



PREDICTING CAR PRICES USING LINEAR REGRESSION

by Sultanah Aldossari & Rawan Alharbi

CONTENT

1 Introduction

2 Dataset
Exploratory Data Analysis

3 Pre-Processing

4 Modeling

5 Conclusion

INTRODUCTION

PROBLEM STATEMENT

- The purchase of a car can take a considerable amount of time.
- It is possible to determine the actual worth of a car based on many factors.
- The aim is to develop Linear regression model based on our data

DATASET

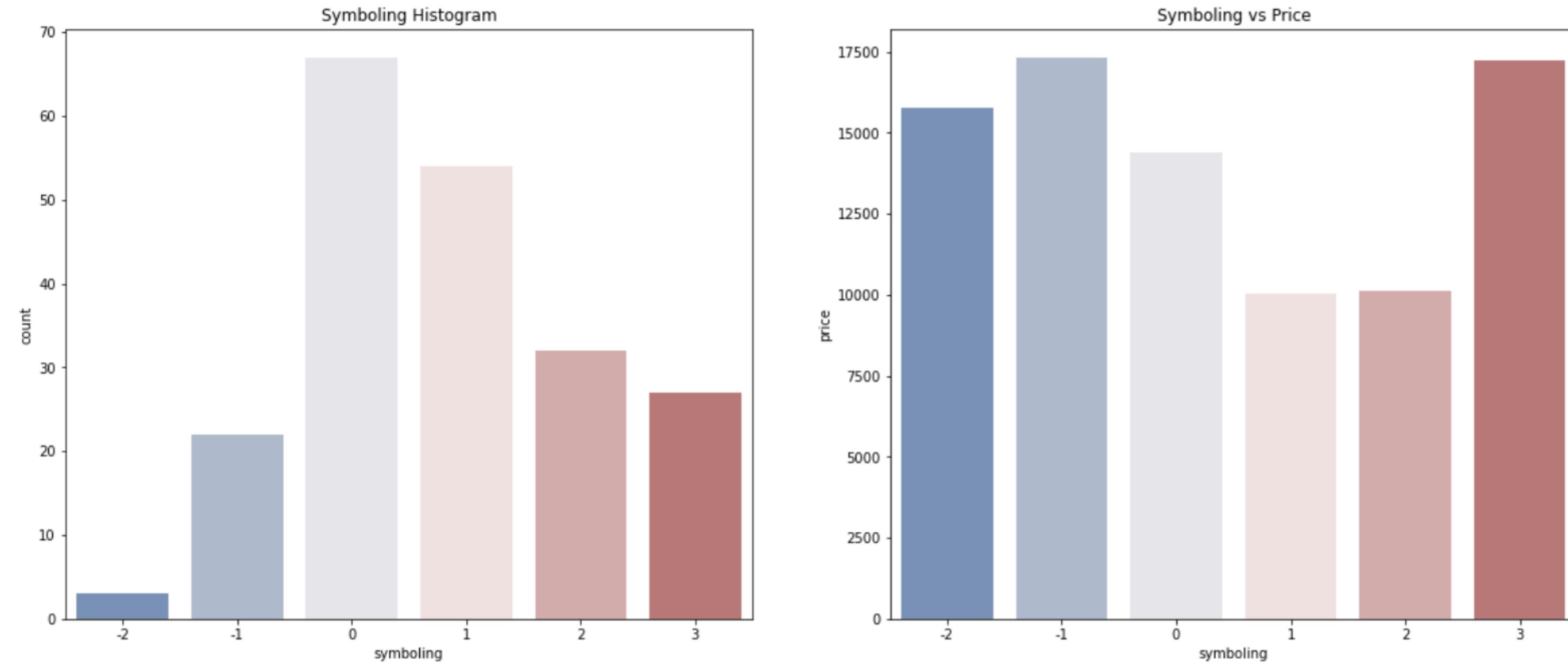
DATA

- Data scraped using Selenium.
- 205 observations, 26 features.
- Price is our target
- The independent variables are divided into Categorical and Numerical variables.

