

A comprehensive report on

House Price Prediction

Including EDA, Recommendation model
and interactive Dashboard

SAEED SHIRANI

Introduction

The dataset contains information such as location, price, number of bedrooms and bathrooms, and more. My primary objective is to create a model that suggests the price of a house. Additionally, I aim to develop a dashboard to assess whether buying a house with certain conditions and a given price is a good decision.

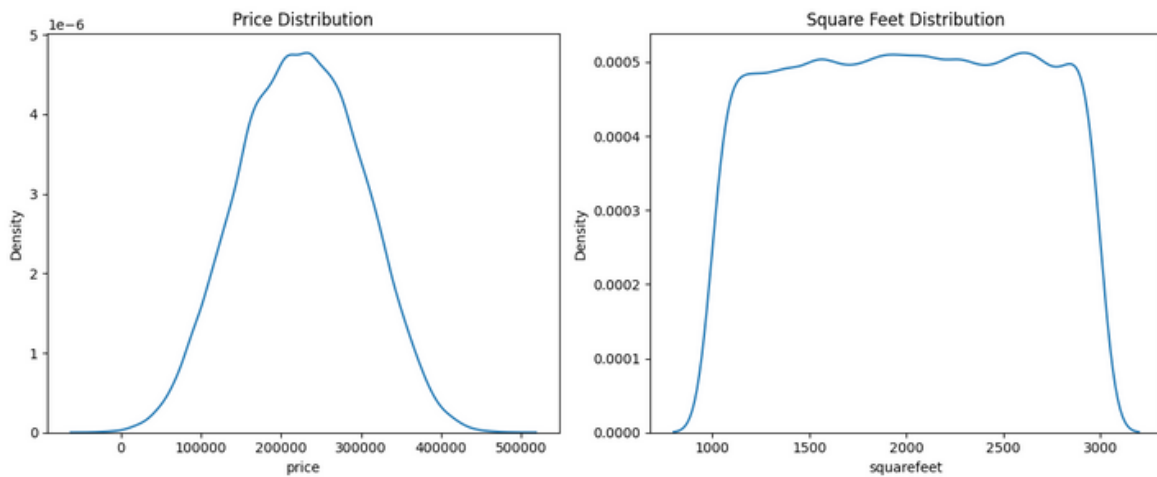
SquareFeet	Bedrooms	Bathrooms	Neighborhood	YearBuilt	Price
2126	4	1	Rural	1969	215355.283618
2459	3	2	Rural	1980	195014.221626
1860	2	1	Suburb	1970	306891.012076
2294	2	1	Urban	1996	206786.787153
2130	5	2	Suburb	2001	272436.239065
2095	2	3	Suburb	2020	198208.803907
2724	2	1	Suburb	1993	343429.319110
2044	4	3	Rural	1957	184992.321268

EDA

In various scenarios, we need to examine how changes in our data impact relationships and what factors influence these relationships. By identifying the connections and understanding how each parameter affects the price of a home, we can develop a model to suggest prices for new homes. Furthermore, we can determine whether buying a home with specific properties is a wise or risky decision.

Data Validation

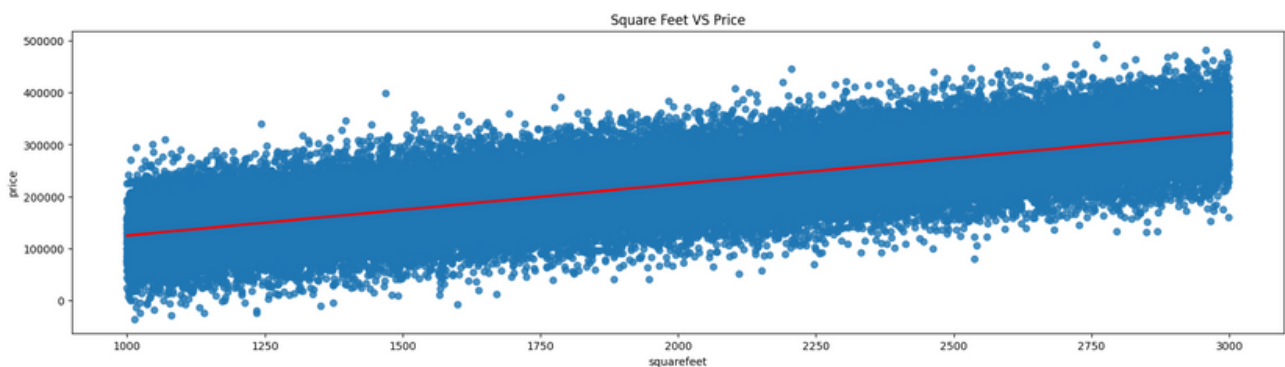
To create an effective model for price recommendations, it's essential to work with a dataset that exhibits normality. Assessing whether our dataset is normal or not is crucial.



