House Price Prediction

Cindy Su
10/1/2021

Motivation

$18,000,000+

2020

12,000,000

2017

Median House Sales Prices
Bay Area

Primary Data Source

CALIFORNIA'S
SAN FRANCISCO
BAY AREA

Sonoma
Napa
Solano
Marin
Contra Costa
San Francisco
Alameda
San Mateo
Santa Clara

- Alameda County
- Santa Clara County

Bay Area House Data

3900+ entries of Data

Median Household Income in California

Modeling Framework

Predict the House value based on the pre-house sale prices

Key Features

- **Home Type**
- **Bedroom**
- **Bathrooms**
- **Home Size**
- **Parking Lot Num**
- **House Age**
- **House Hold Avg Income**
- **City**

Modeling Framework

REMOVE EXTREME VALUES

Sold Price
- ➢ > 5K
- ➢ < 6M

BEDROOM NUM
- ➢ =0
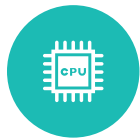
YEAR BUILT
- ➢ =0

Home Size
- ➢ <5000

Home Size

<5000 sqft

Sold Price

<6M

Sqrt Unit price

Key Feature
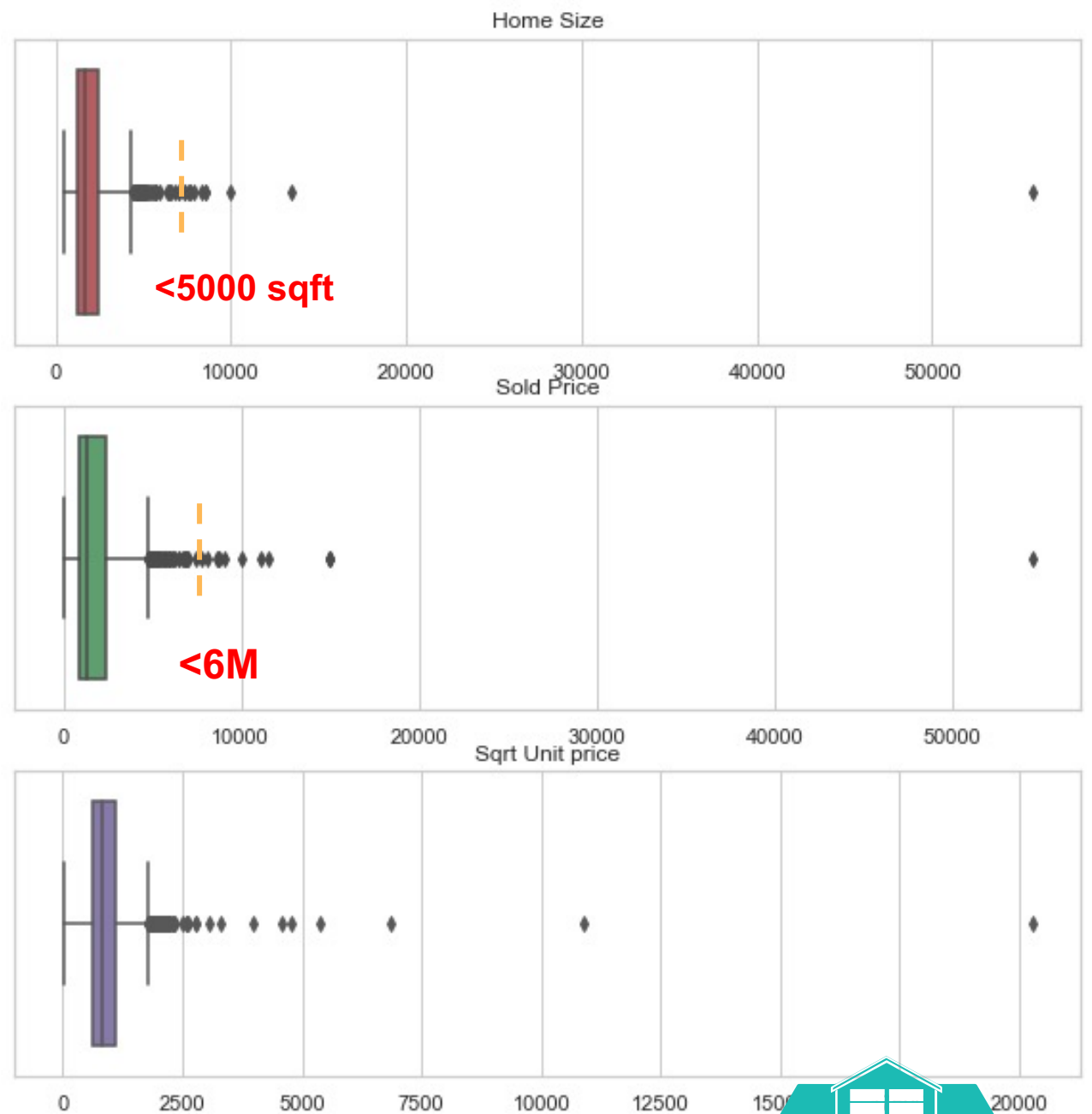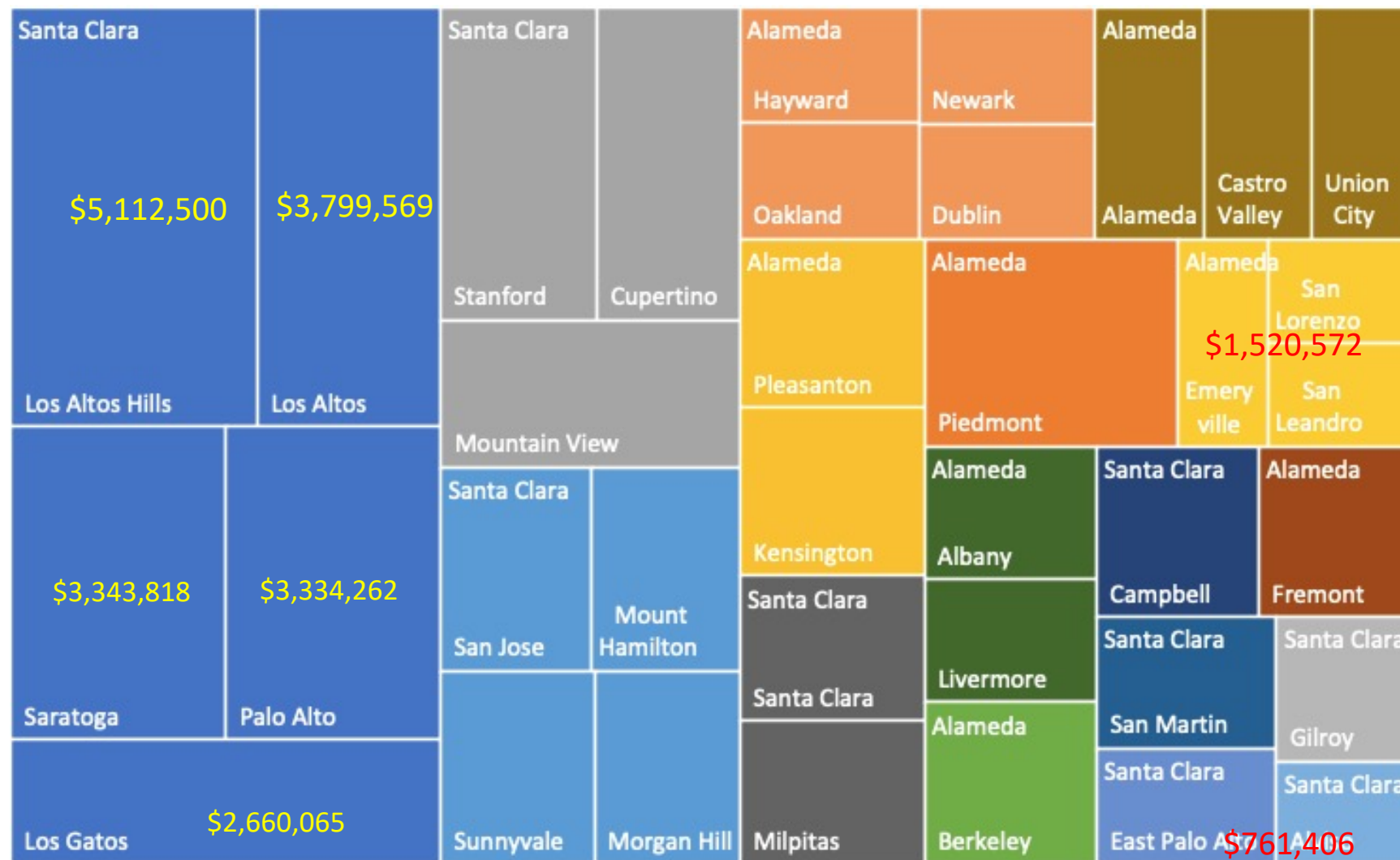
Average Sale Price by City