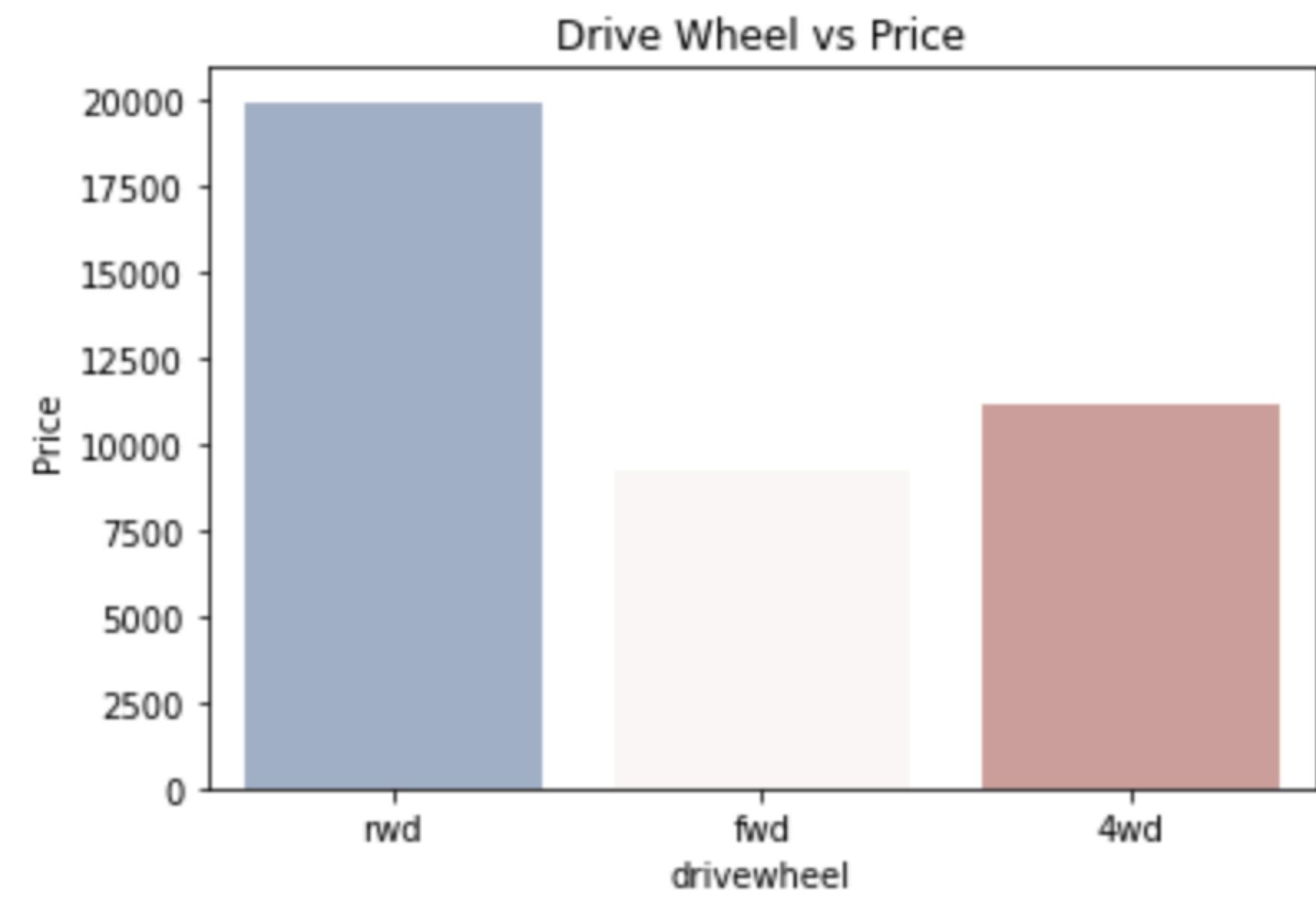
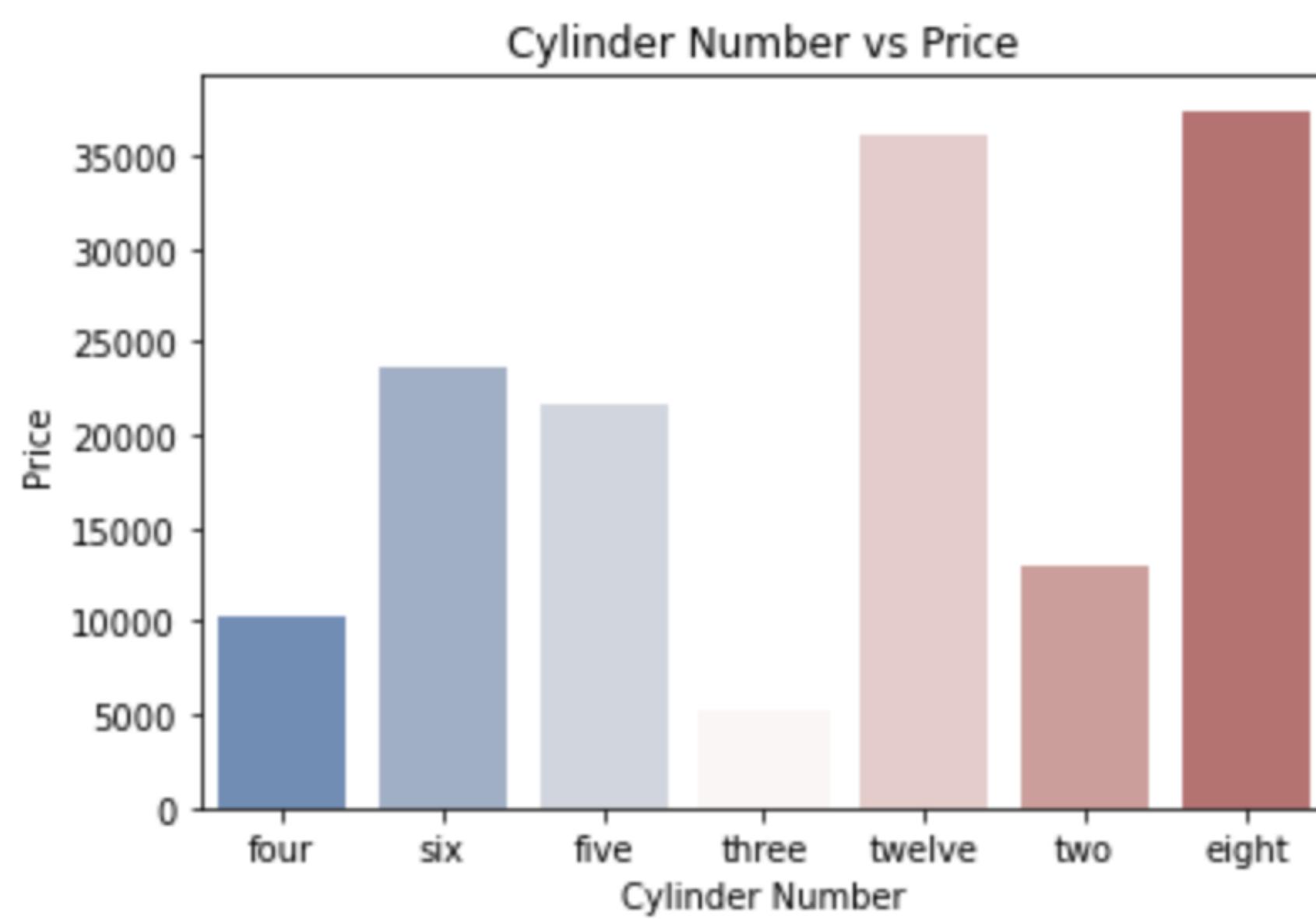
DATASET

