Property Price Predictions in San Francisco Bay Area

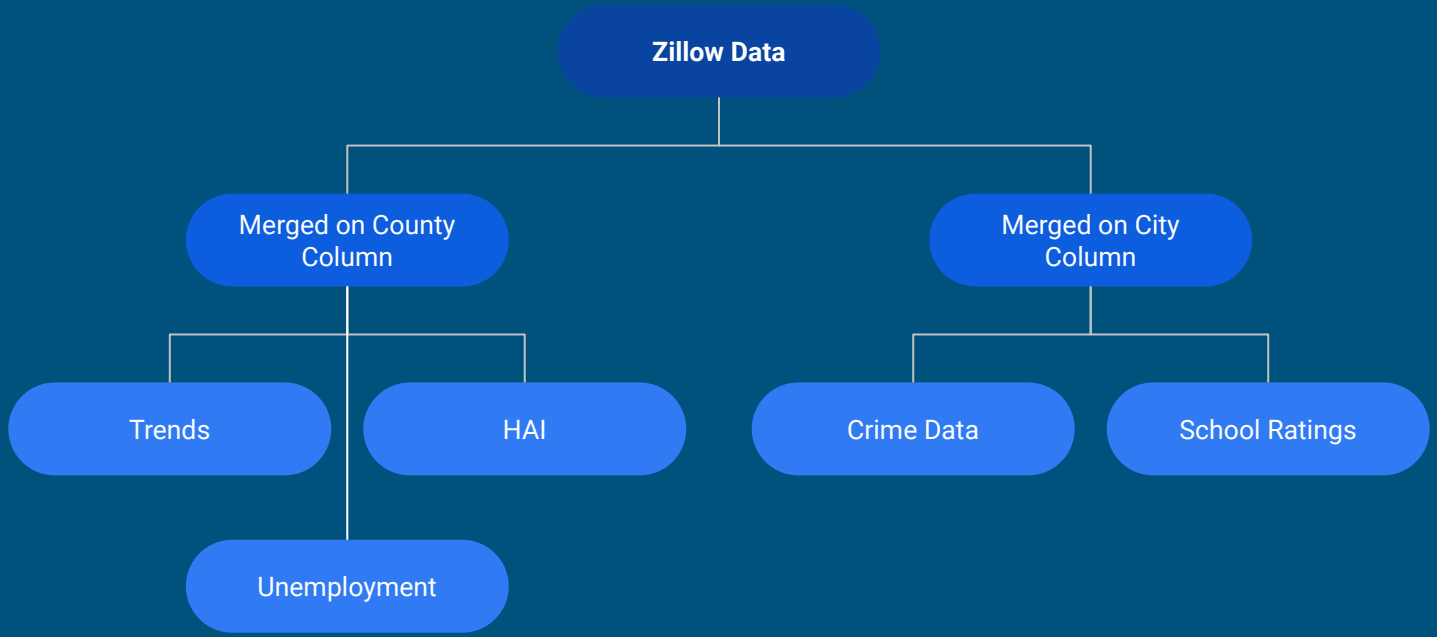
Agne Adukaite Aklifazla
Mentor: Ricardo D. Alanis-Tamez

Data

- Zillow Data using rapid api
- Crime Rates
- School Ratings
- Unemployment Rates
- Housing Affordability Index
- Trends Data (property appreciation in the area)