The price data in our dataset appears to be completely normal. However, the square footage data is not perfectly normal but is approaching normalization. Despite this, the data is still valid for creating a robust price prediction model.

Relationship

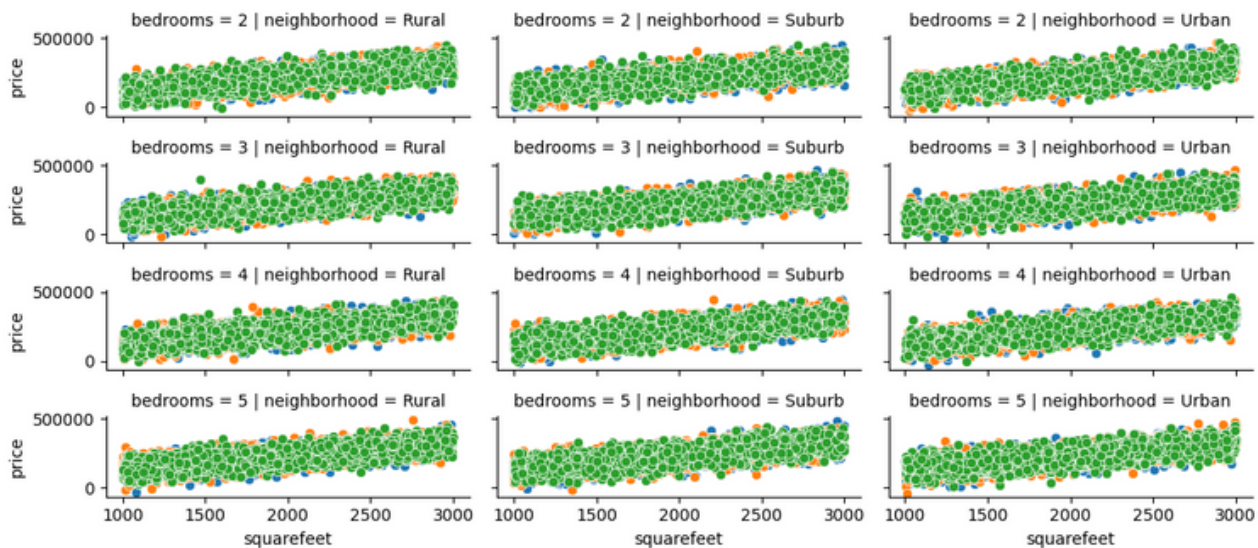
One of the most influential parameters affecting the price of a home is the area of the home. Let's fit a linear regression model to our data to explore this relationship further.



Interesting! There's an obvious and strong relationship between the price and square footage of houses. Let's segment our recorded data based on different categories to see if we can identify any changes in this relationship.

Data Segmentation

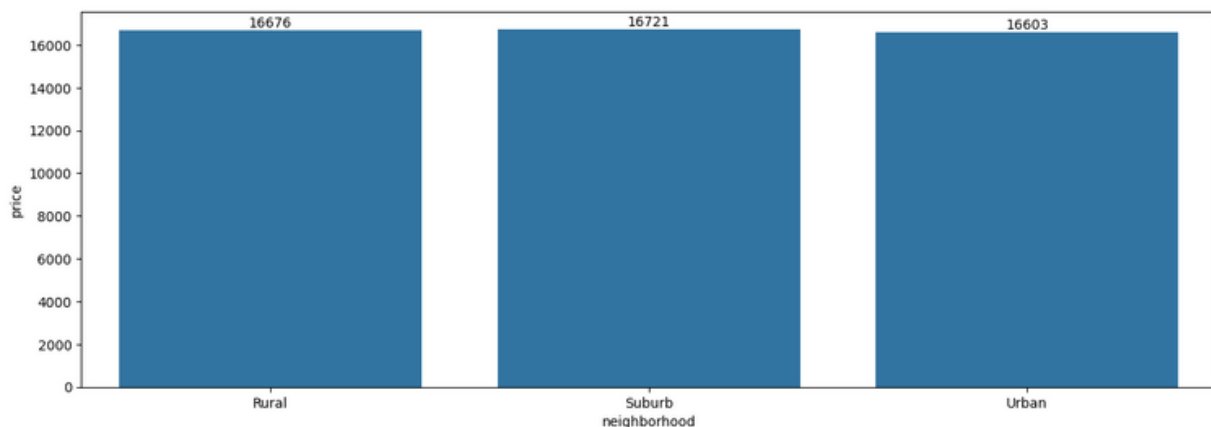
Let's segment our data based on the number of bedrooms, bathrooms, and neighborhoods to observe if there are any changes in the detected linear regression. If changes occur, we'll investigate the causes behind them.