Car Id	Drive Wheel	Car Weight	Stroke	Car Name
Symboling	Engine Location	Engine Type	Compression Ration	Price
Fuel Type	Wheel Base	Cylinder Number	Horse Number	
Aspiration	Car Length	Engine Size	Peak Number	
Door Number	Car Width	Fuel System	CityMpg	
Car Body	Car Height	Bore Ratio	HighwayMpg	

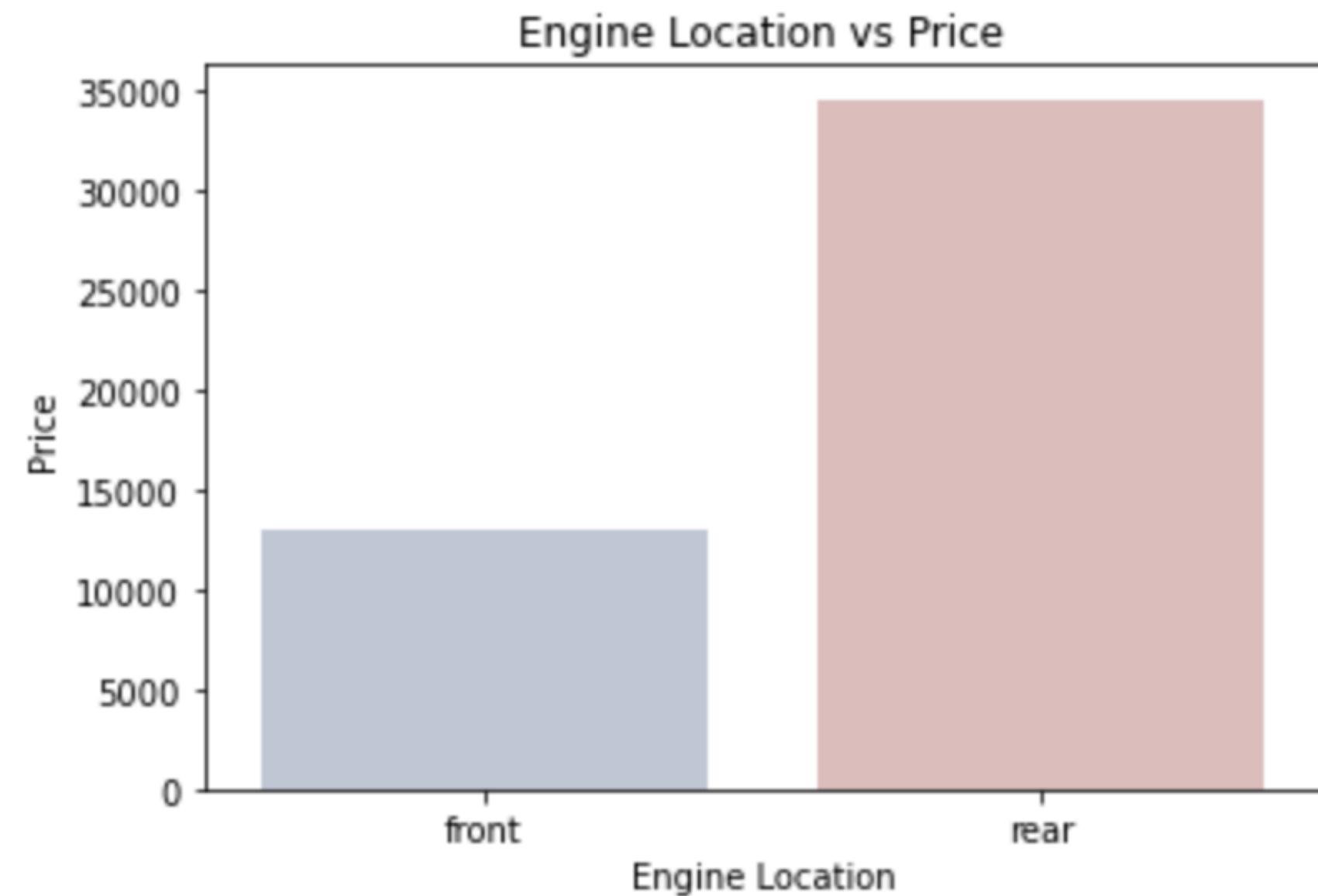
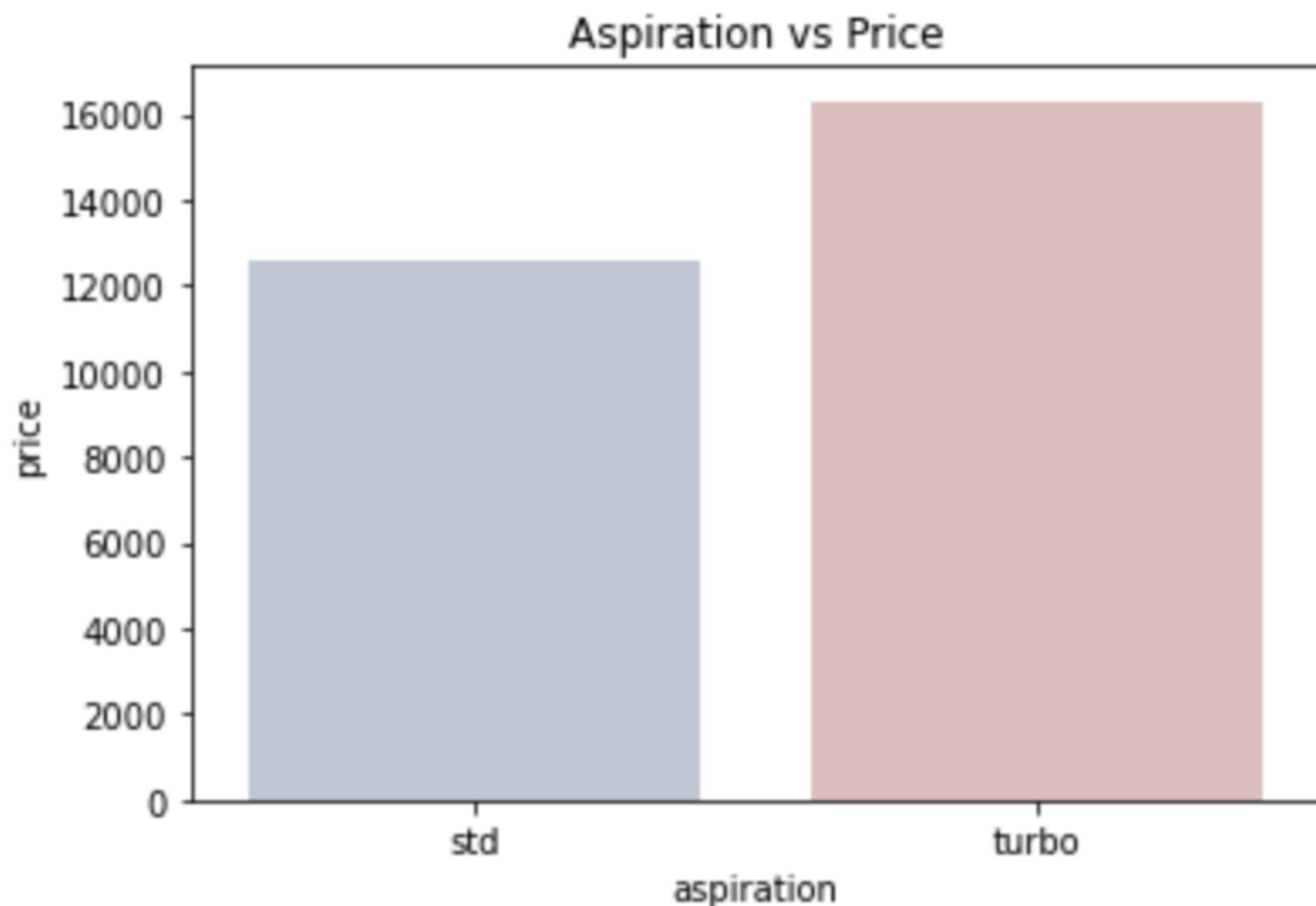
EXPOLARITY DATA ANALYSIS



