

Identifying Business Opportunities for Up-and-Coming Neighborhoods in San Francisco

April 29, 2020

Introduction

Goal- Identify potential business opportunities in San Francisco

- Use real-estate information to find most expensive and cheapest neighborhoods in San Francisco
- Assume more affluent demographic will grow in San Francisco and other neighborhoods will reflect changing consumer profile
- Use Foursquare API to find relative frequencies of these venues

Data

Data used in this project:

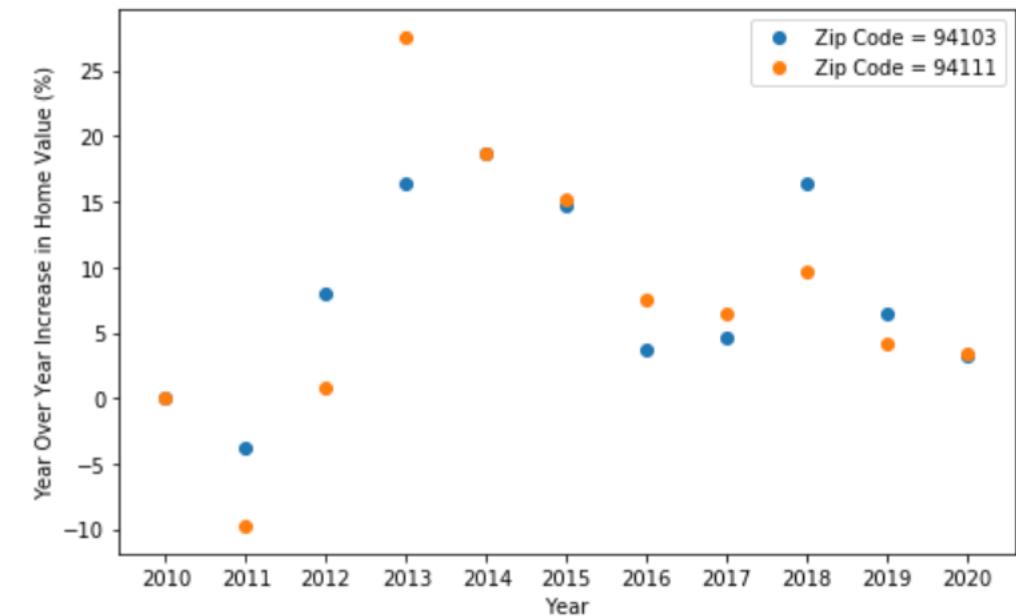
Single-family home values (SFHV) in the US obtained from Zillow

- Filtered by San Francisco zip codes
- SFHV averaged per year and sorted by 2020 values

