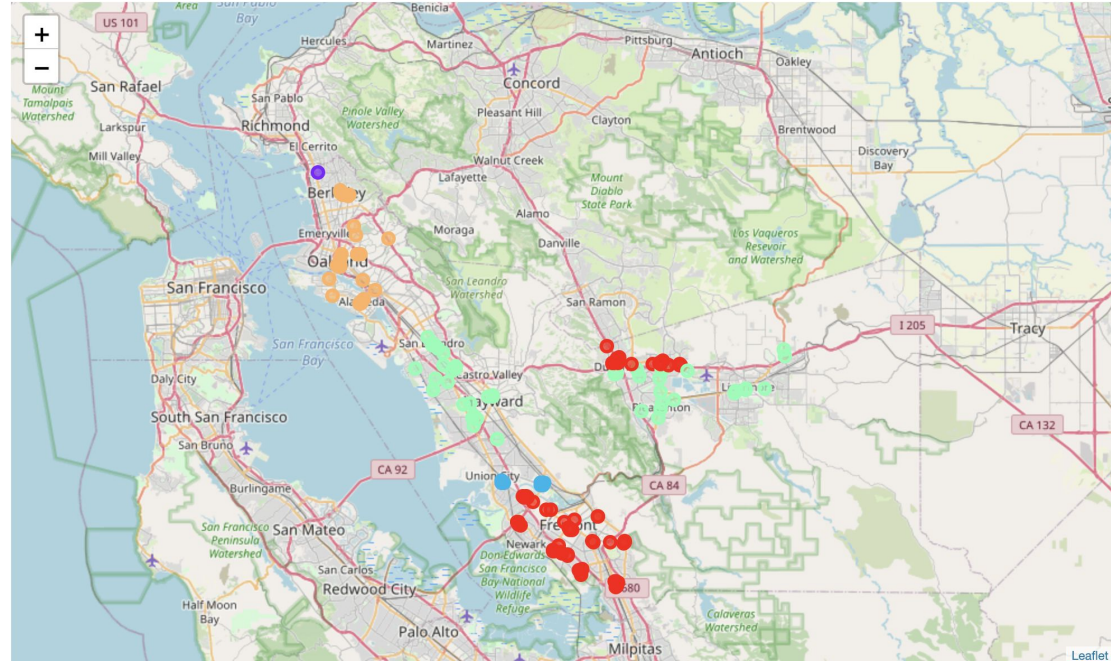
Results of analysis indicate that we should narrow our location down to Alameda County!

K-Means-Clustering: Alameda County

- K-Means-clustering groups like venue locations together
- Allows us to see the characteristics of particular locations

Conclusion and Future Directions

- **The optimal location for opening a new, non-franchised store in the Bay Area is in Union City, Alameda County**

- Future Directions:

    - Analyzing real-estate data in that area

    - Building cost functions including rent, labor costs, etc.

    - Modeling expected revenue

    - Risk management for expanding to future locations