Table of Contents

1. Introduction

For most people, purchasing a house is a big deal in their lives. In San Francisco Bay Area, the house price is always one of the the top 10 in the world. The house price is affected by many factors, like location, school district, size of the lot and the living area, the year of build, the condition of the house, the taste of interior design ect. So each house has its unique price.

For buyers, it’s always difficult to calculate a number to offer to the sellers? Of course the buyers have their own buyer agents who are professional realtors and familiar with the market. But, since the mechanism of the whole housing market in San Francisco Bay Area is based on sellers’ market, sometimes the buyer agents would suggest a higher price than it should be. Buyer agents are paid by 2% to 3% of the house price according to the contract they signed with their clients, that is, the buyers. If the buyer wins the bid and go through a smooth transaction process, his/her buyer agent get paid. The buyer and buyer agent have no idea about what offers the seller get from other potential buyers. How to win a bid? The one with highest price offered to the seller. So for buyer agents, they always an intention to push their clients to offer the highest price. Sometimes, the one who wins the competition has a much higher offer than the second one. How to know the house price and avoid fatty offer, that’s what I try to solve in this project.

The source of the data I will use in this project is from Redfin[1], an online real estate brokerage website.

The project is to target clients like potential home buyers in San Francisco Bay Area. Those people needs to know what factors have influence on house price and what is the estimate price for a particular property they are interested in.

---

[1] https://www.redfin.com/

2. Dataset

https://github.com/Yishi1215/House-Price-In-San-Francisco-Bay-Area/tree/master/Data%20Wrangling

The data comes from Redfin.com. Since Redfin has a limitation on data download, each csv file only contains at most 350 rows of data. All the data files are in the same folder, I used glob package to find all csv files in that folder and then used read_csv function to read every csv file into a data frame. Every data frame was appended to a list. Finally, merge all the data frames together using concat function and save the whole dataset into a new csv file.

The dataset contains properties from four counties in San Francisco Bay Area, there are San Francisco County, San Mateo County, Santa Clara County and Alameda County, representing the city, peninsula, south bay and east bay. All the properties were sold between year 2015 and year 2018. The properties were randomly chosen by Redfin.

Another dataset is city boundaries within the state of California, it's a shapefile. The data is from Government of California website[2].Cities from San Francisco County, San Mateo County, Santa Clara County and Alameda County have been exacted out. This shapefile data is used to draw house price choropleth map to provide better data visualization.

3. Data Facts

3.1. Data volume

After data cleaning, there are 32416 rows and 16 columns in the dataset.

3.2. Data Dictionary

[2] http://www.dot.ca.gov/hq/tsip/gis/datalibrary/Metadata/cities.html

Column Name	Description	Type
SOLD DATE	On which date the property was sold	object
PROPERTY TYPE	Single family house/ townhouse/ condo	object
ADDRESS	Property address	object
CITY	City of the property	object
ZIP	Zip code of the property	object
PRICE IN K	Sold price of the property in thousand	float64
BEDS	Number of bedrooms	float64
BATHS	Number of bathrooms	float64
SQUARE FEET	Living space of the property	float64
LOT SIZE	Lot size of the property	float64
YEAR BUILT	Year of the property was built	float64
$/SQUARE FEET	Sold price/square feet	float64
HOA/MONTH	HOA per month	float64
LATITUDE	Latitude of the property	float64
LONGITUDE	Longitude of the property	float64
COUNTY	County of the property	object