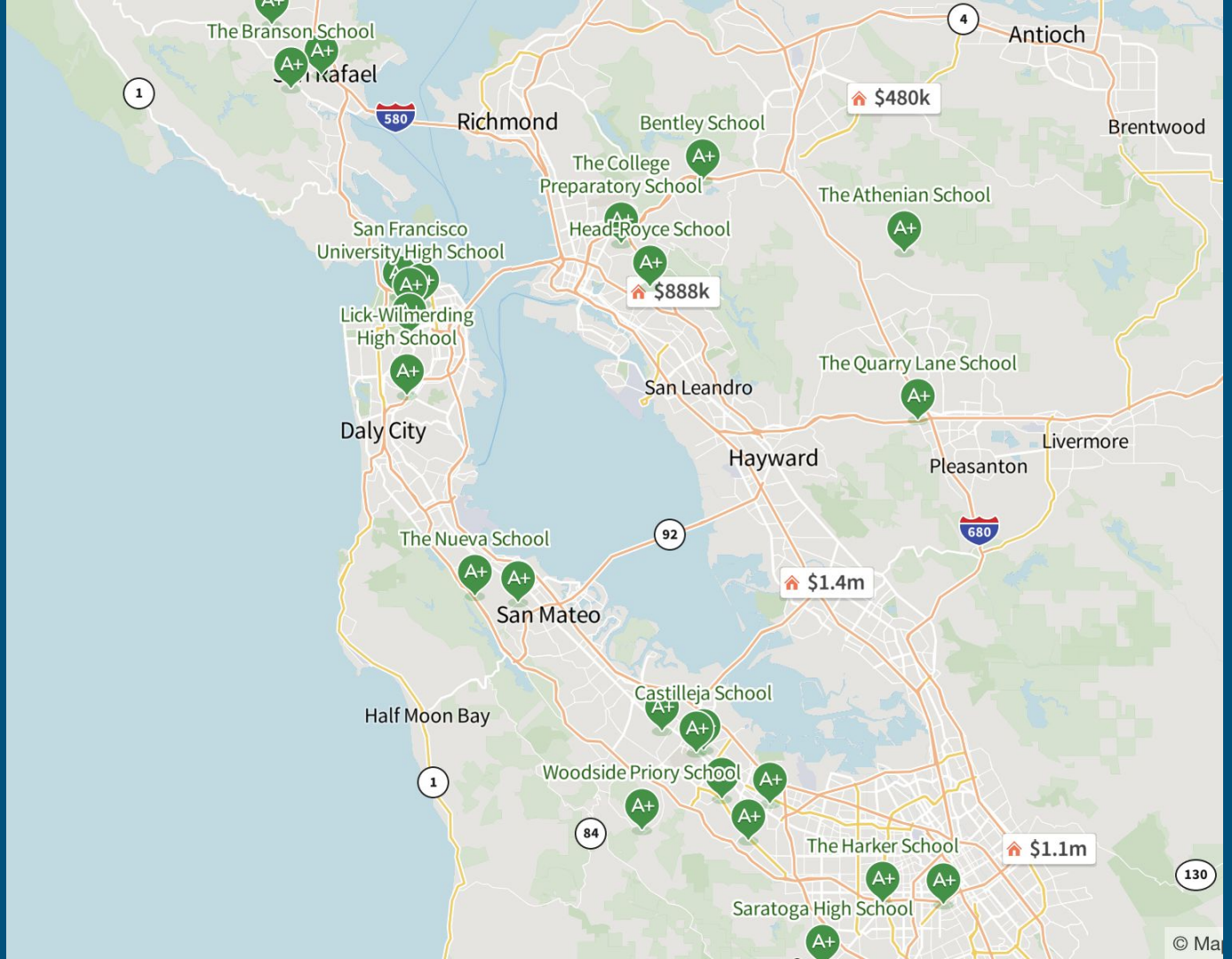
Merged data

2176 entries 22 columns



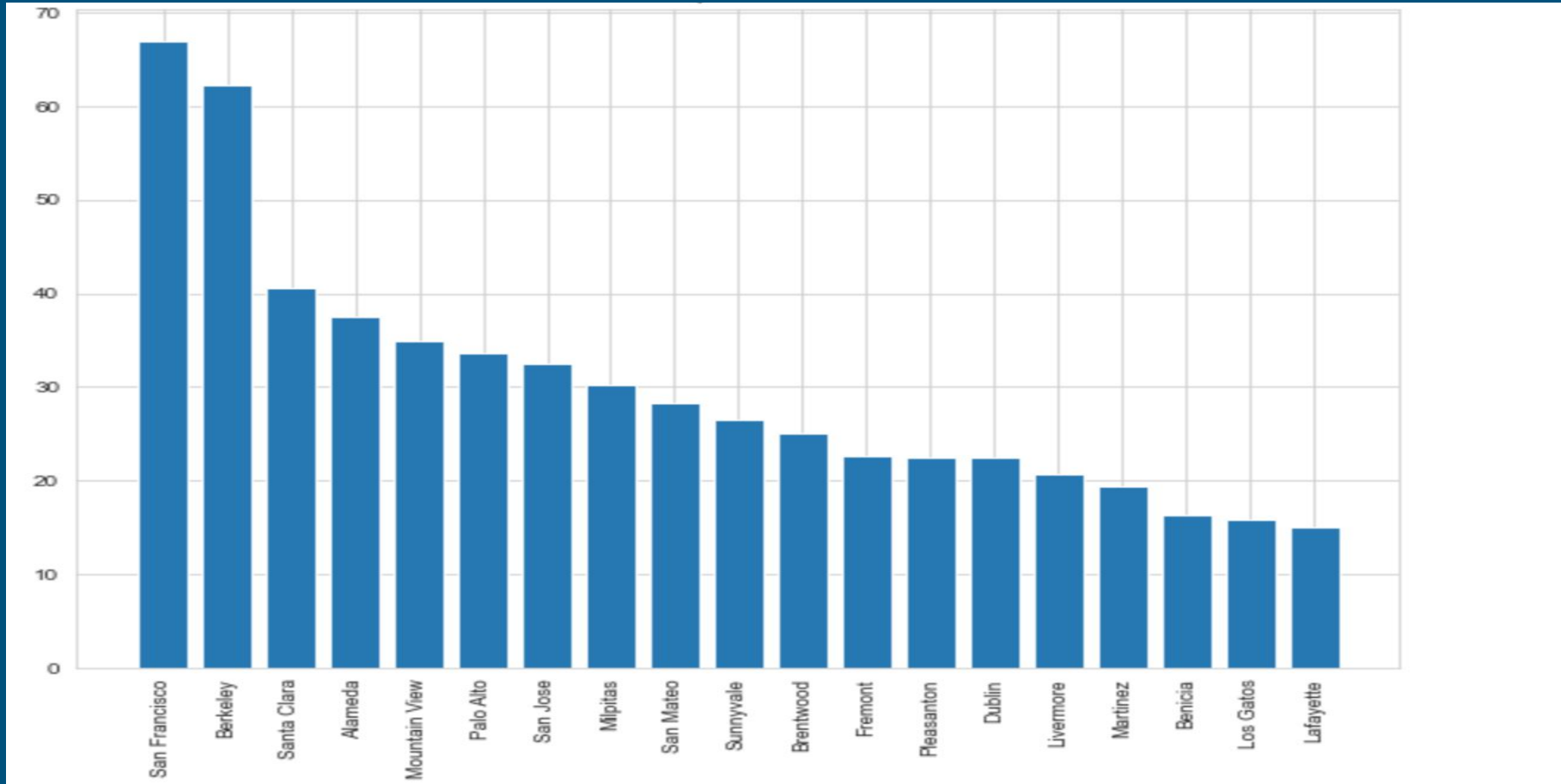
EDA process

1. EDA on outside Data
2. EDA on Zillow Data
3. Merging all data together
4. EDA on full data

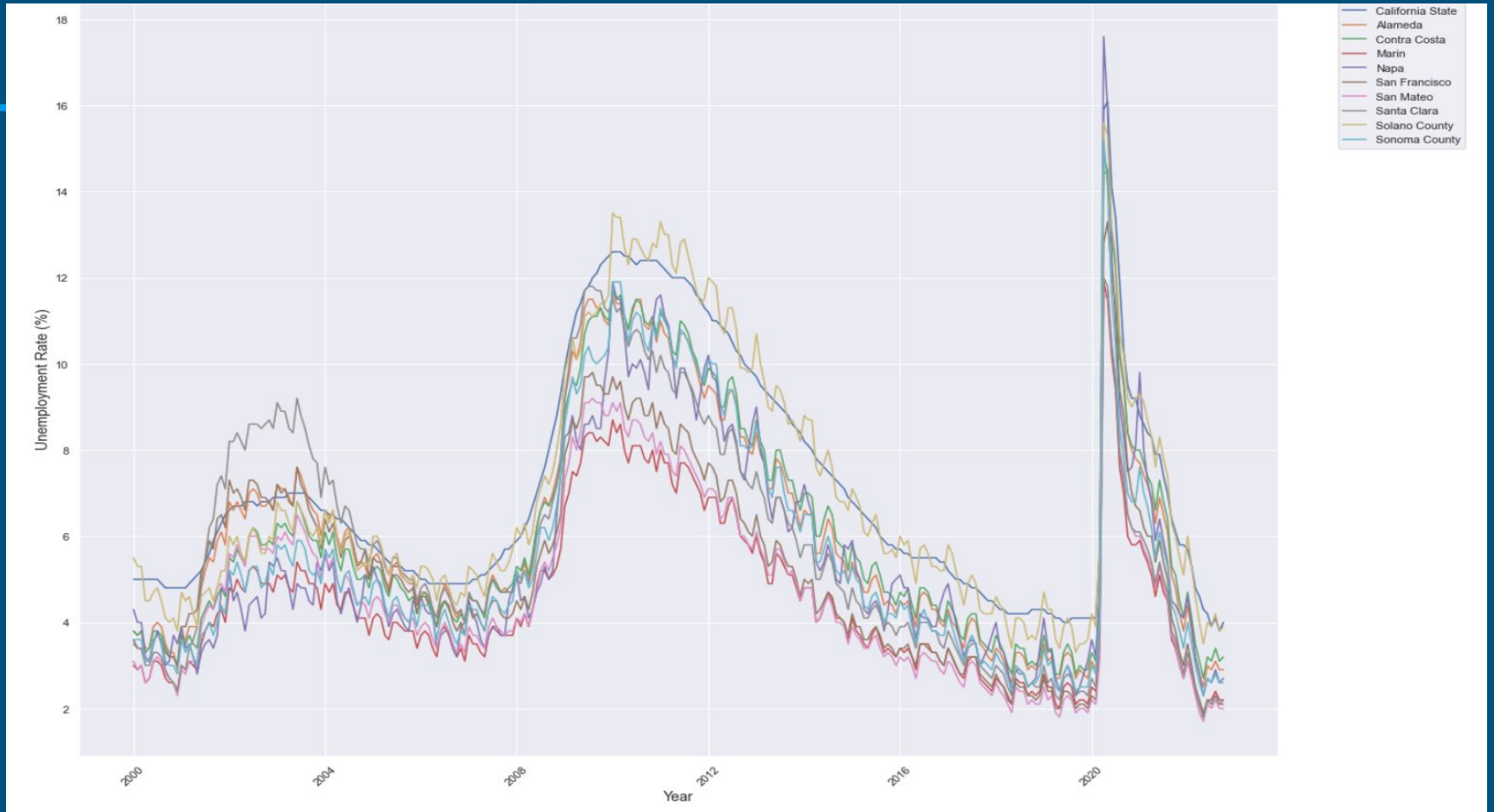


The Branson School
Richmond
Bentley School
The College Preparatory School
Head-Royce School
$888k
$480k
Antioch
Brentwood
The Athenian School
San Francisco University High School
Lick-Wilmerding High School
Daly City
San Leandro
The Quarry Lane School
Hayward
Pleasanton
Livermore
The Nueva School
San Mateo
$1.4m
Castilleja School
Half Moon Bay
Woodside Priory School
The Harker School
$1.1m
Saratoga High School