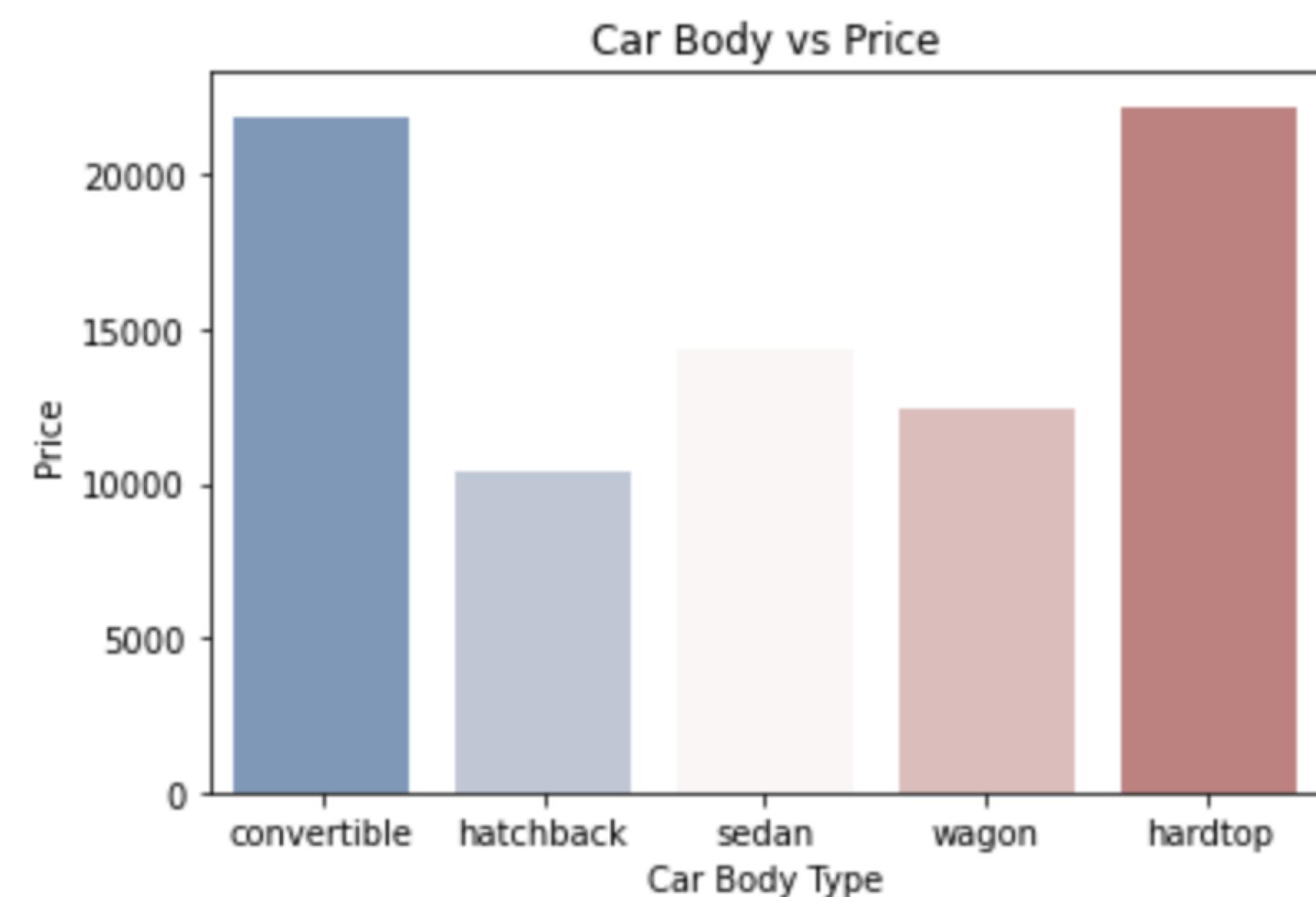