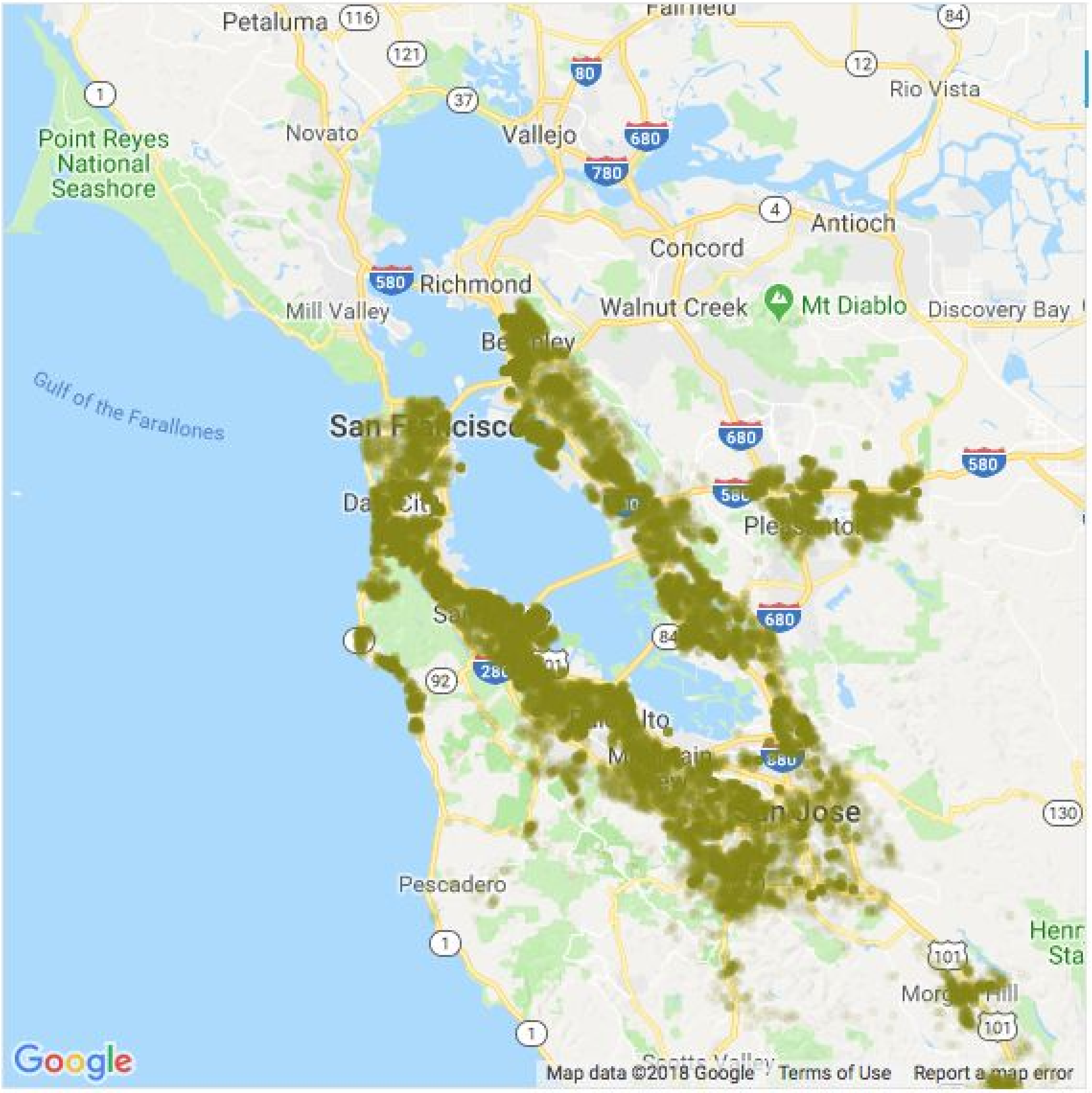
4. Exploratory Data Analysis

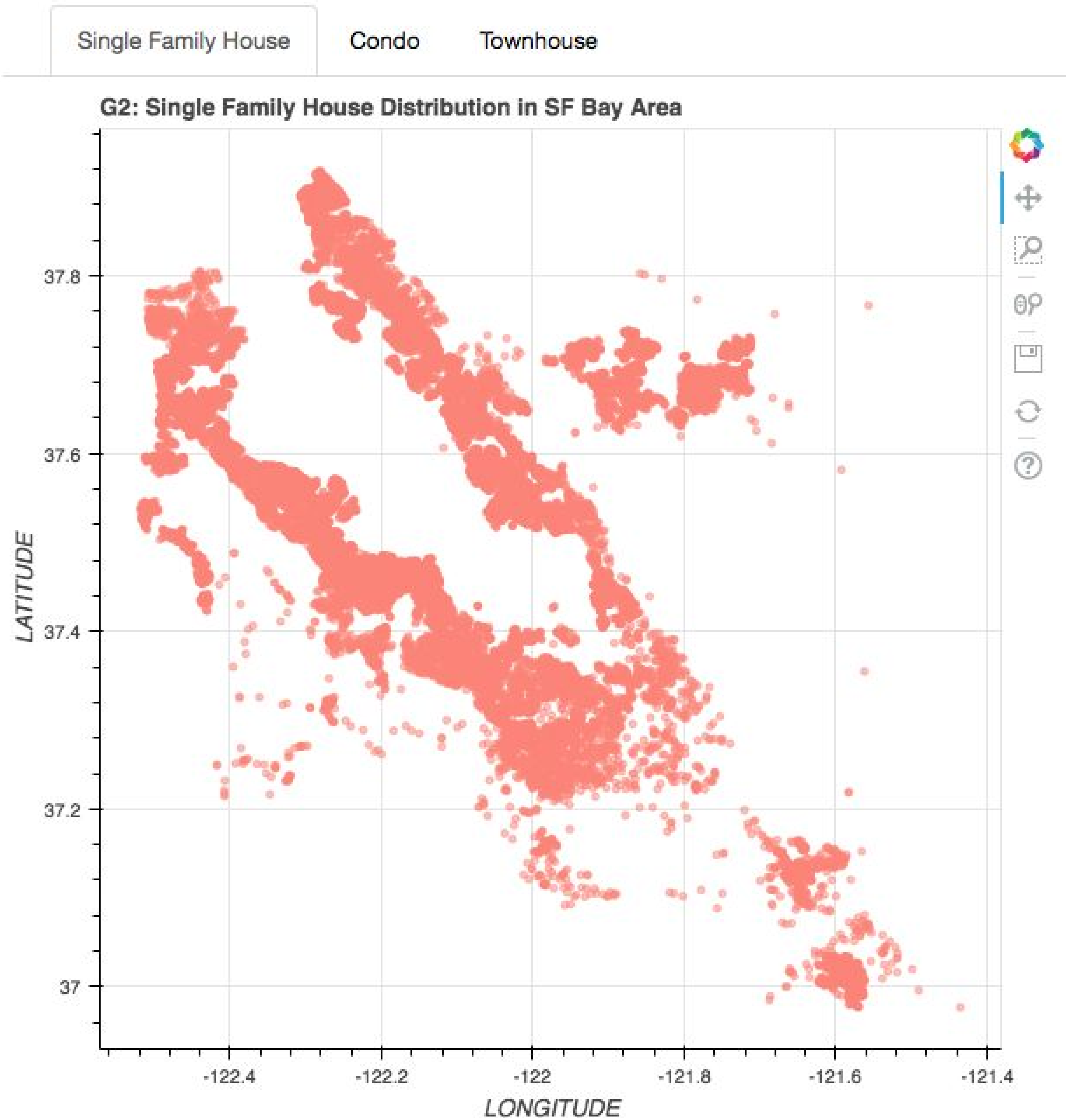
Exploratory data analysis is an approach to analyze datasets to summarize their main characteristics, often with visual methods. A statistical model may be used, but primarily EDA is for seeing what