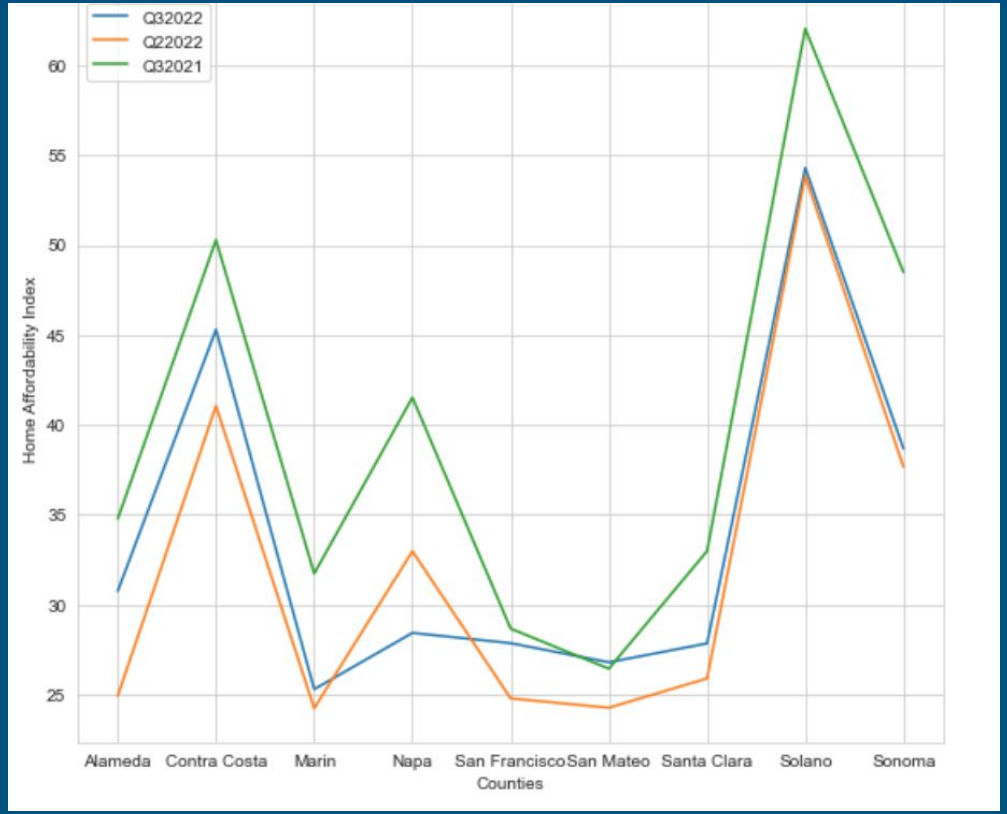
Crime Rates (total crimes per 1000 residents)



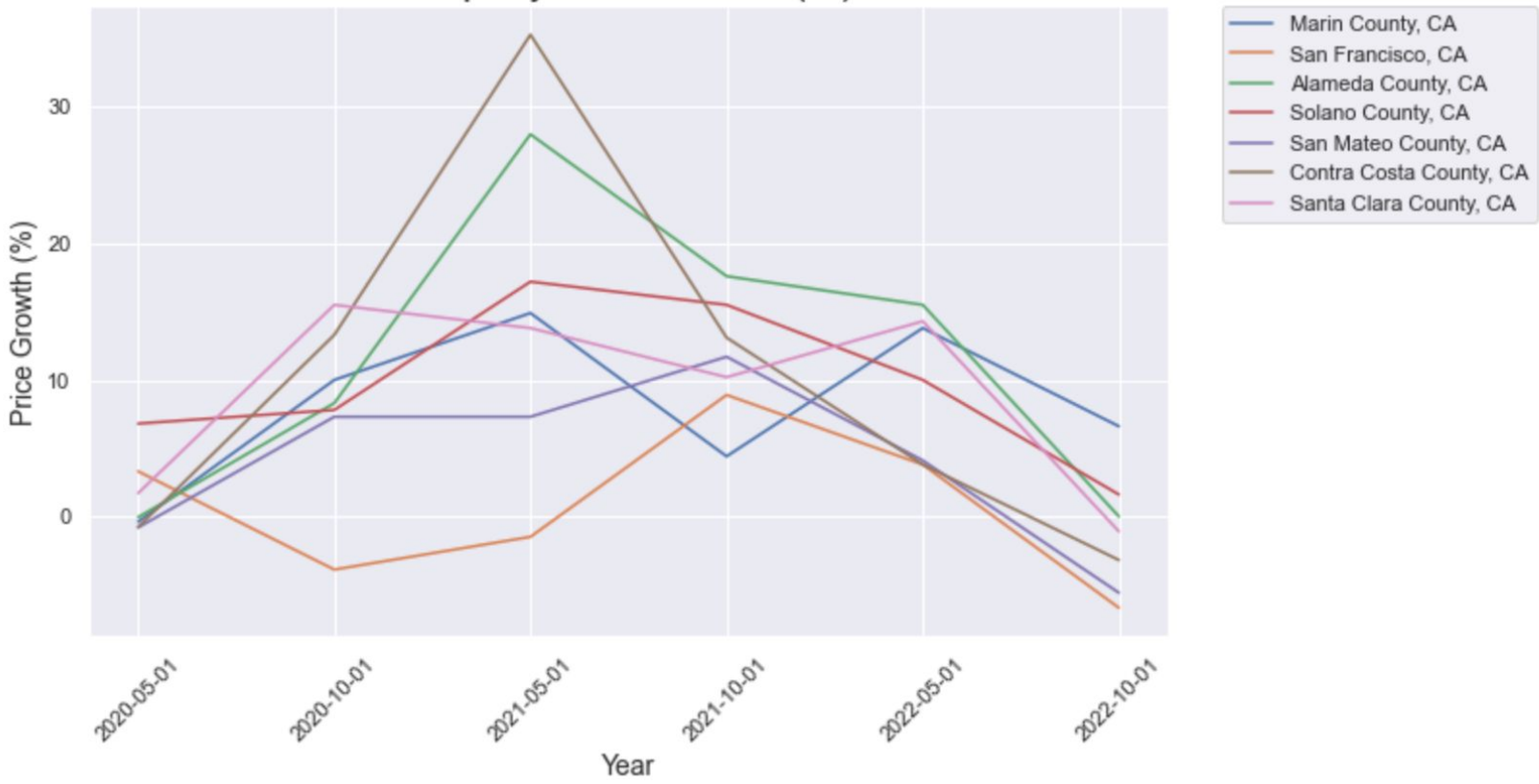
Unemployment Rates



Housing Affordability Index (HAI)