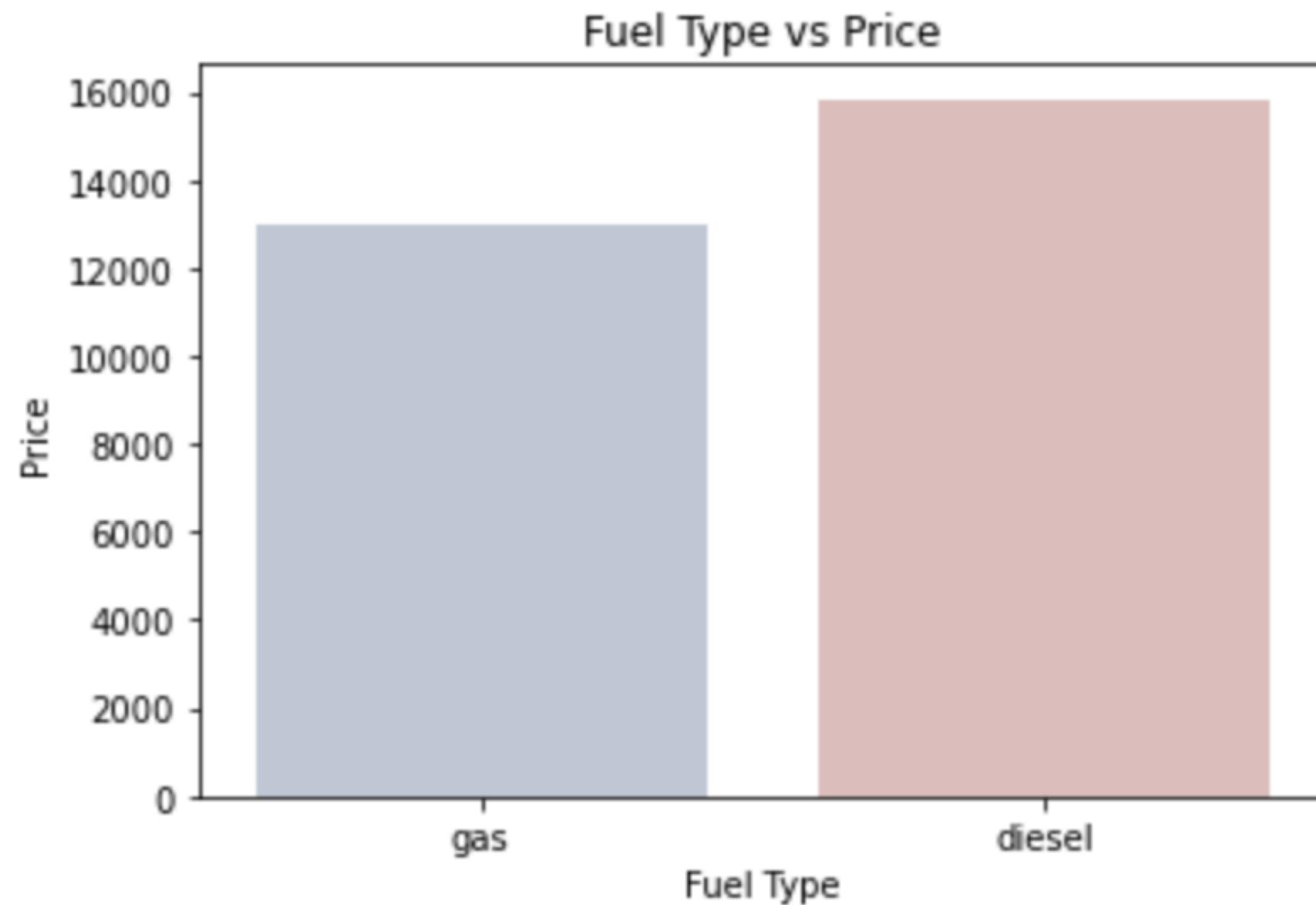
CONT: EXPOLARITY DATA ANALYSIS