the data can tell us prior to the formal modeling or hypothesis testing task.

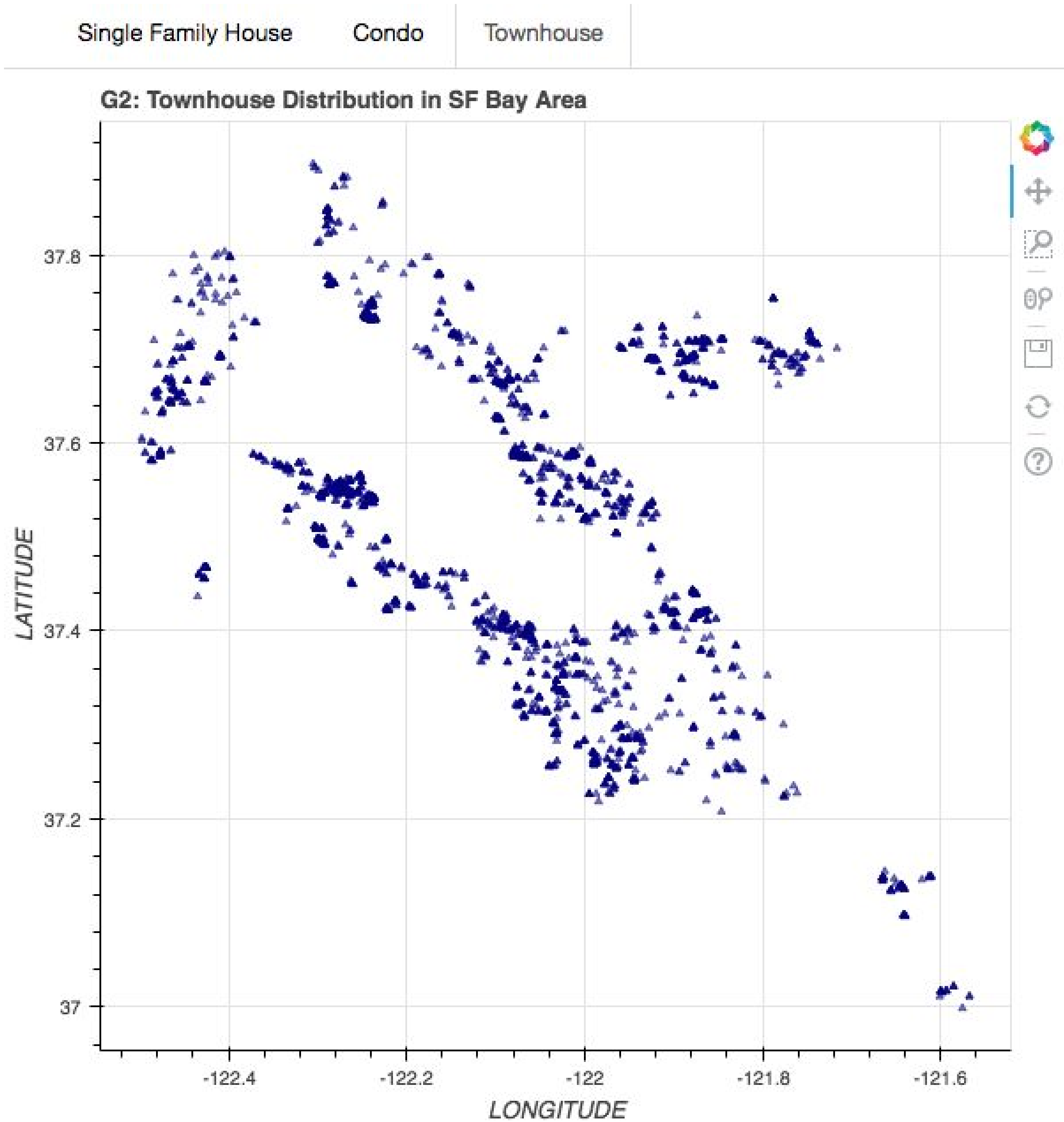
4.1. House Distribution Map

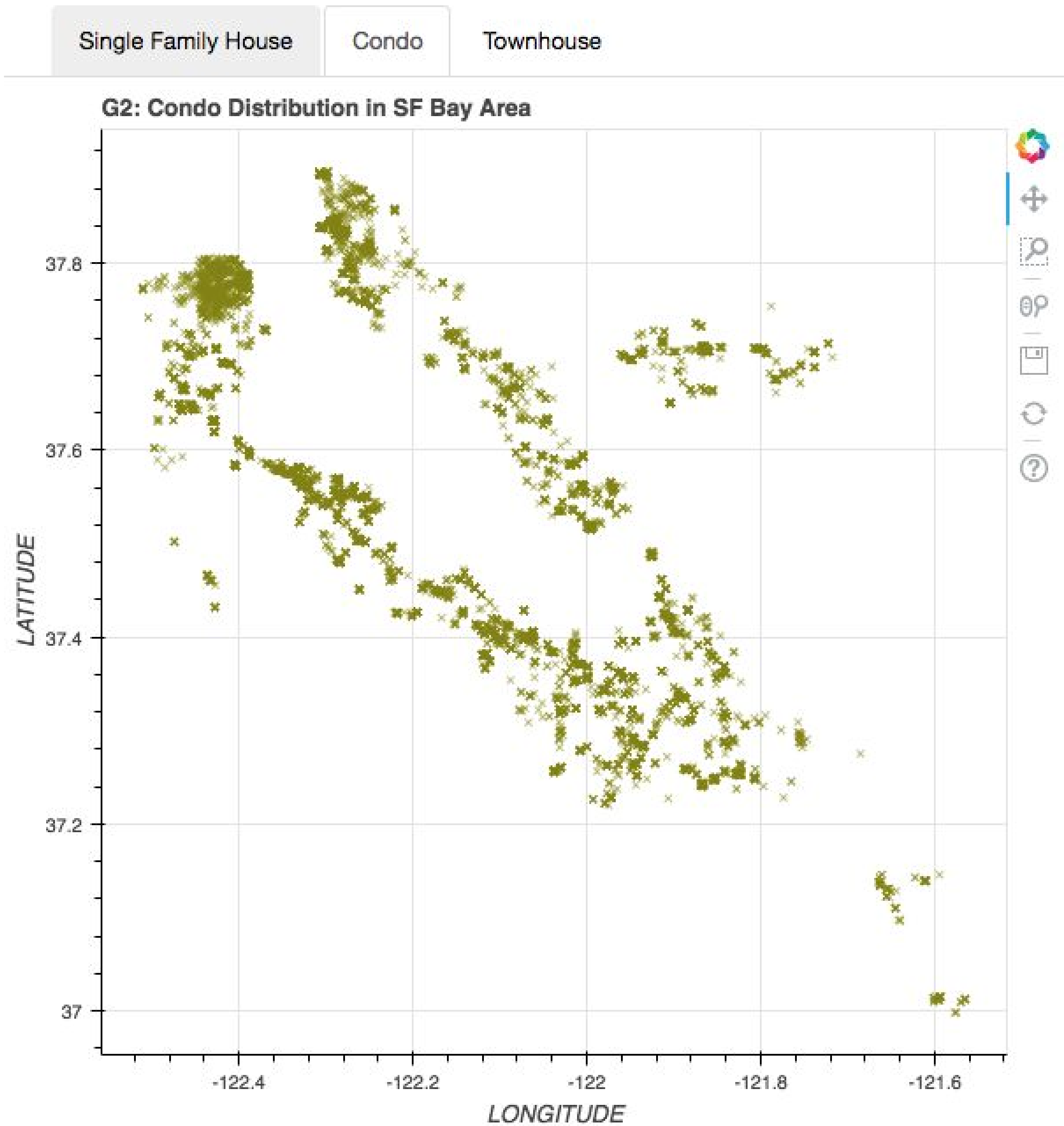
https://github.com/Yishi1215/House-Price-In-San-Francisco-Bay-Area/blob/master/Data%20Visulization/Google_Map.ipynb