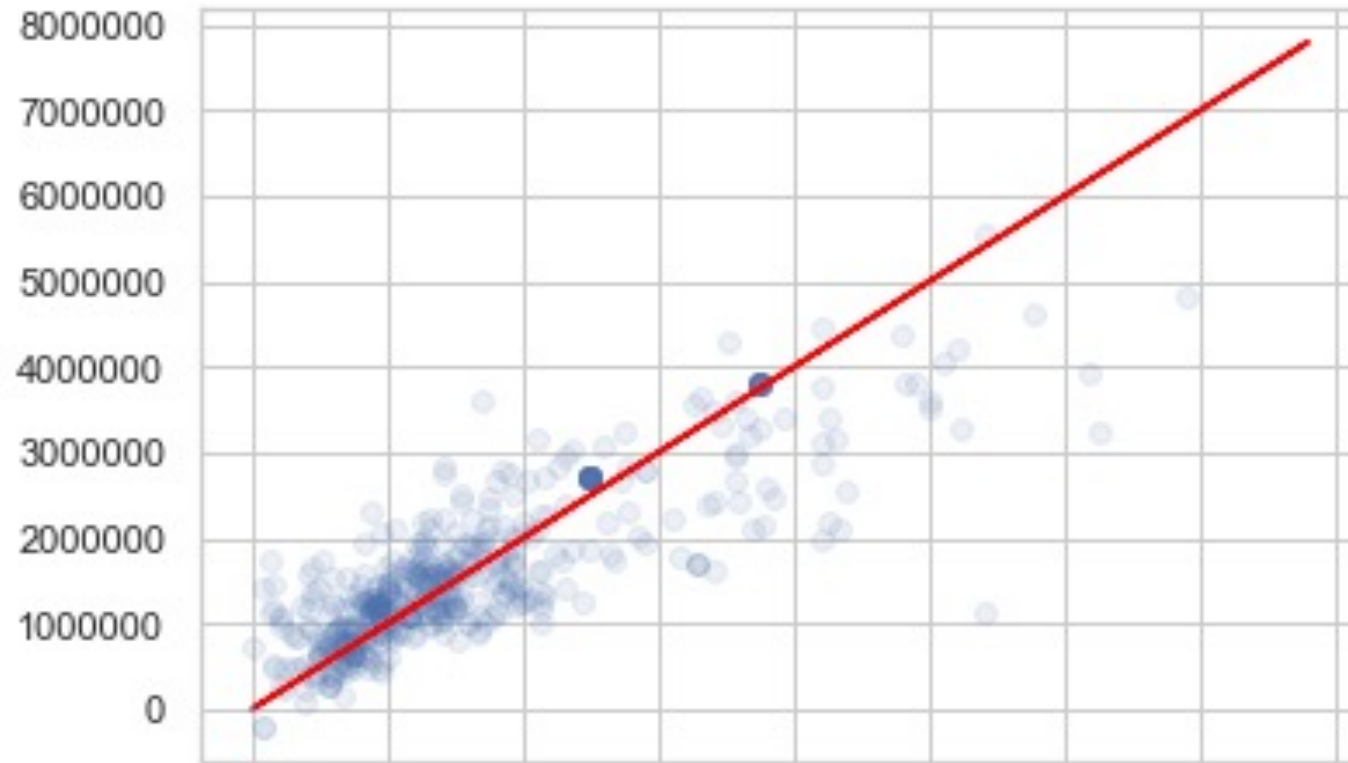
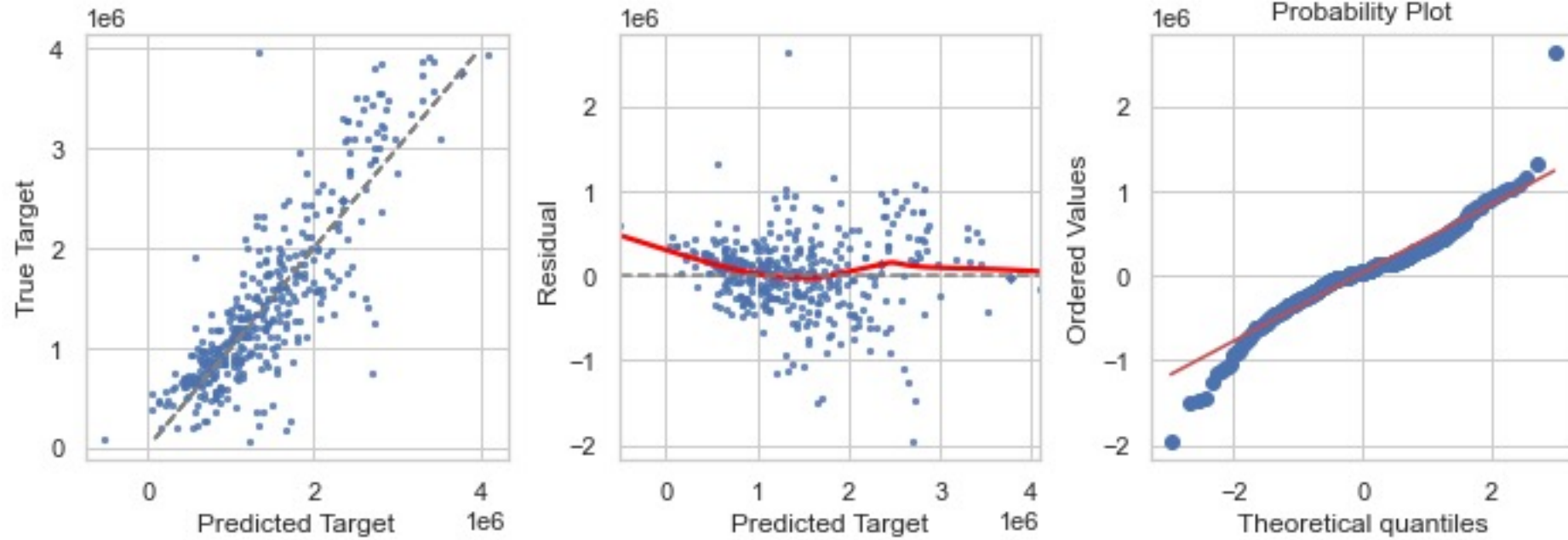Real Estate

Result

R²: 0.81

MAE: 1000+

Result cnt.

Diagnostic Plots

Future Work

Improve the model performance

Access Alternative Features

Separate the Analysis by County

Price Mean
- Alameda County: 1.15M
- Santa Clara County: 2.15M

Thank you