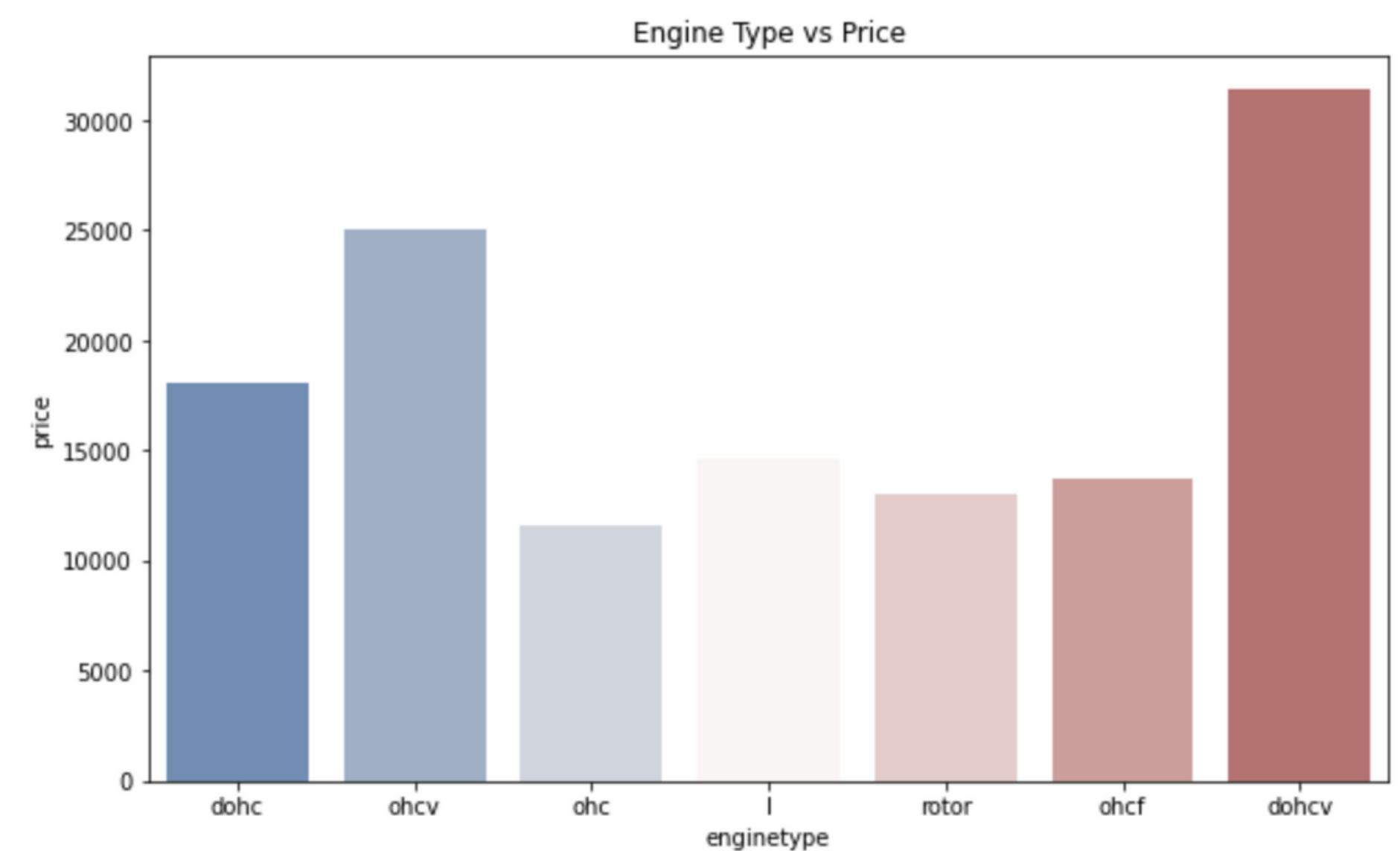
CONT: EXPOLARITY DATA ANALYSIS



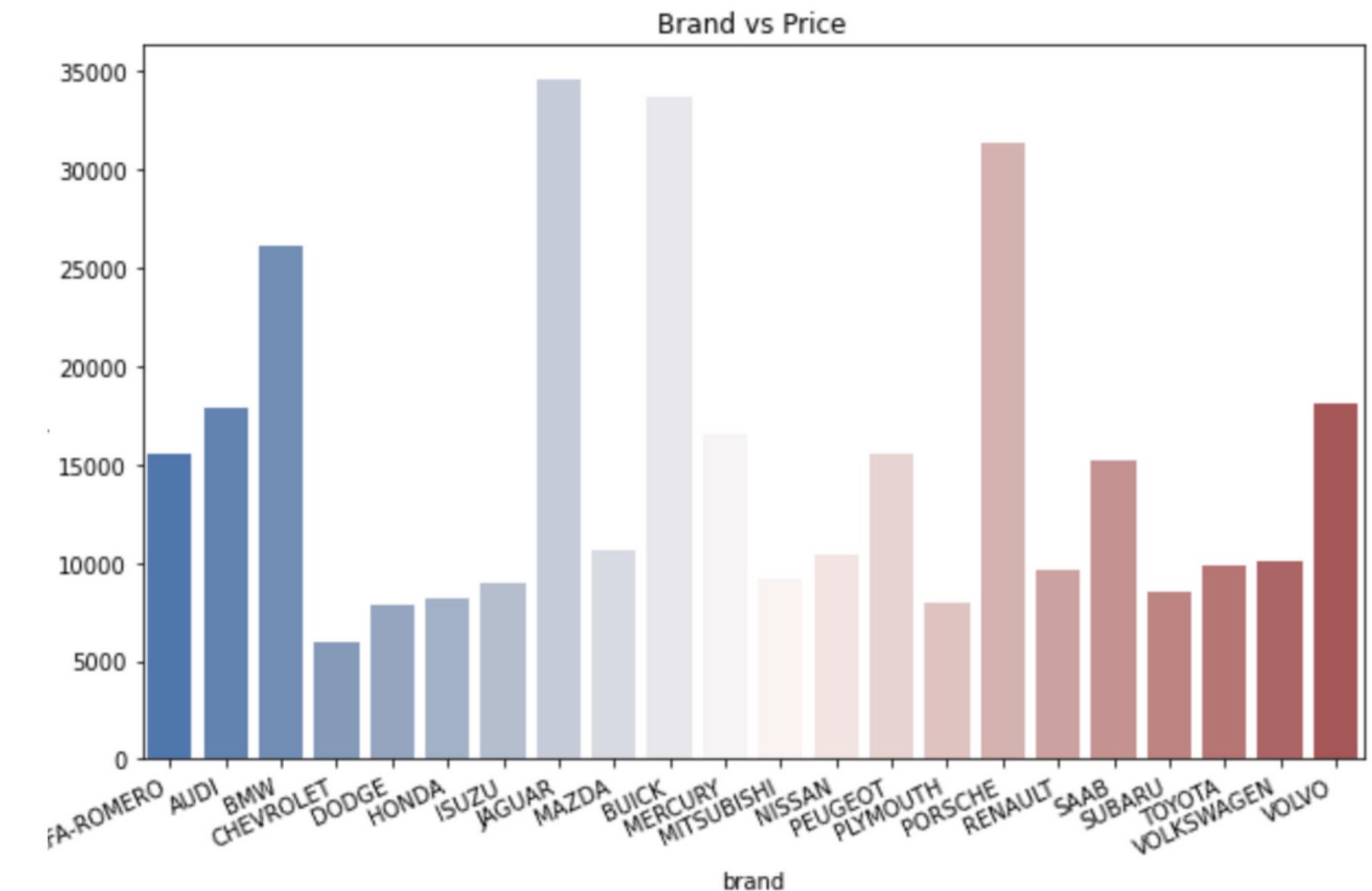