G3: House Distribution in SF Bay Area




Single Family House
Condo
Townhouse
G2: Single Family House Distribution in SF Bay Area
37.8
37.6
37.4
37.2
37
LATITUDE
-122.4
-122.2
-122
-121.8
-121.6
-121.4
LONGITUDE


Single Family House
Condo
Townhouse
G2: Townhouse Distribution in SF Bay Area
LATITUDE
37.8
37.6
37.4
37.2
37
-122.4
-122.2
-122
-121.8
-121.6
LONGITUDE



According to the house distribution plot on Google map, the properties are concentrated along San Francisco Bay and demonstrate a banding distribution with two sets direction.

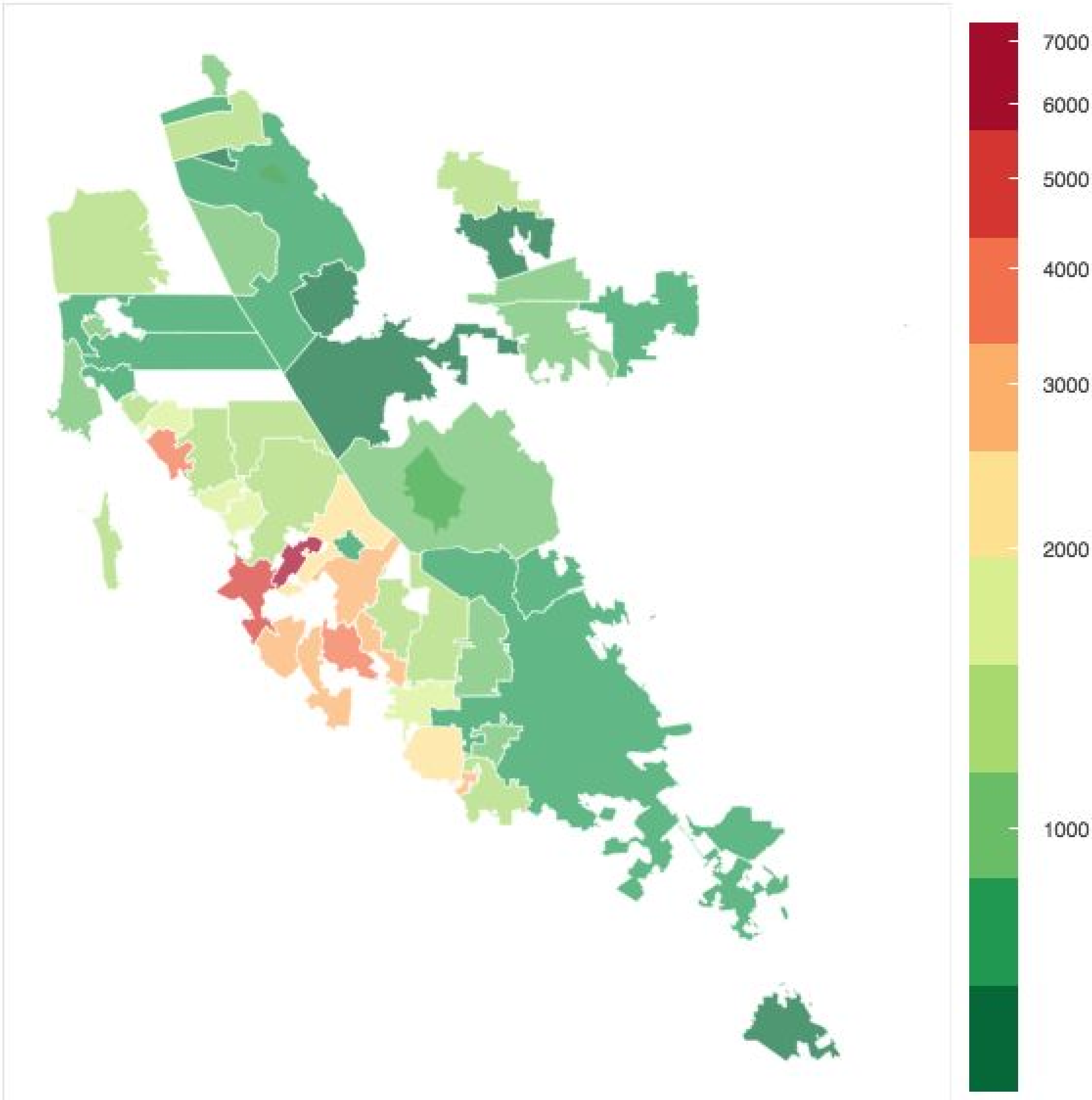
The number of single family house is larger than the other two types of property, which reflects preference of residents in this area.

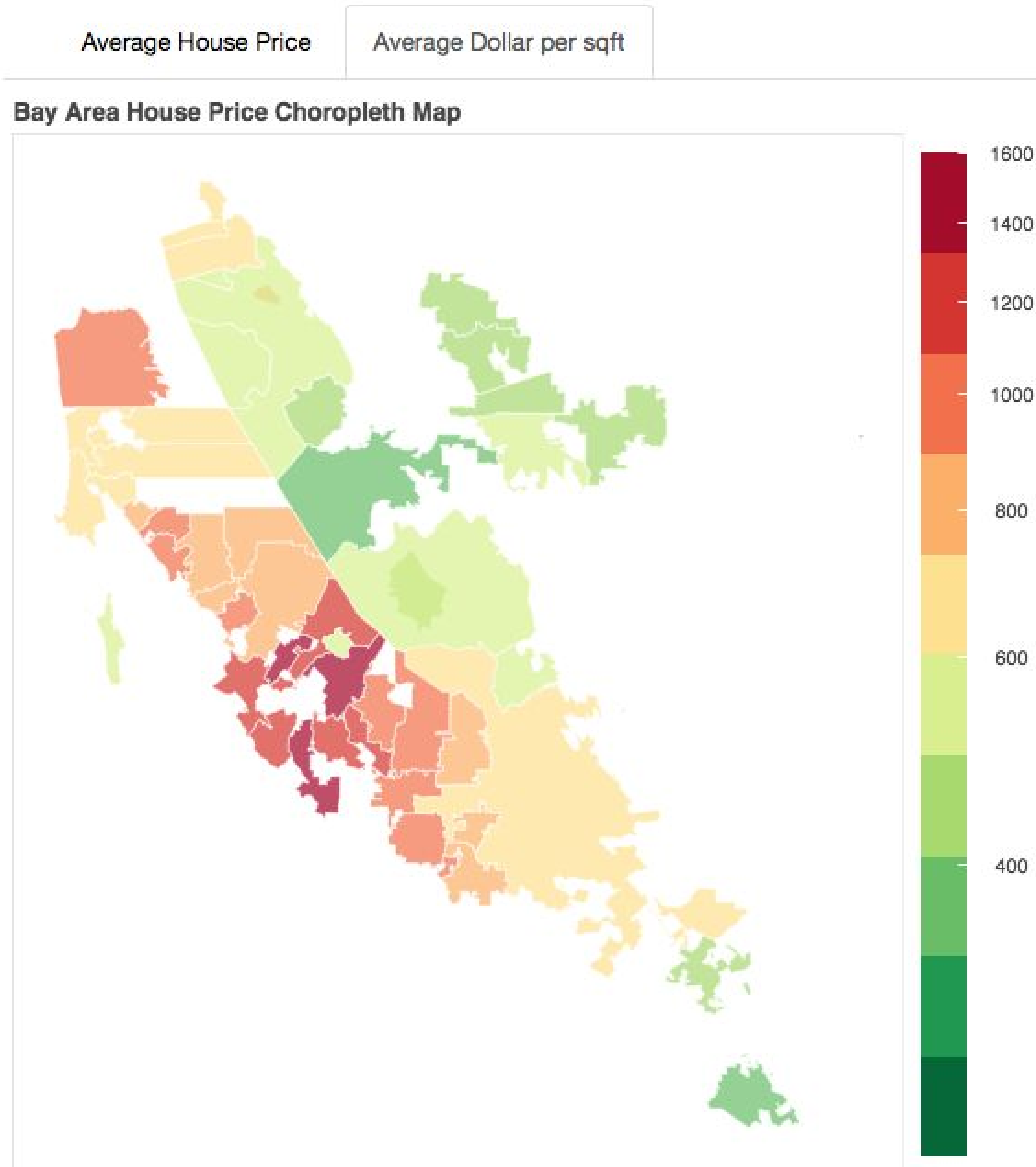
4.2. House Price Choropleth Map

https://github.com/Yishi1215/House-Price-In-San-Francisco-Bay-Area/blob/master/Data%20Visulization/Bay%20Area%20House%20Price%20Map.ipynb