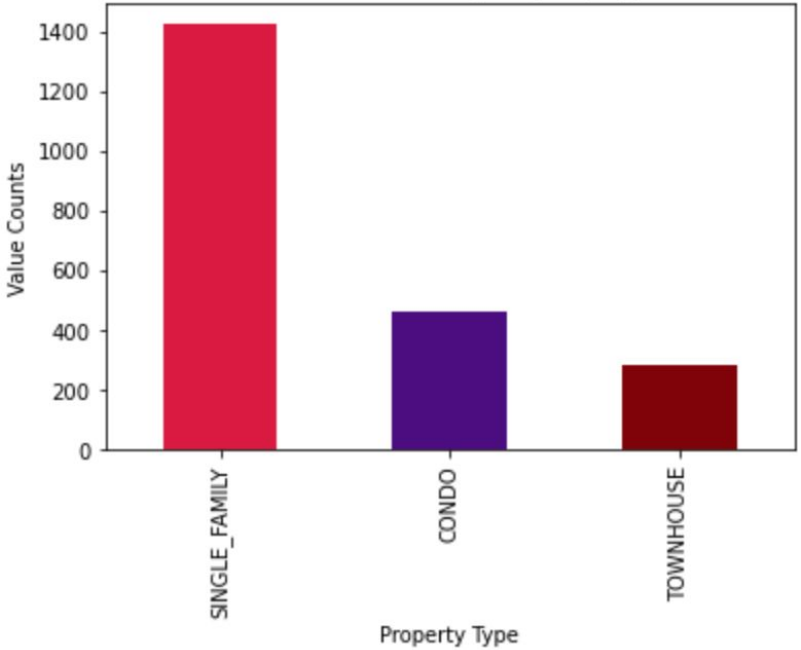
Trends Data

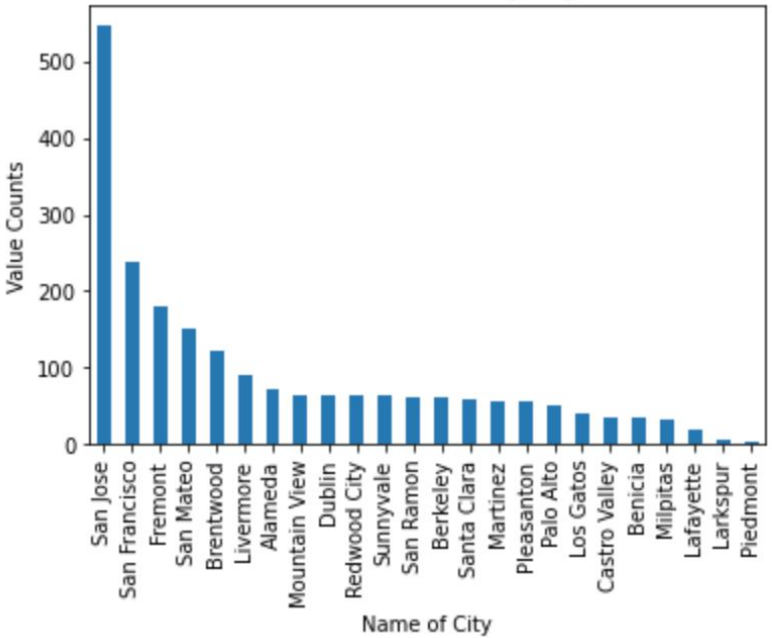

Price Growth (%)
30
20
10
0
2020-05-01
2020-10-01
2021-05-01
2021-10-01
2022-05-01
2022-10-01
Year
Marin County, CA
San Francisco, CA
Alameda County, CA
Solano County, CA
San Mateo County, CA
Contra Costa County, CA
Santa Clara County, CA

Zillow data (filters used for selection)

- Location - city
- Type of property. Selected SINGLE_FAMILY (which included condos, townhouses, and single family houses).
- Minimum number of bedrooms = 2.
- Maximum living area sqf of 3000.