Foursquare API

- Highest occurrence of venue type in each respective zip code

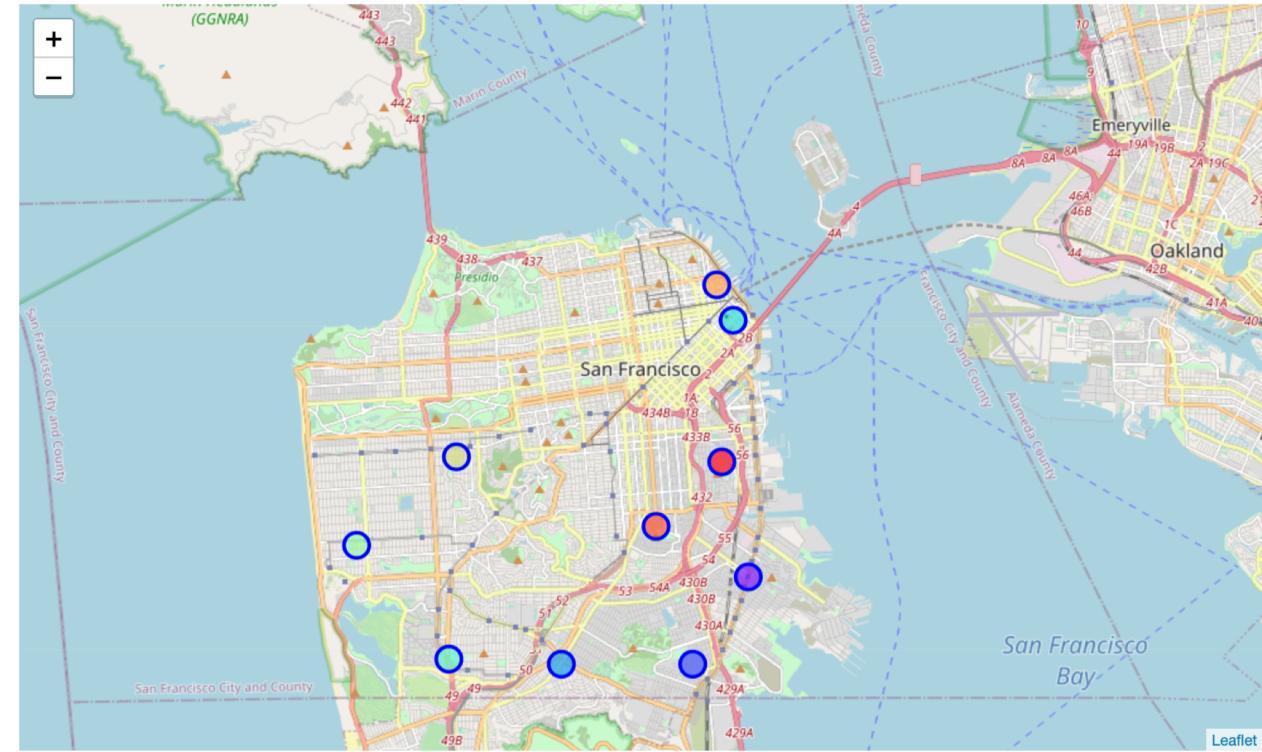
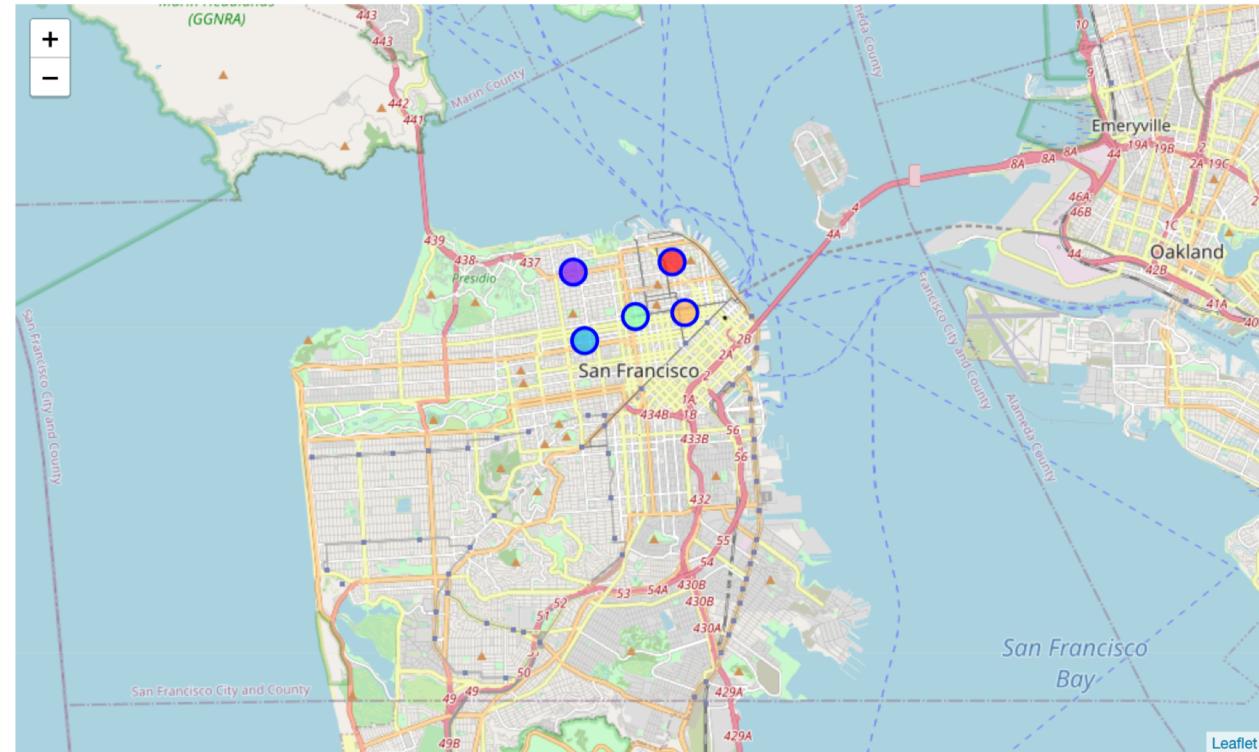
Comparison of most expensive and cheapest zip code



