



House Sales in King County Regression

Seattle: May 2014-May 2015



TEAM MEMBERS

ML-GROUP 15

01 SAURAV LABADE

02 ISHU GUPTA

03 BUSHRA DOSANI

04 RAJESH KUMAR
SAHU

AGENDA

INTRODUCTION

5-STEP PLAN

INSIGHTS

MODEL COMPARISON

APP DEVELOPMENT

INTRODUCTION Regression

Regression models are used to predict the price of a house in seatatle based on the data from May 2014 and 2015



5 Step Plan

01

AIM

To make our evaluations based on every basic parameter that is considered while determining price

02

importing Libraries

pandas, numpy, sklearn, seaborn, matplotlib

03

Data Cleaning

Null values, noise and outliers were removed.

04

MODELS

Linear Regression, Random Forest Regressor, XGBoost, Polynomial Regression (degree 2, 3)l

05

Web App

Pycharm:index.html & main.py

Insights



● UNIVARIATE ANALYSIS

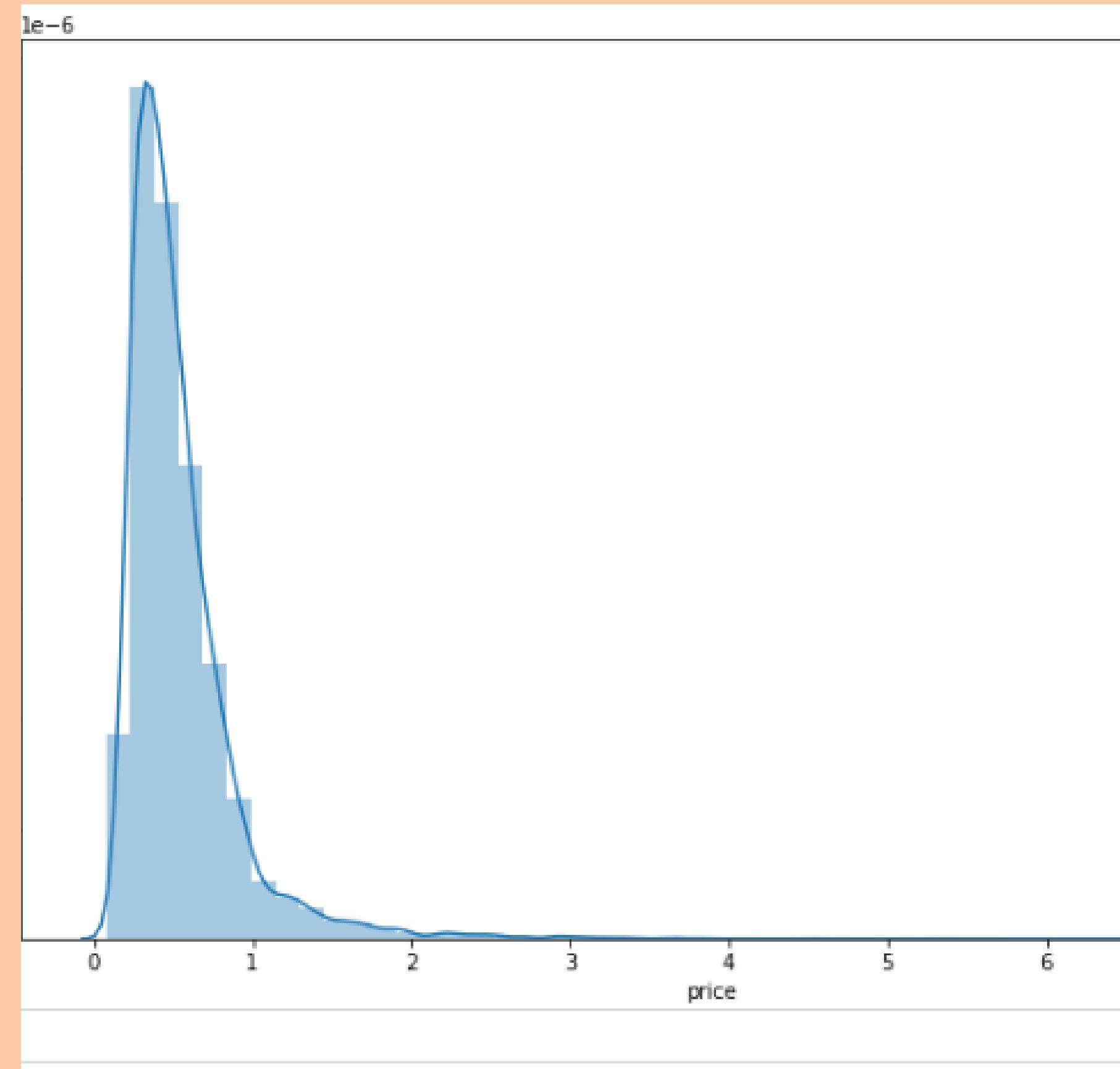
● BIVARIATE ANALYSIS

● CHISQUARE ANALYSIS

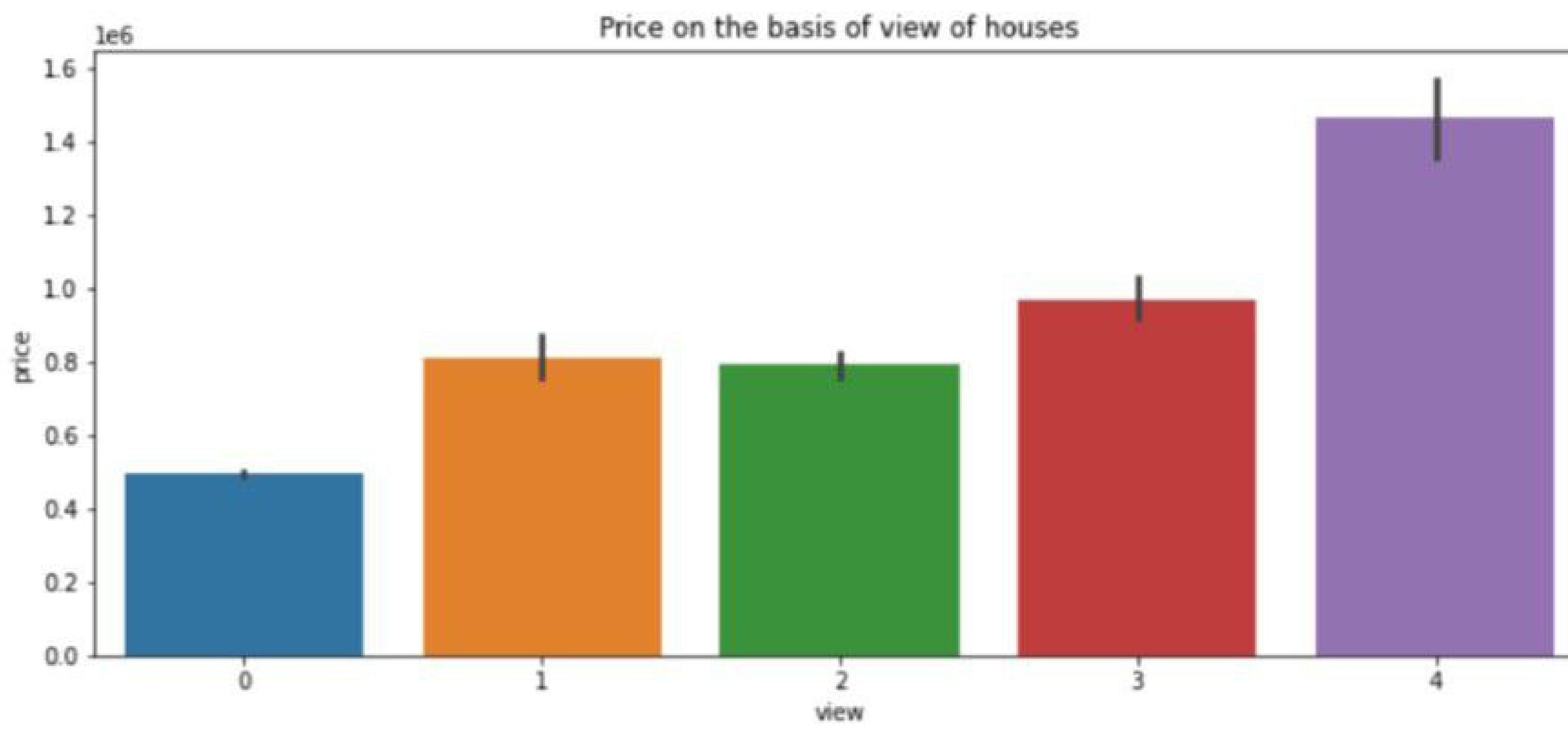
● MULTICOLLINEARITY

UNIVARIATE ANALYSIS

Most Properties sold are of
near 1 million Prices in Seattle



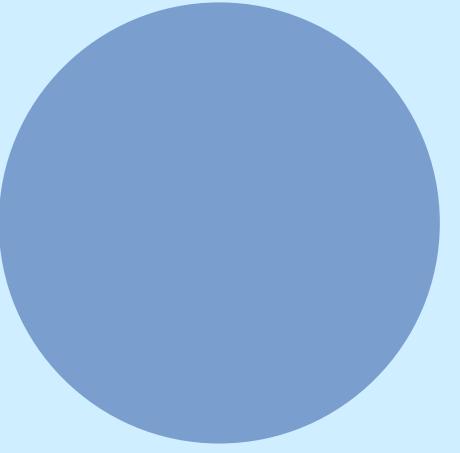
Price on the basis of view of houses



PRICE VS VIEW

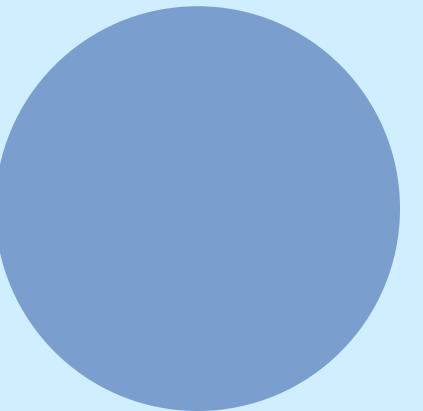
THE PROPERTY PRICE IS
DIRECTLY RELATED TO THE
VIEW IN SEATTLE

CHI SQUARE ANALYSIS



**View-
Waterfront**

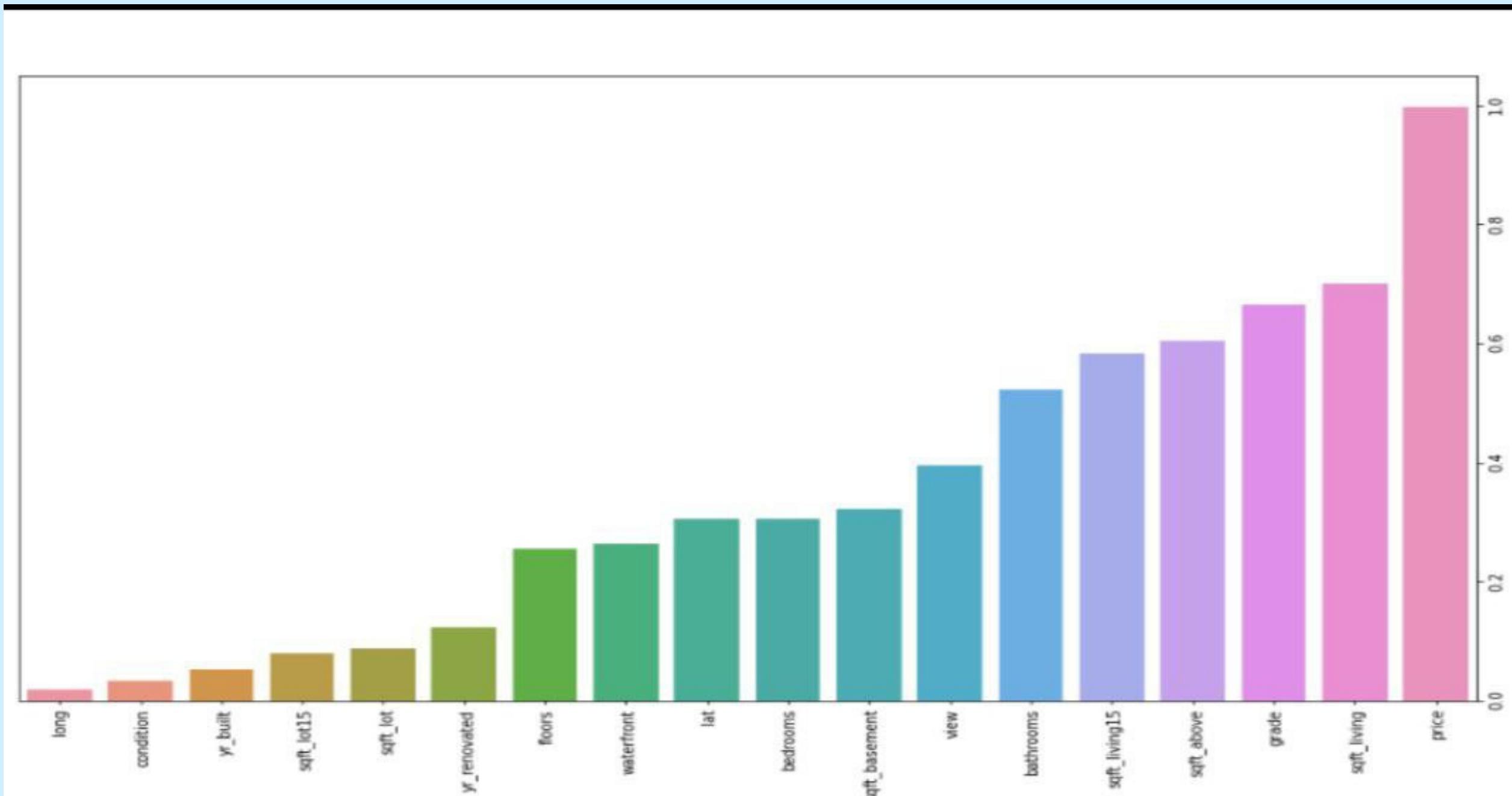
7575.56



**Condition-
grade**

2225.62

MULTICOLLINEARITY