Zillow Data Distribution





Correlation between numerical and categorical values

price/living area = 0.57

bathrooms/living area = 0.7

bathrooms/bedrooms = 0.56

bedrooms/living area = 0.69

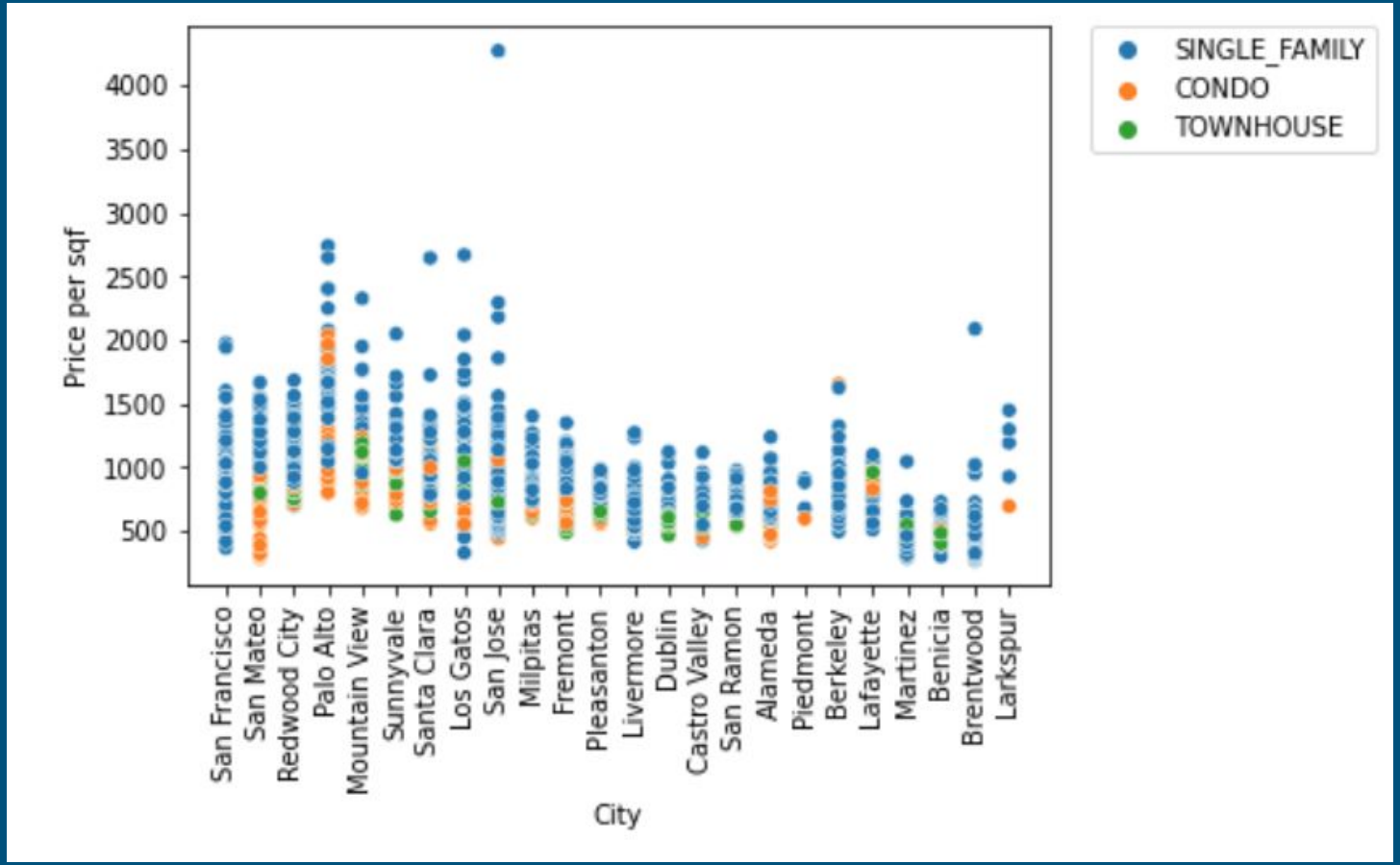
price/listing sub type is new home = 0.66

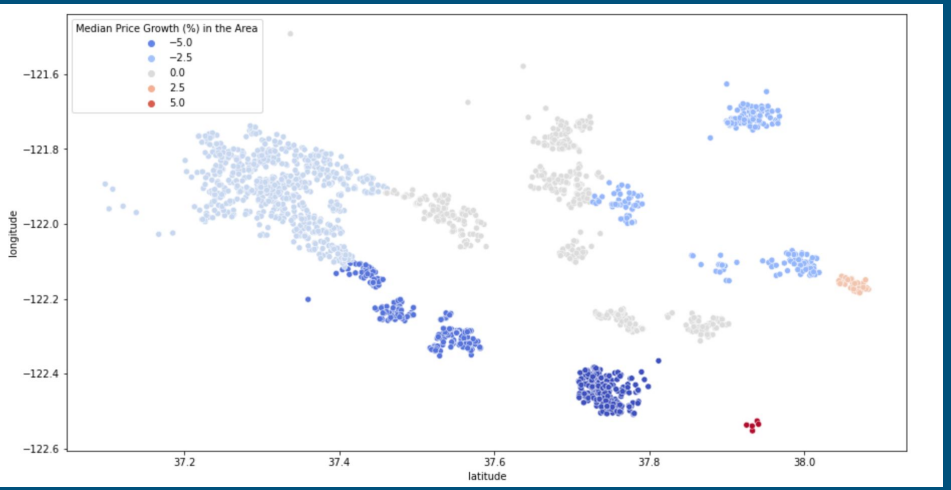
price/new construction type = 0.62

price/zpid(zillow property id)=0.68

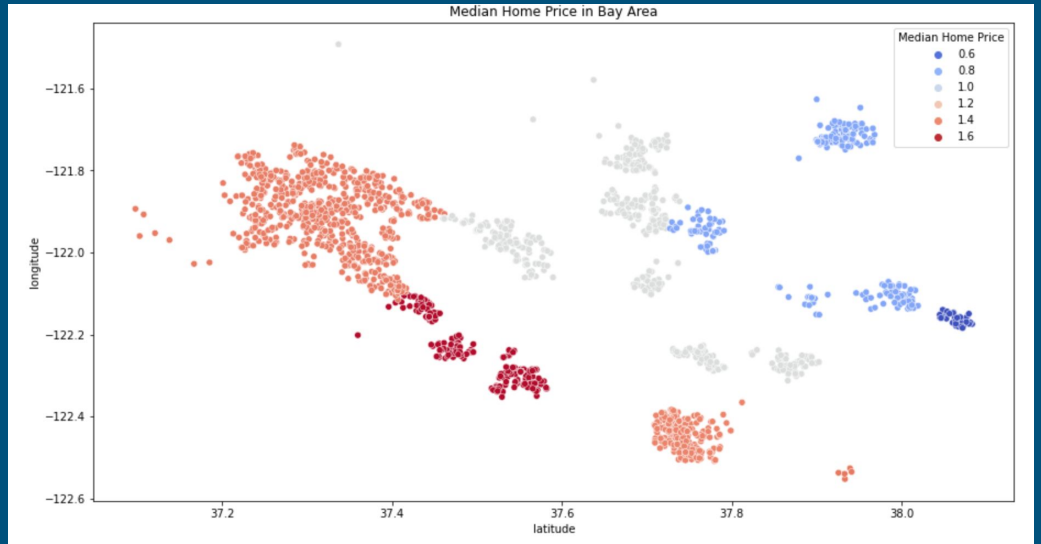
price/address = 0.68

Added price per sqf column




Median Price Growth (%) in the Area
−5.0
−2.5
0.0
2.5
5.0
longitude
latitude
−121.6
−121.8
−122.0
−122.2
−122.4
−122.6
37.2
37.4
37.6
37.8
38.0