Average House Price | Average Dollar per sqft

Bay Area House Price Choropleth Map

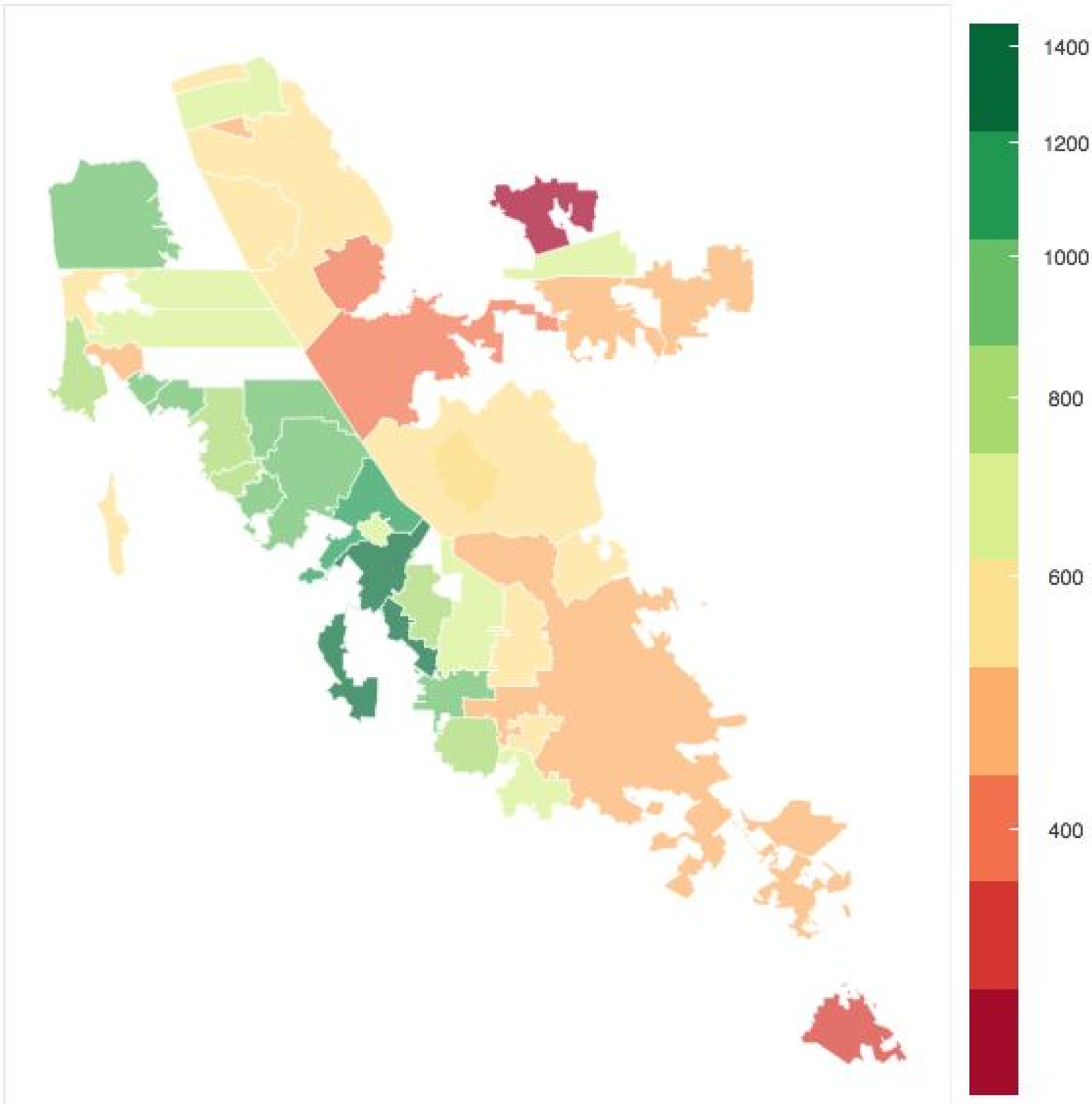


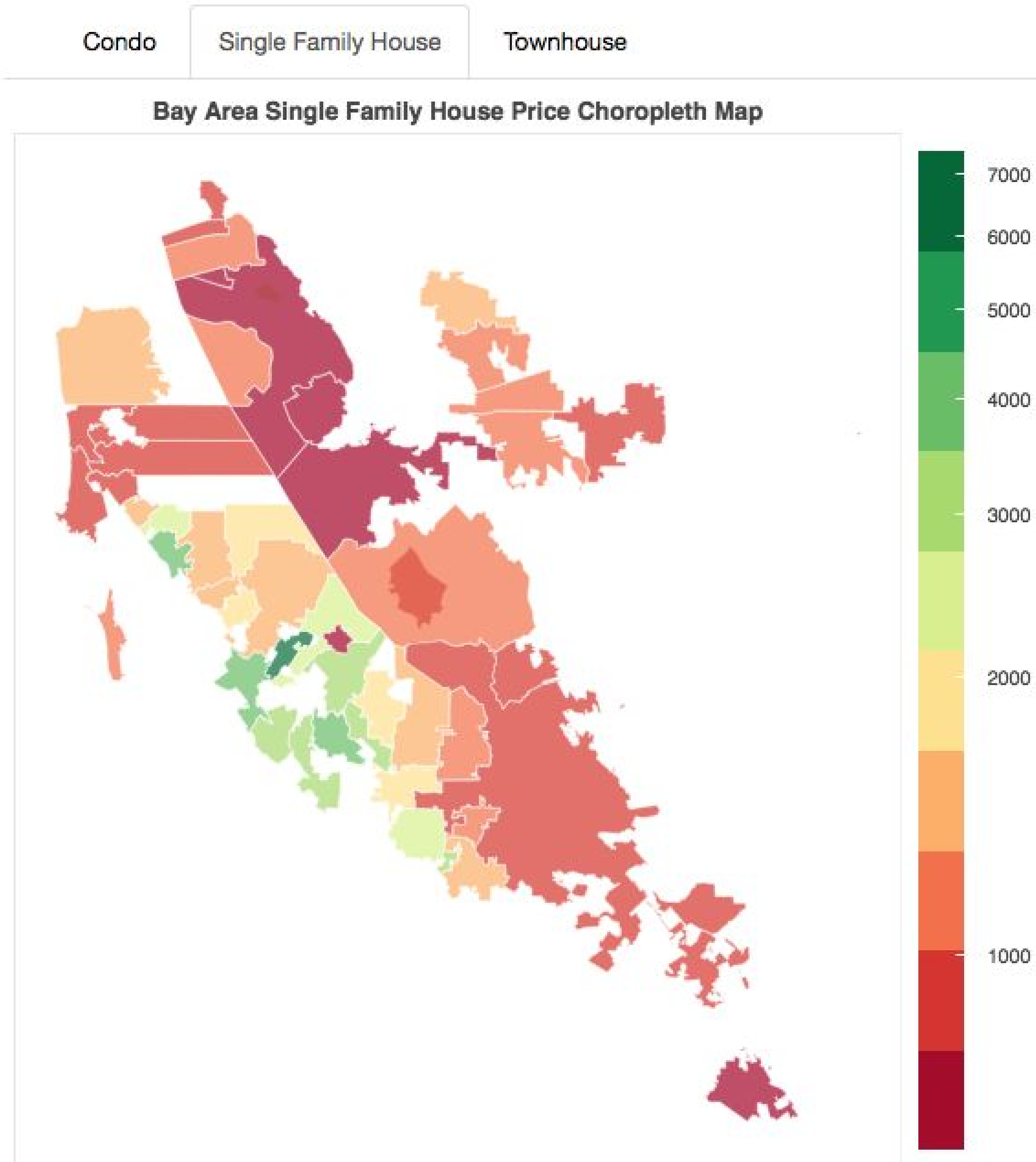


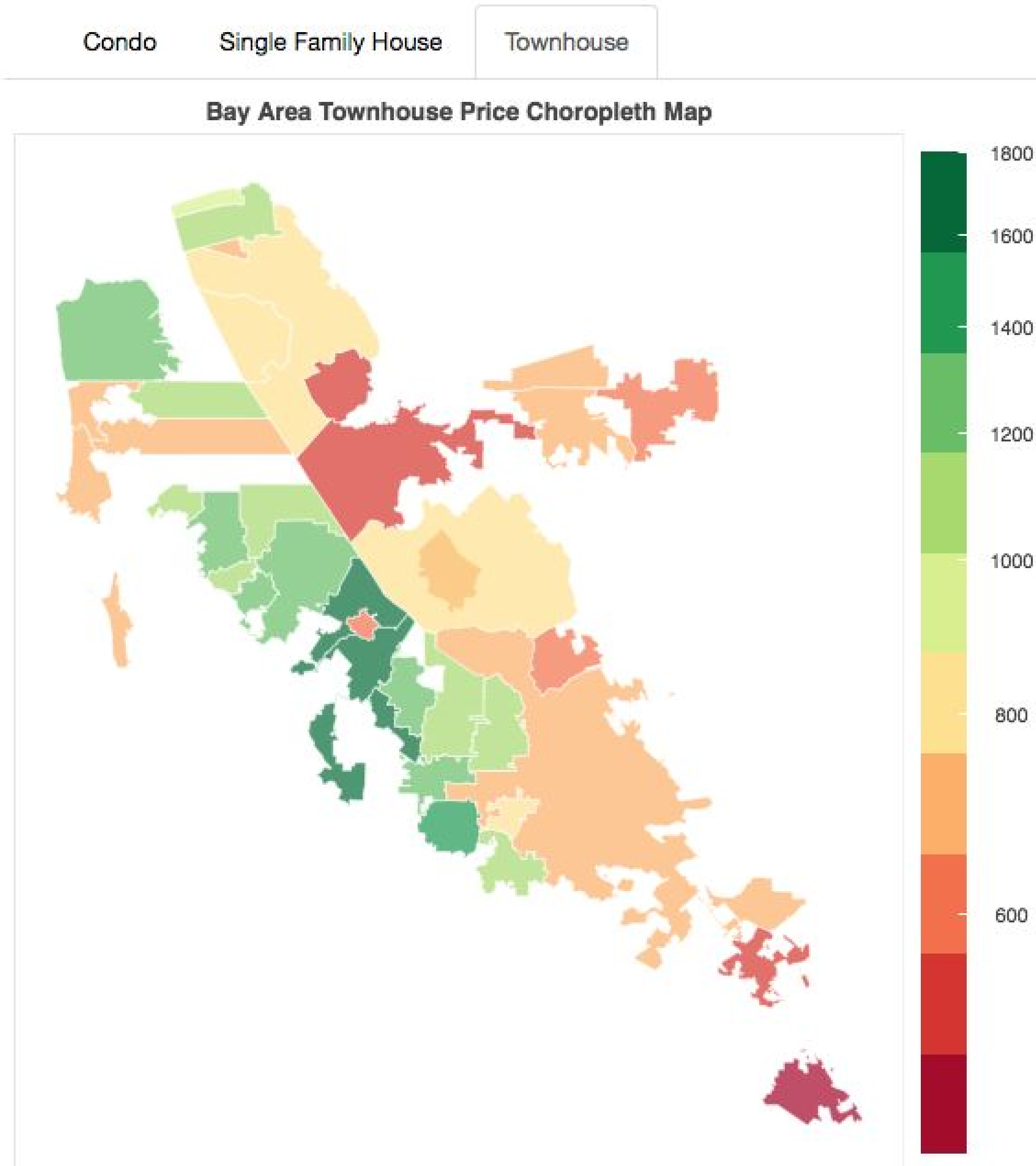
In the first map, tab 1 is the average house price based on city. The high-end properties are gathered in Peninsular and South Bay cities like Hillsborough, Palo Alto, Atherton, Woodside, Menlo Park, Portola Valley, Los Altos, Los Altos Hills, Saratoga, Monte Sereno. Tab 2 is the average dollar per square footage based on city. From San Francisco to South Bay Area are heat up. Unit price of East Bay Areas is lower.

Condo Single Family House Townhouse

Bay Area Condo Price Choropleth Map



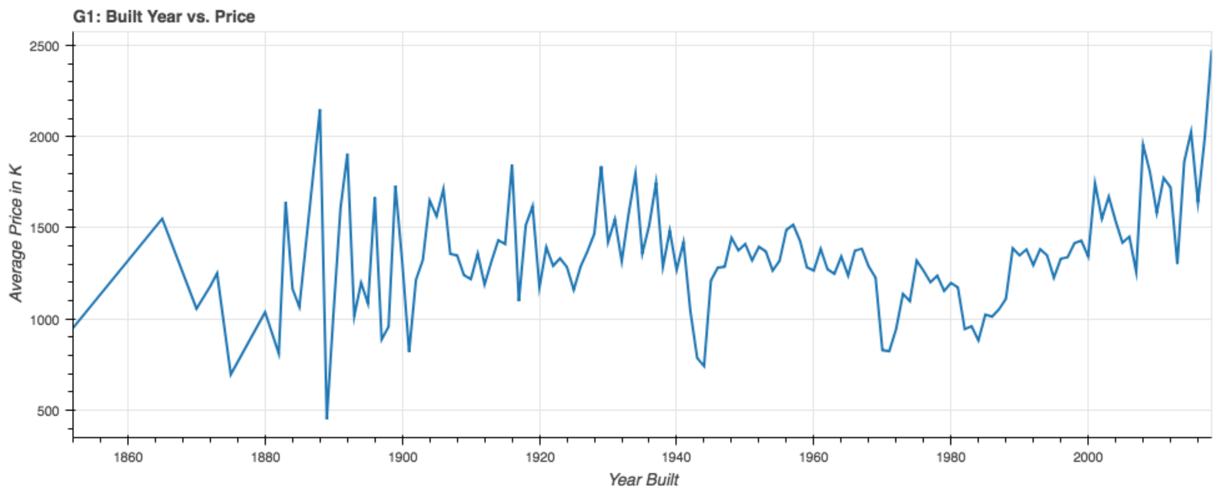

Condo
Single Family House
Townhouse
Bay Area Single Family House Price Choropleth Map
7000
6000
5000
4000
3000
2000
1000