Interestingly, while the top zip code has seen a large increase in SFHV compared to the bottom zip code, no clear trends are observed in the year-over-year increases between the two.

For the sake of argument let's assume that the top 5 zip codes in terms of single family home value represent the more affluent demographic behind the skyrocketing housing prices in San Francisco. Over time we can reasonably expect that the home values of these zip codes will plateau as the SFHV of other zipcodes will rise as this demographic spreads to cheaper areas.

Location of top 5 and bottom 10 zip codes



Normalized Results of Relative Frequency of Venues

	Category	Top 5 Average	94105	94107	94110	94111	94112	94116	94122	94124	94132	94134
0	Italian Restaurant	0.0	-0.034	-0.034	-0.034	-0.034	-0.034000	-0.034000	-0.034	-0.034000	-0.034000	-0.034000
1	Coffee Shop	0.0	0.090	0.020	-0.030	0.070	-0.030000	0.048947	0.020	0.060909	-0.030000	-0.030000
2	Wine Bar	0.0	-0.024	-0.024	-0.024	-0.024	-0.024000	-0.024000	-0.024	-0.024000	-0.024000	-0.024000
3	Gym / Fitness Center	0.0	-0.024	-0.024	-0.024	-0.024	-0.024000	-0.024000	-0.024	-0.024000	-0.024000	-0.024000
4	Park	0.0	-0.022	0.018	-0.022	-0.022	-0.022000	-0.022000	-0.022	-0.022000	-0.022000	-0.022000
5	Pizza Place	0.0	-0.022	-0.022	-0.022	-0.022	0.068909	-0.022000	-0.022	0.023455	-0.022000	-0.022000
6	Bakery	0.0	-0.022	-0.022	0.018	-0.022	0.038606	-0.022000	0.038	0.068909	-0.022000	0.029282
7	Cocktail Bar	0.0	-0.018	-0.018	0.022	-0.018	-0.018000	-0.018000	-0.018	-0.018000	-0.018000	-0.018000
8	Café	0.0	-0.016	0.034	-0.016	-0.016	-0.016000	-0.016000	-0.016	-0.016000	0.031619	-0.016000
9	Men's Store	0.0	-0.016	-0.016	-0.016	0.024	-0.016000	-0.016000	-0.016	-0.016000	-0.016000	-0.016000

Italian restaurant may be inflated due to proximity of top 5 zip codes to Little Italy

ZipCode	ΔRelative Frequency	
0	94110	-0.03
1	94112	-0.03
2	94132	-0.03
3	94134	-0.03
4	94107	0.02
5	94122	0.02
6	94116	0.0489474
7	94124	0.0609091
8	94111	0.07
9	94105	0.09

Taking a look at coffee shop least competitive markets would be in:

- 94110
- 94112
- 94132
- 94134

Future Directions

Assumptions:

- Top 5 zip code demographic will be reflected in other zip codes due to expected cap in SFHV
 - Can verify using additional data sources including census data (household income, etc.)
- Relative frequency is best metric for relative business opportunity/competition
 - Foursquare API limited to 100 results & may be better to raster latitude/longitude coordinates within zip code to maximize results
 - Total and available commercial real estate may influence occurrence of these types of businesses (e.g. commercial district vs highly residential will skew results)
 - May be able to normalize by additional data e.g. total commercial real estate per zip code