Price per sqf ($) Distribution in Bay Area
Price per sqf ($)
800
1600
2400
3200
4000
longitude
latitude
−121.6
−121.8
−122.0
−122.2
−122.4
−122.6
37.2
37.4
37.6
37.8
38.0


Median Home Price in Bay Area
Median Home Price
0.6
0.8
1.0
1.2
1.4
1.6
longitude
latitude
−121.6
−121.8
−122.0
−122.2
−122.4
−122.6
37.2
37.4
37.6
37.8
38.0

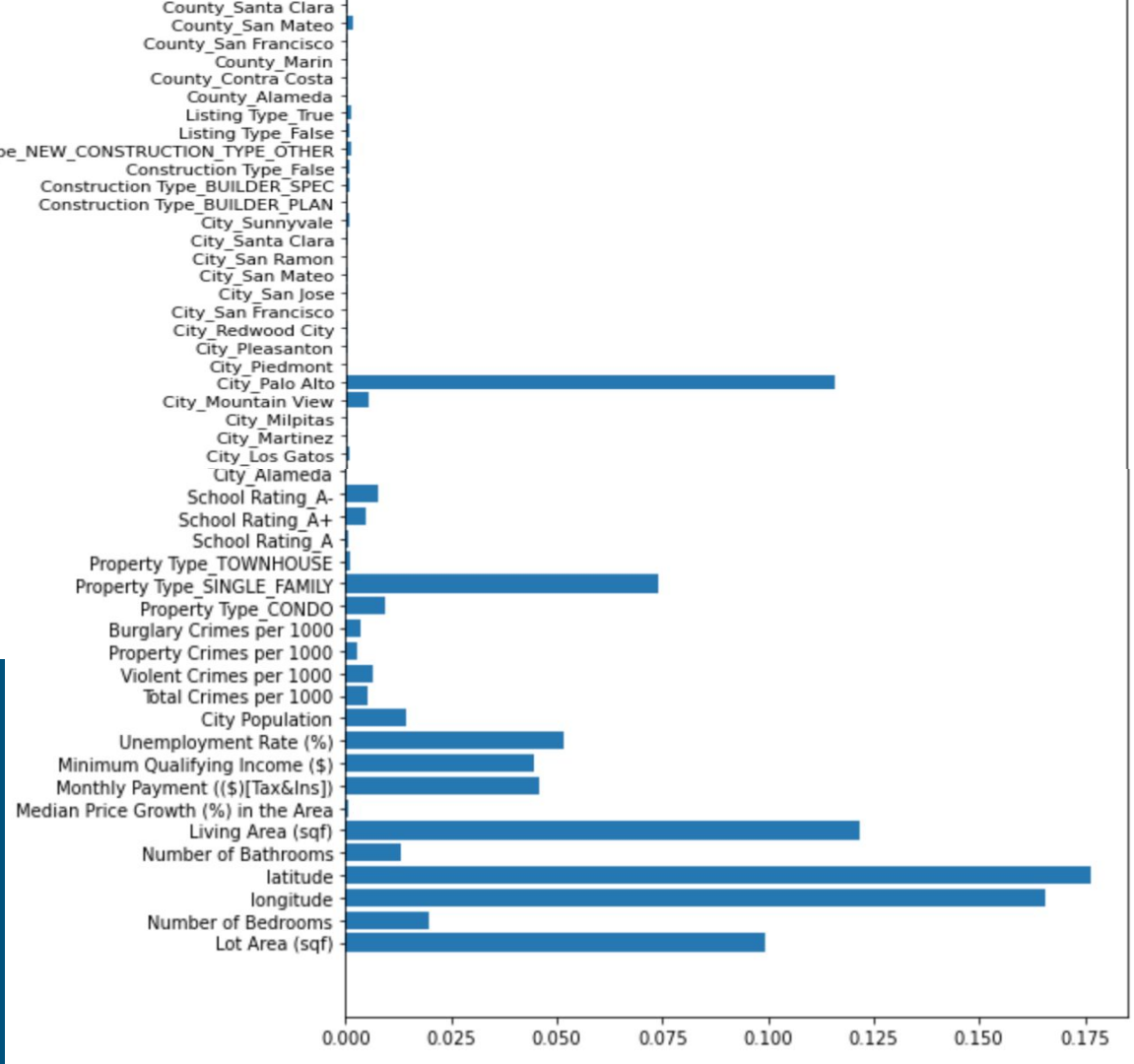
Machine Learning Algorithms

Algorithms

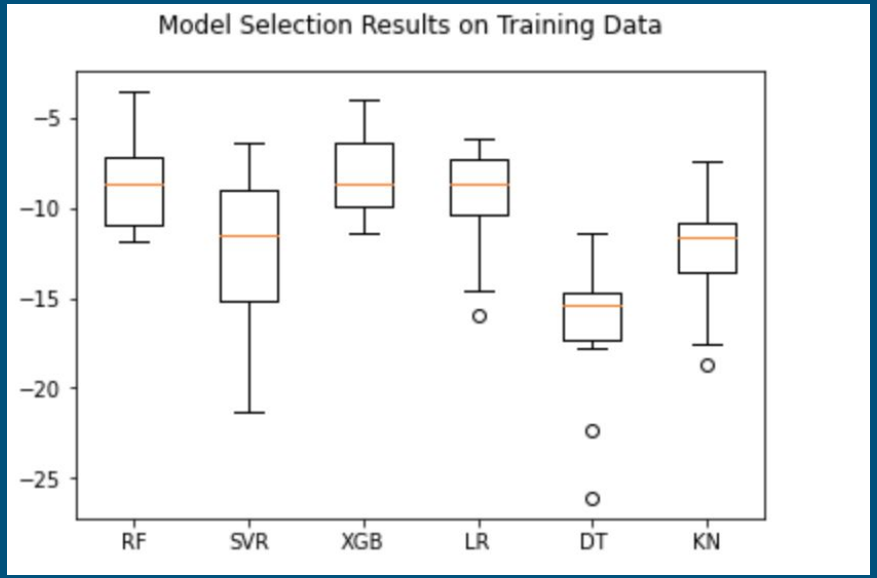
- Random Forest Regressor
- Grid Search CV
- XGBoost
- LInear Regressor
- Linear Regression using 6 folds