sqft_above,sqft_basement and sqft_living shows some level of multicollinearity..But we can't drop them as they are highly correlated with the Price.

MODEL COMPARISON



RANDOM FOREST



XGBOOST



LINEAR
REGRESSION



SVR



POLYNOMIAL
REGRESSION

	Model	R-Squared Score
1	Random Forest	85.01171
2	XGBoost	82.67395
4	Polynomial Regression(D-2)	75.75459
5	Polynomial Regression(D-3)	75.75459
0	Linear Regression	65.48679
3	SVR	-4.79359

Random Forest Regressor:Highest
R2_Score

House Price Predictor

The screenshot shows a web-based user interface for predicting house prices. At the top, a header reads "Welcome to House Price Predictor". Below the header are four input fields arranged in a 2x2 grid. The top-left field is labeled "Select the grade:" and contains the value "12". The top-right field is labeled "Enter Bedrooms:" and contains the value "7". The bottom-left field is labeled "Enter Number of Bathrooms:" and contains the value "4". The bottom-right field is labeled "Enter Square Feet:" and contains the value "6000". Below these fields is a large blue button with the text "Predict Price" in white. At the bottom of the form, the predicted price is displayed as "Prediction: 2038570.0".

Random Forest Regressor Model was used to create a webapp which can predict the House price based on Parameters:
grade,sqft,bedrooms and bathrooms in Seattle.



**Thanks for
listening!**