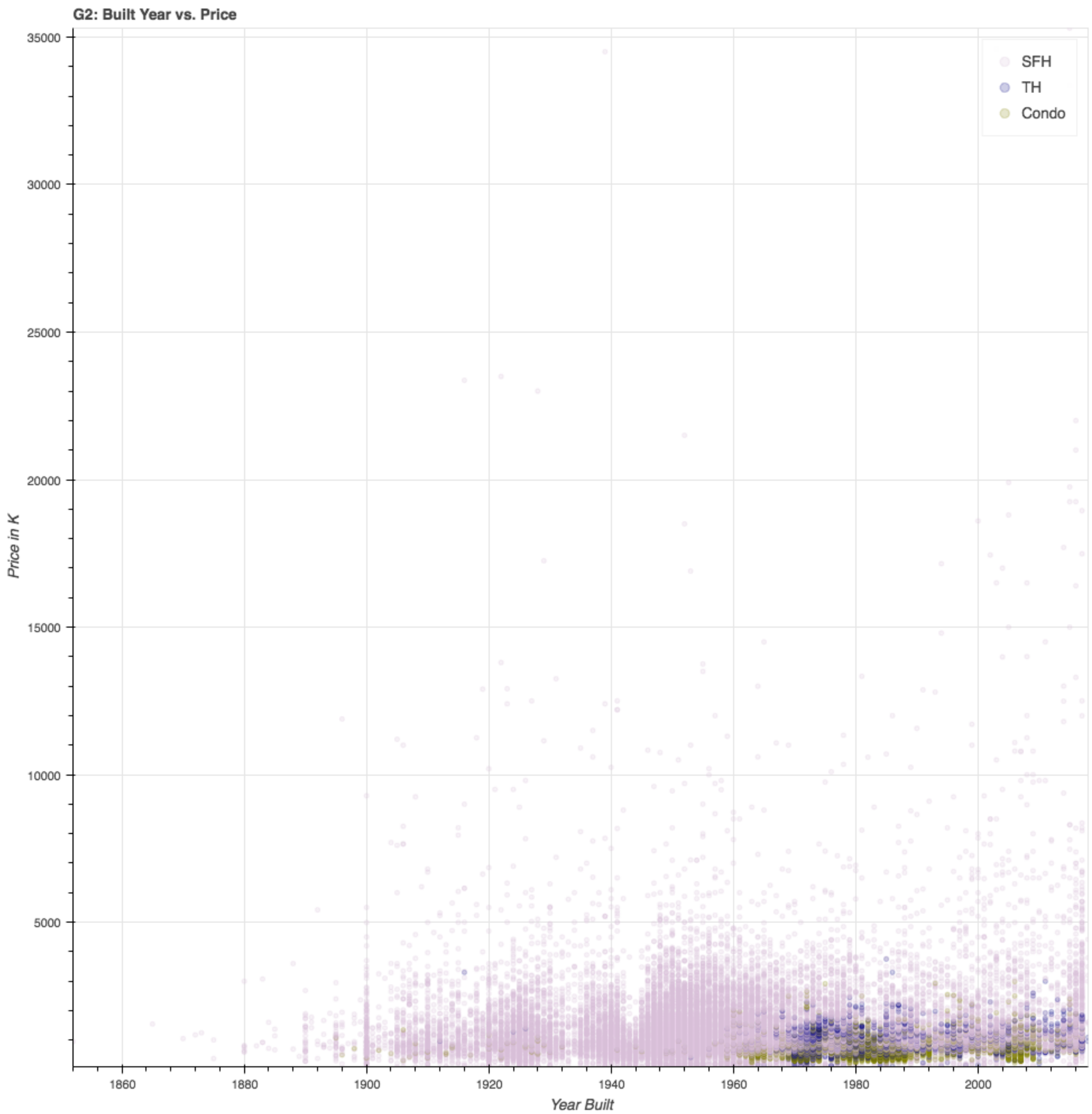


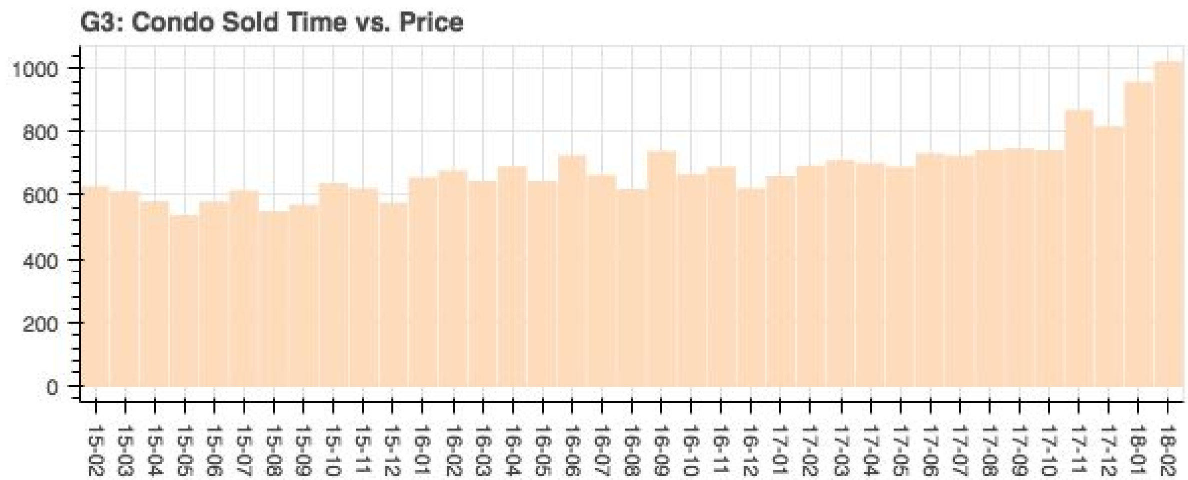
In the second map, tab 1 is the heat map of average condo price based on city. The high price is between San Francisco to Saratoga. Tab 2 is the heat map of average single family house price based on city. The high price is between Burlingame to Saratoga. Tab 3 is the heat map of average townhouse price based on city. The high price distribution is similar to that of condo. San Francisco is a unique area with high average price on townhouse and condos compared to other cities but low single family house price.

4.3. House Price by Build/Sold Year

https://github.com/Yishi1215/House-Price-In-San-Francisco-Bay-Area/blob/master/Data%20Visulization/Year%20%20vs.%20Price.ipynb


G1: Built Year vs. Price
Average Price in K
Year Built
2500
2000
1500
1000
500
1860
1880
1900
1920
1940
1960
1980
2000


G2: Built Year vs. Price
35000
30000
25000
20000
15000
10000
5000
Price in K
SFH
TH
Condo
1860
1880
1900
1920
1940
1960
1980
2000
Year Built



From G1, the average price before 1920 is very steep since there are just few properties built before 1920 in our dataset. From G2, we can see that most properties sold at the price under than 4 million. The properties built before 1900 are very rare in the market;

Townhouses and condos were not popular before 1960; Most single family houses sold in the latest three years were built between 1920 to 1960. The price of townhouse and condo is concentrate in the section below 2 million. Most high price properties were built in recent several years. From G3, for single family house, the average price of each month didn't change a lot, always around 1500K. The down time is every winter. For condo, the trend is that the average price of each month is increasing. For townhouse, the average price of each month is increasing, but the increasing rate is not as high as that of condo.

4.4. House Price by County and Property Type

https://github.com/Yishi1215/House-Price-In-San-Francisco-Bay-Area/blob/master/Data%20Visulization/Price%20by%20County%20%26%20Property%20Type.ipynb

Through the bar chart, we can find the following trends. No matter what property type it is, the average price in Alameda County is always the lowest. San Mateo County and Santa Clara County share a lot of similarities in house price. The average price of single family house is much higher than townhouse and condo; The difference of average price between townhouse and condo is bigger than that of other two counties; The difference of average price of each property type is smaller compared to other counties, which means people in San Francisco city doesn't have a strong preference on single family house as residents in other counties do.

5. Linear Regression

https://github.com/Yishi1215/House-Price-In-San-Francisco-Bay-Area/blob/master/Machine%20Learning/ML.ipynb

5.1. Building the Prediction Model – Choosing the Attributes

It is time to go beyond the exploratory analysis and build a model for prediction. We need to look at all the available attributes in the dataset and pick the variables that are potential candidates for the model.

After carefully checking the attributes, we pick the following:

PROPERTY TYPE	categorical
CITY	categorical
ZIP	categorical
PRICE IN K	numerical
BEDS	numerical
BATHS	numerical
SQUARE FEET	numerical
LOT SIZE	numerical
YEAR BUILT	numerical
HOA/MONTH	numerical
COUNTY	categorical
LOT	numerical
HOA	numerical

SOLD DATE was removed in regression analysis because it's a date object; ADDRESS was removed since address of every property is identical; LATITUDE and LONGITUDE are also removable due to the same reason; $/SQUARE FEET is the value equals to PRICE IN K divided by SQUARE FEET.

There are some townhouses and condos in our dataset without LOT SIZE or marked with a

LOT SIZE shared by all the properties in the small community. For townhouse, two different solutions to deal with the missing/wrong value problem have been used in the project. A threshold has been set for townhouse. If LOT SIZE is smaller than 3500 sqft, the original LOT SIZE will remain the same; If LOT SIZE is greater than 3500 sqft, it will be replaced by the mean LOT SIZE of all the townhouses with LOT SIZE below 3500 sqft. This approach is prepared for the prediction model which contains all the properties with different property types. When training the model for just townhouse, LOT SIZE is not one of the features. For condos, LOT SIZE is set to 0.

Since most of the single family houses have no HOA/MONTH, the missing value is set to 0. Missing HOA/MONTH of other kind of properties is set to the mean HOA/MONTH of the same type of properties. Missing BEDS, BATHS, YEAR BUILT are all set the the mean of the same type of properties.

LOT and HOA are two columns created based on whether there is missing values in LOT SIZE or HOA/MONTH. If exist, the value is 1, otherwise, is 0.

PROPERTY TYPE, CITY, ZIP AND COUNTY will be utilized as dummy variables. Dummy variables are independent variables which take the value of either 0 or 1.

5.2. Perform the Regression

In this section, I tried three different approaches to train the linear regression model. First approach is to train three different models for each property type (single family house, townhouse and condo); The second one is to train three different models based on three different types of location (zip, city, county); The last one is a general model contain different types of property and all types of location.

5.2.1. Based on different property type, only zip is included for location

		SFH	TH	Condo
LinearRegression()	RMSE	615.9830	163.7677	143.7450
	R Squared	0.7042	0.8461	0.8499
RidgeCV()	RMSE	**615.1894**	**163.2328**	**143.7823**
	R Squared	**0.7042**	**0.8460**	**0.8496**
	Alpha	**0.0050**	**0.0116**	**0.0202**
LassoCV()	RMSE	615.5369	163.4276	144.0289
	R Squared	0.7041	0.8457	0.8486
	Alpha	0.0066	0.0202	0.0202

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2}$$

sqrt((615.1894 * 615.1894 * 26792 + 163.2328 * 163.2328 * 2203 + 143.7823 * 143.7823 * 3314) / (26792 + 2203 + 3314)) = 563.7106

5.2.2. Use different location info but keep three property types together

		County	City	Zip
LinearRegression()	RMSE	768.6123	642.7695	318280779212354.1
	R Squared	0.5849	0.7190	0.7147
RidgeCV()	RMSE	768.3835	642.5487	94969.2994
	R Squared	0.5851	0.7191	0.7148
	Alpha	0.0050	0.0050	0.0050
LassoCV()	RMSE	768.3360	**642.4308**	644.7071
	R Squared	0.5850	**0.7190**	0.7146
	Alpha	0.0153	**0.0087**	0.0087

5.2.3. Keep all location and property type info

LinearRegression()	RMSE	3505353926924246.0
	R Squared	0.7297
RidgeCV()	RMSE	102539.5530
	R Squared	0.7305
	Alpha	0.005
LassoCV()	RMSE	**633.8348**
	R Squared	**0.7301**
	Alpha	**0.0116**

6. Conclusion

6.1. Analysis

In regression session, there are three sets of models and each set includes three models: simple linear regression, ridge and lasso. In each set, we choose the best fitting model, and then compared the three best fitting models. The best solution is to train separate models for each type of property and choose ridge regression for single family house and townhouse, condo. This approach makes root mean square error the smallest.

There are huge differences among single family house, townhouse and condo. A general model which includes these three types of property would sacrifice accuracy to fit all the three property types. According to the results above, it's better to seperate them and customize a model and alpha level for the specific type of property.