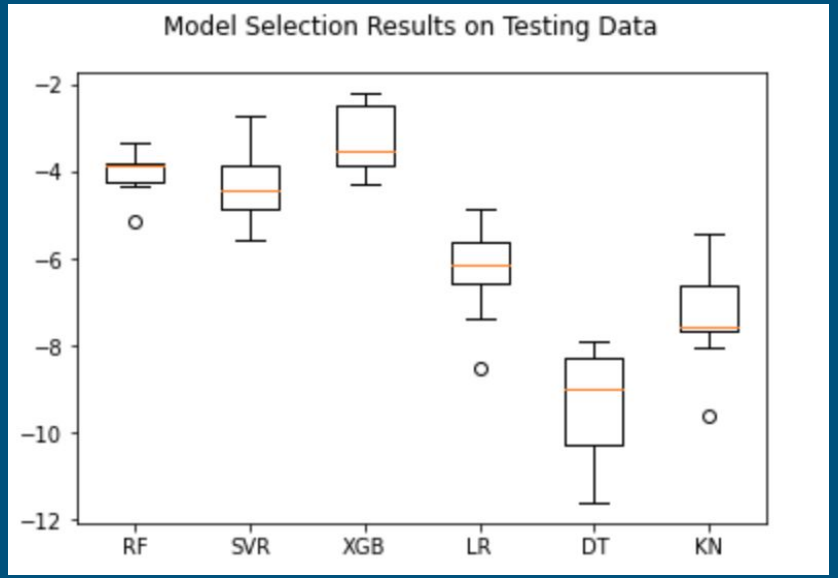
Evaluation

- R2
- Mean squared error (+/-)
- Mean absolute error
- Cross validation


County_Santa Clara
County_San Mateo
County_San Francisco
County_Marin
County_Contra Costa
County_Alameda
Listing Type_True
Listing Type_False
Construction Type_NEW_CONSTRUCTION_TYPE_OTHER
Construction Type_False
Construction Type_BUILDER_SPEC
Construction Type_BUILDER_PLAN
City_Sunnyvale
City_Santa Clara
City_San Ramon
City_San Mateo
City_San Jose
City_San Francisco
City_Redwood City
City_Pleasanton
City_Piedmont
City_Palo Alto
City_Mountain View
City_Milpitas
City_Martinez
City_Los Gatos
City_Alameda
School Rating_A-
School Rating_A+
School Rating_A
Property Type_TOWNHOUSE
Property Type_SINGLE_FAMILY
Property Type_CONDO
Burglary Crimes per 1000
Property Crimes per 1000
Violent Crimes per 1000
Total Crimes per 1000
City Population
Unemployment Rate (%)
Minimum Qualifying Income ($)
Monthly Payment (($)[Tax&Ins])
Median Price Growth (%) in the Area
Living Area (sqf)
Number of Bathrooms
latitude
longitude
Number of Bedrooms
Lot Area (sqf)
0.000
0.025
0.050
0.075
0.100
0.125
0.150
0.175

ML models evaluation





Suggestions

Predicted Price per sqf	Lot Area (sqf)	Property Price ($)	Bedrooms	Bathrooms	Living Area (sqf)	Median Price Growth (%) in the Area	Median Home Price	Monthly Payment (($)[Tax&Ins])	Minimum Qualifying Income ($)	Total Crimes per 1000	Property Type	School Rating	City	County
$549	2.2	1040000	3	3	2131	-3.2	753310	4320	129600	25.09	Single Family	A-	Brentwood	Contra Costa
$924	8507.268	2850000	4	3	2860	-5.6	1641350	9420	282600	28.29	Single Family	A+	San Mateo	San Mateo
$518	3049.2	825000	3	3	1515	-3.2	753310	4320	129600	19.44	Townhouse	A-	Martinez	Contra Costa
$758	2996.928	1850000	4	3	2508	-6.7	1411000	8100	243000	66.99	Single Family	A-	San Francisco	San Francisco
$548	4791.6	999000	4	2	1804	0	1071000	6150	184500	62.26	Single Family	A+	Berkeley	Alameda
$1,419	5928.516	2398000	3	2	1610	-5.6	1641350	9420	282600	28.29	Single Family	A+	San Mateo	San Mateo
$943	2613.6	949000	2	2	1035	0	1071000	6150	184500	37.5	Single Family	A	Alameda	Alameda
$419	566.28	1025000	5	3	2522	-3.2	753310	4320	129600	19.44	Single Family	A-	Martinez	Contra Costa
$1,049	5662.8	2800000	4	3	2623	0	1071000	6150	184500	37.5	Single Family	A	Alameda	Alameda
$750	0	759850	2	1	947	-1.1	1434800	8230	246900	24.46	Condo	A+	Sunnyvale	Santa Clara