There doesn't seem to be any significant difference in the detected linear relationship. As you can see, the linear relationship between square footage and price remains very strong and obvious even after segmentation.

Abundance Based on location

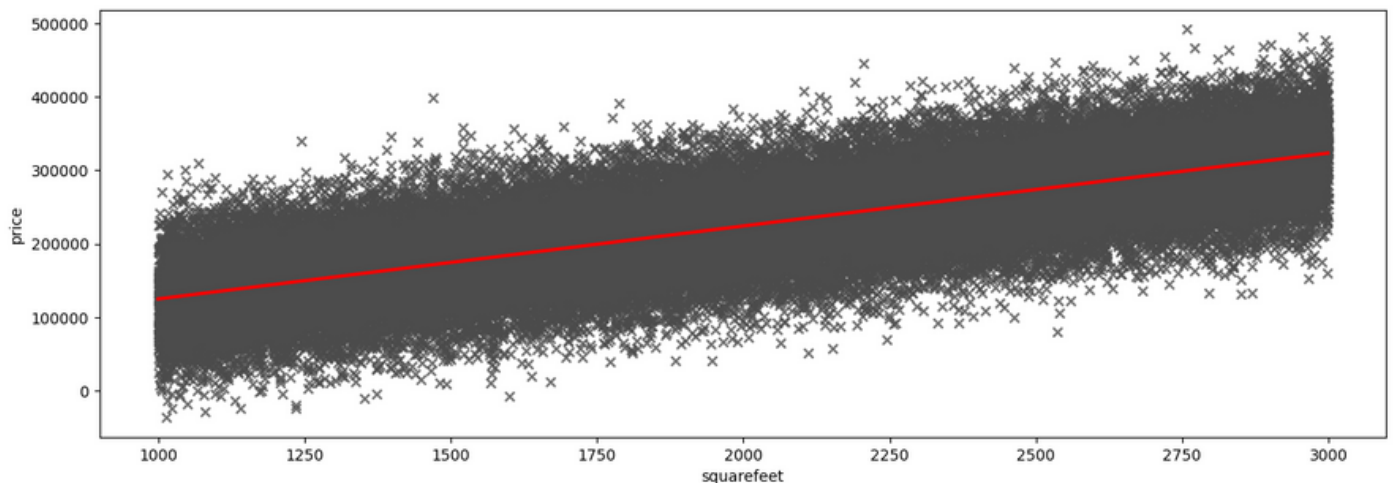
Let's see if there is any difference in the number of houses across different locations.



There is no significant difference between the number of houses in each neighborhood.