For the single family house, coefficients of living space, lot size, number of bathroom are positive which mean they are influence factors of the house price, especially the size of living space. Location is also a key factor. 94027, 94301, 94022, 94025, 94024, 94010, 94306, 96062, 94028, 95070 are the top 10 zipcode with highest coefficients.

For townhouse, coefficients of size of living space, number of bathroom, number of bedrooms, HOA and year of build are all positive. However, size of living space plays most significant role among these factors. 94025, 95014, 94022, 94043, 94404, 94301, 94306, 94040, 95070, 94133 are the top 10 zipcode with highest coefficients.

For condos, coefficients of size of living space, number of bathroom, number of bedrooms, HOA and year of build are all positive. Size of living space plays most significant role among these factors. 95148, 94115, 94022, 94301, 94306, 94025, 95014, 94402, 94040, 94043 are the top 10 zipcode with highest coefficients. However, zip 95148 has a extremely high coefficient score due to small data point in the area.

6.2. Limitations

Some cities with few sold property in the dataset may have less accurate prediction.
If the dataset contains neighborhood name and remodel year, it would help us improve the accuracy.

7. Attachment

7.1. Single family house price coefficients table using ridge regression

	Coefficients	Feature
2	856.2477336	SQUARE FEET
3	62.66154996	LOT SIZE
1	9.08007462	BATHS
4	-89.9229582	YEAR BUILT
0	-119.9324139	BEDS
17	316.6461986	ZIP_94027
58	177.7829346	ZIP_94301
14	167.073072	ZIP_94022
16	134.1066853	ZIP_94025
15	121.3539101	ZIP_94024
7	119.4757922	ZIP_94010
62	117.6555969	ZIP_94306
28	117.6466086	ZIP_94062
18	84.91410866	ZIP_94028
130	69.9407196	ZIP_95070
118	66.81746425	ZIP_95014
37	65.79016935	ZIP_94087
22	54.11904273	ZIP_94040
52	52.92117483	ZIP_94123
32	48.08236645	ZIP_94070
59	40.80975618	ZIP_94303
24	36.07537743	ZIP_94043
64	34.1588393	ZIP_94402
121	33.51512756	ZIP_95030

49	32.00292596	ZIP_94118
27	31.05007743	ZIP_94061
55	29.87274174	ZIP_94131
5	29.53582771	ZIP_94002
128	27.49881791	ZIP_95051
147	27.18497493	ZIP_95129
45	25.80629647	ZIP_94114
23	24.85011895	ZIP_94041
36	22.68757954	ZIP_94086
60	22.66481212	ZIP_94304
122	22.08248188	ZIP_95032
38	21.77639739	ZIP_94089
65	20.54162795	ZIP_94403
66	20.03822398	ZIP_94404
42	17.1524629	ZIP_94109
19	14.60694327	ZIP_94030
148	13.24786781	ZIP_95130
30	12.41316522	ZIP_94065
48	12.33714043	ZIP_94117
35	12.16569345	ZIP_94085
43	10.954452	ZIP_94110
120	10.08578296	ZIP_95025
54	9.290192772	ZIP_94127
46	8.188379247	ZIP_94115
142	7.67987273	ZIP_95124

41	5.932564154	ZIP_94107
117	4.956715413	ZIP_95008
149	4.361844991	ZIP_95131
136	3.042927592	ZIP_95118
47	1.300750728	ZIP_94116
39	1.269903201	ZIP_94102
50	-0.410394194	ZIP_94121
127	-0.542428007	ZIP_95050
143	-0.609397898	ZIP_95125
138	-0.66717174	ZIP_95120
135	-0.758076544	ZIP_95117
56	-0.802270711	ZIP_94132
129	-0.981031901	ZIP_95054
40	-1.069521591	ZIP_94103
116	-1.080509477	ZIP_95002
155	-1.955999354	ZIP_95139
63	-2.046749016	ZIP_94401
157	-2.111297027	ZIP_95391
51	-2.317040709	ZIP_94122
113	-2.739627088	ZIP_94710
137	-2.891256039	ZIP_95119
115	-3.036791466	ZIP_94980
114	-3.082698016	ZIP_94979
141	-3.283263735	ZIP_95123
150	-3.322899751	ZIP_95132

61	-3.717188397	ZIP_94305
154	-4.07758857	ZIP_95138
29	-4.09267325	ZIP_94063
153	-4.784046236	ZIP_95136
152	-4.895326281	ZIP_95135
88	-5.488062645	ZIP_94586
101	-6.513280845	ZIP_94612
107	-6.923573429	ZIP_94704
26	-7.00442054	ZIP_94060
33	-7.192062469	ZIP_94074
13	-7.300918352	ZIP_94021
146	-7.556471014	ZIP_95128
102	-7.598317407	ZIP_94618
57	-7.816271476	ZIP_94134
132	-7.925391295	ZIP_95111
71	-7.960718803	ZIP_94539
139	-8.359160394	ZIP_95121
144	-8.628226631	ZIP_95126
151	-10.13571859	ZIP_95133
156	-10.22146968	ZIP_95148
98	-12.09226135	ZIP_94609
131	-13.1575892	ZIP_95110
140	-13.38269964	ZIP_95122
126	-14.01914445	ZIP_95046
21	-14.2298048	ZIP_94038

70	-14.56286863	ZIP_94538
105	-14.83370063	ZIP_94702
112	-14.94025466	ZIP_94709
44	-15.22520763	ZIP_94112
99	-15.31648422	ZIP_94610
53	-15.99893887	ZIP_94124
109	-17.10600375	ZIP_94706
134	-17.71761213	ZIP_95116
20	-17.90904559	ZIP_94037
31	-17.96042978	ZIP_94066
87	-18.00321354	ZIP_94580
96	-18.10050003	ZIP_94607
79	-20.0485276	ZIP_94552
6	-21.02542179	ZIP_94005
110	-21.05708959	ZIP_94707
133	-21.35652162	ZIP_95112
106	-21.35886193	ZIP_94703
10	-21.64042295	ZIP_94018
68	-22.37721174	ZIP_94502
92	-22.92213071	ZIP_94602
9	-23.42264433	ZIP_94015
97	-23.59418224	ZIP_94608
80	-24.80911742	ZIP_94555
12	-25.41086002	ZIP_94020
95	-26.02621999	ZIP_94606

76	-26.28389842	ZIP_94546
108	-26.55769669	ZIP_94705
25	-26.64003301	ZIP_94044
145	-26.66994597	ZIP_95127
69	-26.84102378	ZIP_94536
100	-27.0361552	ZIP_94611
124	-27.17750707	ZIP_95035
111	-29.3883164	ZIP_94708
34	-31.09151568	ZIP_94080
103	-31.32192099	ZIP_94619
8	-31.69835064	ZIP_94014
86	-32.42216912	ZIP_94579
123	-33.16865265	ZIP_95033
104	-34.06880147	ZIP_94621
91	-34.29065595	ZIP_94601
75	-40.61385538	ZIP_94545
93	-41.28789735	ZIP_94603
90	-44.73813623	ZIP_94588
85	-47.29767144	ZIP_94578
81	-50.60683584	ZIP_94560
74	-52.22005965	ZIP_94544
94	-52.3073977	ZIP_94605
11	-52.86510008	ZIP_94019
67	-58.57396068	ZIP_94501
78	-60.2956293	ZIP_94551

72	-64.96983462	ZIP_94541
89	-68.3132972	ZIP_94587
84	-73.68242814	ZIP_94577
82	-78.94442687	ZIP_94566
83	-79.43221971	ZIP_94568
73	-81.0977401	ZIP_94542
125	-81.76700529	ZIP_95037
77	-94.21266883	ZIP_94550
119	-115.6888683	ZIP_95020

7.2. Townhouse price coefficients table using ridge regression

	Coefficients	Feature
2	131.7346078	SQUARE FEET
1	54.42581661	BATHS
3	36.33409914	YEAR BUILT
0	14.60044954	BEDS
4	10.65848789	HOA/MONTH
13	118.6668022	ZIP_94025
72	71.88921432	ZIP_95014
11	63.59787928	ZIP_94022
16	62.8609593	ZIP_94043
40	56.78167779	ZIP_94404
34	54.8494159	ZIP_94301

36	51.78481708	ZIP_94306
14	45.59187262	ZIP_94040
81	43.80317632	ZIP_95070
32	38.01614476	ZIP_94133
26	35.2314631	ZIP_94087
35	34.49242063	ZIP_94303
98	33.67903539	ZIP_95129
21	32.69030657	ZIP_94065
39	31.01227132	ZIP_94403
15	26.7339319	ZIP_94041
75	26.11960745	ZIP_95032
25	24.05788292	ZIP_94086
79	23.1310474	ZIP_95051
12	21.98198743	ZIP_94024
7	20.57897385	ZIP_94010
74	19.13301628	ZIP_95030
38	17.92519085	ZIP_94402
80	17.17738003	ZIP_95054
22	16.55817441	ZIP_94070
20	12.06805646	ZIP_94063
18	11.15865821	ZIP_94061
24	11.09046722	ZIP_94085
5	11.04597292	ZIP_94002
70	10.80421287	ZIP_94709
27	7.623886841	ZIP_94089

86	5.416130092	ZIP_95117
78	5.358305049	ZIP_95050
68	4.045407036	ZIP_94703
37	3.027277488	ZIP_94401
97	2.231275431	ZIP_95128
89	1.480896391	ZIP_95120
45	1.237923304	ZIP_94539
94	1.012469268	ZIP_95125
82	0.352684954	ZIP_95110
31	0.221407979	ZIP_94131
69	-0.660616916	ZIP_94706
84	-0.879652973	ZIP_95112
29	-1.233162648	ZIP_94111
93	-1.737297804	ZIP_95124
95	-2.220834494	ZIP_95126
87	-2.745191635	ZIP_95118
8	-3.012672124	ZIP_94014
88	-3.220743636	ZIP_95119
103	-3.809950564	ZIP_95134
100	-3.929849588	ZIP_95131
102	-3.959434333	ZIP_95133
105	-5.469658563	ZIP_95138
30	-5.537004284	ZIP_94124
99	-5.537972161	ZIP_95130
59	-5.876975717	ZIP_94586

66	-6.230885658	ZIP_94618
71	-6.573483585	ZIP_95008
96	-7.791470973	ZIP_95127
6	-8.271434419	ZIP_94005
106	-8.424376098	ZIP_95139
64	-8.473998061	ZIP_94606
107	-8.512104949	ZIP_95148
51	-8.879612486	ZIP_94552
85	-9.237290647	ZIP_95116
92	-9.767821627	ZIP_95123
62	-9.821877873	ZIP_94603
65	-10.14150782	ZIP_94608
90	-10.16269991	ZIP_95121
58	-10.32700057	ZIP_94579
52	-10.40605822	ZIP_94555
104	-10.43129516	ZIP_95136
91	-11.25247918	ZIP_95122
67	-11.99504275	ZIP_94619
23	-11.99786333	ZIP_94080
83	-13.08666441	ZIP_95111
101	-14.01979346	ZIP_95132
10	-14.48267178	ZIP_94019
17	-16.70394217	ZIP_94044
9	-19.22164693	ZIP_94015
41	-19.35930906	ZIP_94501

44	-20.88092786	ZIP_94538
43	-21.23281833	ZIP_94536
33	-22.57191442	ZIP_94134
57	-22.80677195	ZIP_94578
49	-23.77124774	ZIP_94550
54	-24.63595313	ZIP_94566
55	-24.75427404	ZIP_94568
48	-28.8089844	ZIP_94545
76	-29.66169618	ZIP_95035
47	-31.79466668	ZIP_94544
56	-33.40107848	ZIP_94577
73	-33.62566072	ZIP_95020
63	-36.15725424	ZIP_94605
61	-42.15854966	ZIP_94588
42	-45.1288362	ZIP_94502
53	-46.17644945	ZIP_94560
60	-52.96294759	ZIP_94587
77	-54.53038268	ZIP_95037
50	-55.55721714	ZIP_94551
46	-67.22083474	ZIP_94541
28	-44599.93198	ZIP_94102
19	-44599.93198	ZIP_94062

7.3. Condo price coefficients table using ridge regression

	Coefficients	Feature
2	144.6640955	SQUARE FEET
1	38.84442471	BATHS
3	20.52097198	YEAR BUILT
0	10.36645415	BEDS
4	6.677333993	HOA/MONTH
34	4.21914E+13	ZIP_94110
126	-9.69262E+12	ZIP_95148
36	-5.59637E+13	ZIP_94117
37	-5.59637E+13	ZIP_94121
30	-5.59637E+13	ZIP_94102
33	-5.59637E+13	ZIP_94109
29	-5.59637E+13	ZIP_94089
88	-5.59637E+13	ZIP_94708
109	-5.59637E+13	ZIP_95119
38	-5.59637E+13	ZIP_94122
79	-5.59637E+13	ZIP_94612
39	-7.91326E+13	ZIP_94131
80	-7.91326E+13	ZIP_94618
43	-7.91326E+13	ZIP_94305
87	-7.91326E+13	ZIP_94707
73	-7.91326E+13	ZIP_94606
71	-7.91326E+13	ZIP_94602
76	-7.91326E+13	ZIP_94609
72	-7.91326E+13	ZIP_94605

83	-7.91326E+13	ZIP_94703
125	-7.91326E+13	ZIP_95138
55	-7.91326E+13	ZIP_94542
32	-9.69026E+13	ZIP_94107
81	-9.69026E+13	ZIP_94619
67	-9.69026E+13	ZIP_94579
12	-1.11877E+14	ZIP_94024
74	-1.11877E+14	ZIP_94607
90	-1.11877E+14	ZIP_94710
70	-1.11877E+14	ZIP_94601
82	-1.25063E+14	ZIP_94702
40	-1.25063E+14	ZIP_94134
50	-1.25063E+14	ZIP_94502
57	-1.25063E+14	ZIP_94545
60	-1.25063E+14	ZIP_94551
89	-1.36979E+14	ZIP_94709
105	-1.36979E+14	ZIP_95113
20	-1.36979E+14	ZIP_94062
110	-1.36979E+14	ZIP_95121
21	-1.47932E+14	ZIP_94063
84	-1.47932E+14	ZIP_94704
59	-1.47932E+14	ZIP_94550
31	-1.54424E+14	ZIP_94105
16	-1.58122E+14	ZIP_94041
85	-1.58122E+14	ZIP_94705

78	-1.58122E+14	ZIP_94611
122	-1.58122E+14	ZIP_95134
69	-1.58122E+14	ZIP_94588
63	-1.58122E+14	ZIP_94566
58	-1.58122E+14	ZIP_94546
66	-1.67688E+14	ZIP_94578
94	-1.76732E+14	ZIP_95030
77	-1.76732E+14	ZIP_94610
53	-1.76732E+14	ZIP_94539
61	-1.76732E+14	ZIP_94555
119	-1.8533E+14	ZIP_95131
5	-1.93542E+14	ZIP_94002
124	-1.93542E+14	ZIP_95136
116	-1.93542E+14	ZIP_95127
18	-2.01414E+14	ZIP_94044
56	-2.01414E+14	ZIP_94544
52	-2.16288E+14	ZIP_94538
22	-2.23347E+14	ZIP_94065
120	-2.30186E+14	ZIP_95132
107	-2.30186E+14	ZIP_95117
86	-2.36824E+14	ZIP_94706
115	-2.43276E+14	ZIP_95126
111	-2.43276E+14	ZIP_95122
123	-2.43276E+14	ZIP_95135
54	-2.43276E+14	ZIP_94541

114	-2.49558E+14	ZIP_95125
121	-2.55682E+14	ZIP_95133
64	-2.55682E+14	ZIP_94568
28	-2.73211E+14	ZIP_94087
108	-2.73211E+14	ZIP_95118
51	-2.73211E+14	ZIP_94536
41	-2.78803E+14	ZIP_94301
19	-2.78803E+14	ZIP_94061
8	-2.78803E+14	ZIP_94014
10	-2.84281E+14	ZIP_94019
106	-2.84281E+14	ZIP_95116
65	-2.84281E+14	ZIP_94577
6	-3.00097E+14	ZIP_94005
113	-3.05181E+14	ZIP_95124
118	-3.10178E+14	ZIP_95129
104	-3.10178E+14	ZIP_95112
103	-3.10178E+14	ZIP_95111
93	-3.10178E+14	ZIP_95020
117	-3.24692E+14	ZIP_95128
68	-3.29382E+14	ZIP_94587
13	-3.34004E+14	ZIP_94025
95	-3.43051E+14	ZIP_95032
46	-3.56172E+14	ZIP_94402
47	-3.56172E+14	ZIP_94403
9	-3.56172E+14	ZIP_94015

14	-3.60435E+14	ZIP_94030
44	-3.64645E+14	ZIP_94306
42	-3.68804E+14	ZIP_94303
26	-3.68804E+14	ZIP_94085
97	-3.68804E+14	ZIP_95037
100	-3.72914E+14	ZIP_95054
62	-3.72914E+14	ZIP_94560
102	-3.80995E+14	ZIP_95110
27	-4.11641E+14	ZIP_94086
75	-4.11641E+14	ZIP_94608
101	-4.22524E+14	ZIP_95070
23	-4.33114E+14	ZIP_94066
24	-4.56798E+14	ZIP_94070
92	-4.60074E+14	ZIP_95014
7	-4.6655E+14	ZIP_94010
25	-4.6655E+14	ZIP_94080
11	-4.79216E+14	ZIP_94022
98	-5.00539E+14	ZIP_95050
91	-5.48534E+14	ZIP_95008
48	-5.59157E+14	ZIP_94404
99	-5.72122E+14	ZIP_95051
49	-5.79743E+14	ZIP_94501
15	-5.92193E+14	ZIP_94040
45	-6.11507E+14	ZIP_94401
112	-6.2784E+14	ZIP_95123

17	-6.34689E+14	ZIP_94043
96	-7.00848E+14	ZIP_95035
35	-1.70195E+15	ZIP_94115