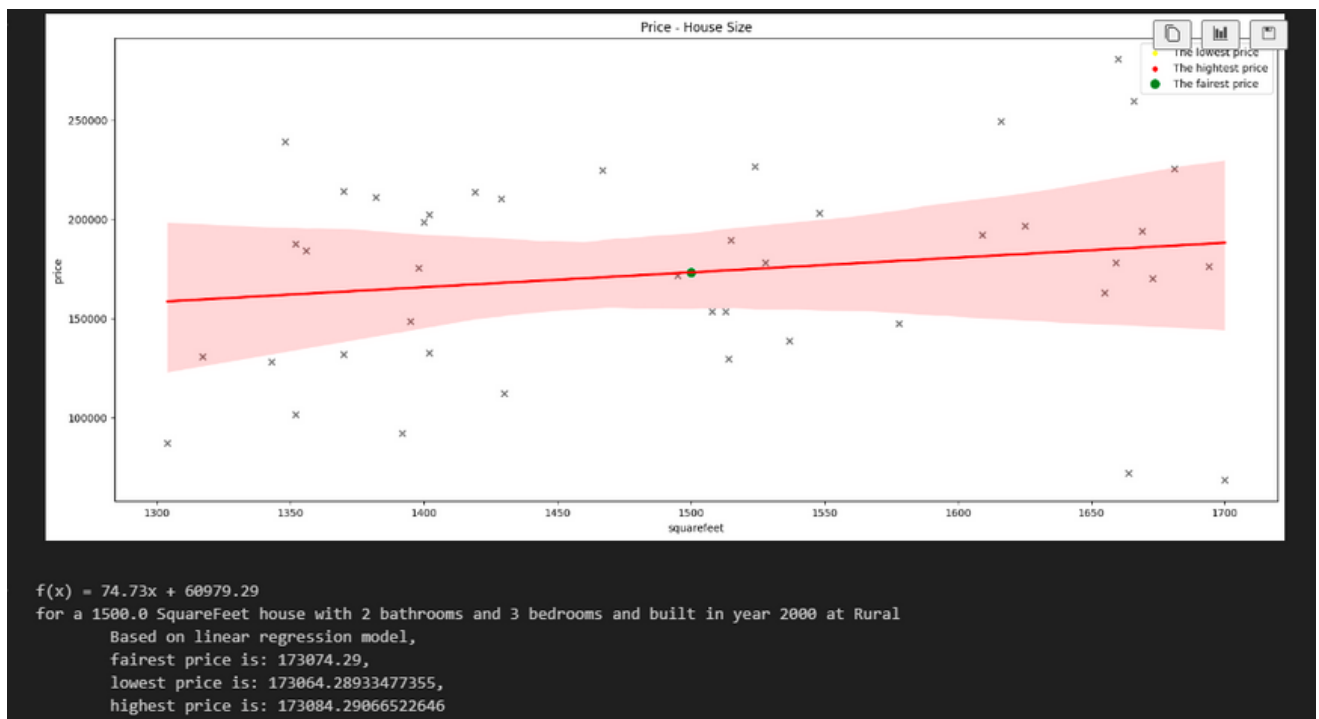
Creating model

Let's perform regression analysis on the entire dataset once again:



Linear Regretion on a part of datasets

Imagine a scenario where you want to buy a house with 2 bedrooms, 1 bathroom, an area of 1500 sq ft, and a build year between 2000 and 2010. You can select a collection of data points within intervals that closely match your expectations and perform a linear regression. In this case, you would have a line like this:

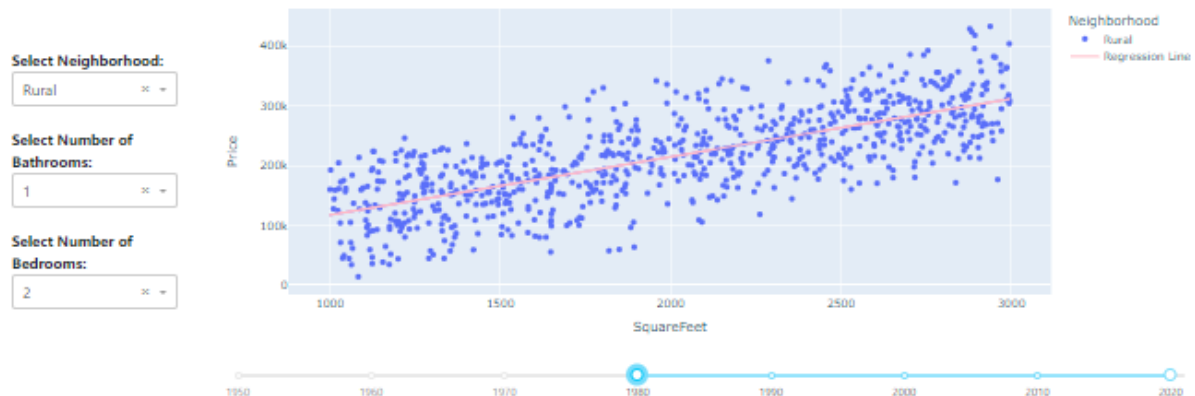


If a house falls below the line, you should definitely consider buying it, provided there are no fundamental problems with the property. The price is fair if the data point is close to the line. However, if a data point is significantly above the line, it's advisable not to buy it as the offered price is likely unfair.

Creating a tool for data selection

For a more dominant selection of data points, I have personally developed a dashboard using the Plotly-Dash library in Python. This picture provides a good overview of how this dashboard appears.

House Pricing Application



Average Price per Neighborhood



Price Prediction

Provide the house's square footage for a price estimation:

I explained and demonstrated the functionality of this dashboard in a video. To watch the video, please click on the provided link: [Project Explanation](#)