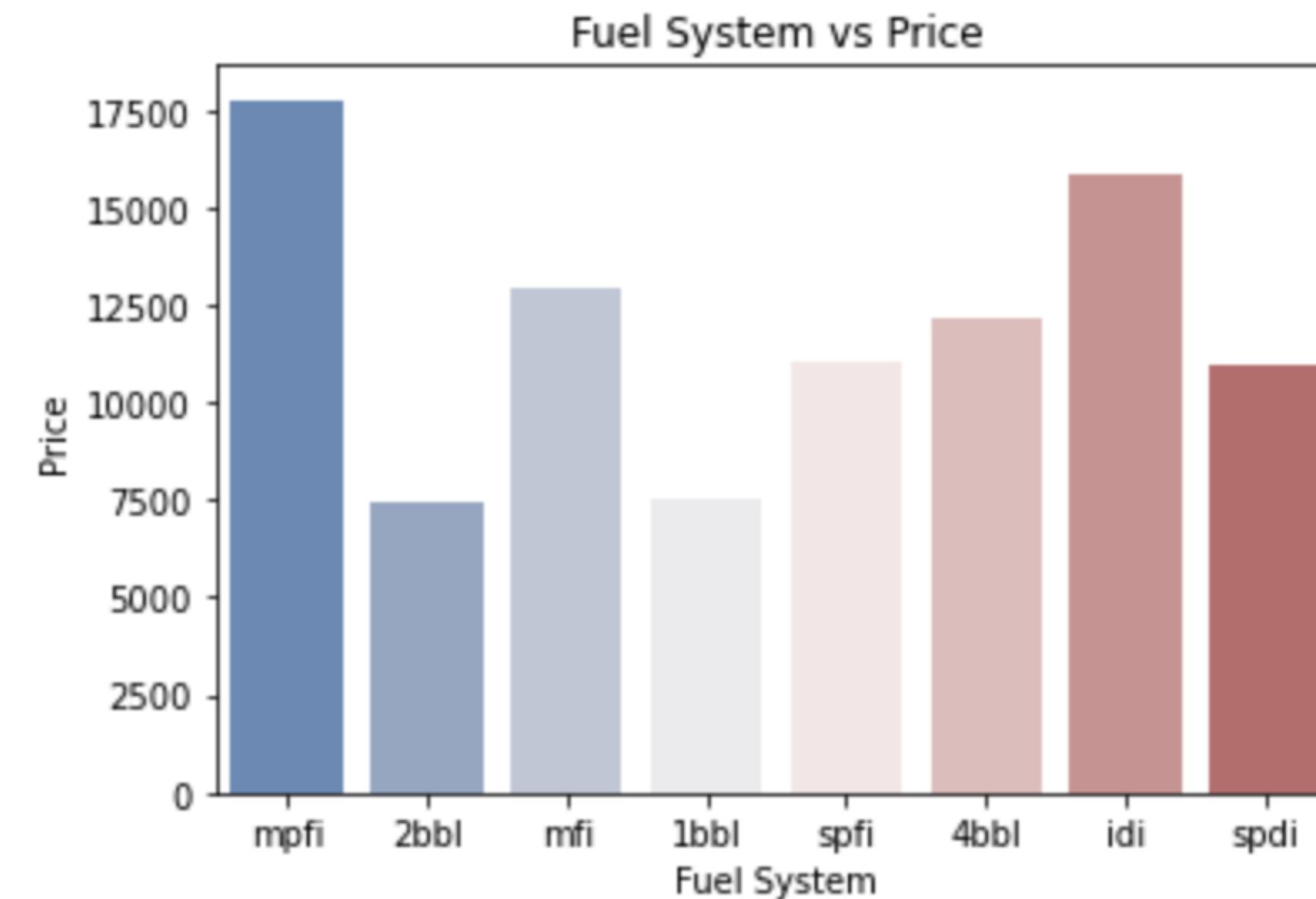
CONT: EXPOLARITY DATA ANALYSIS



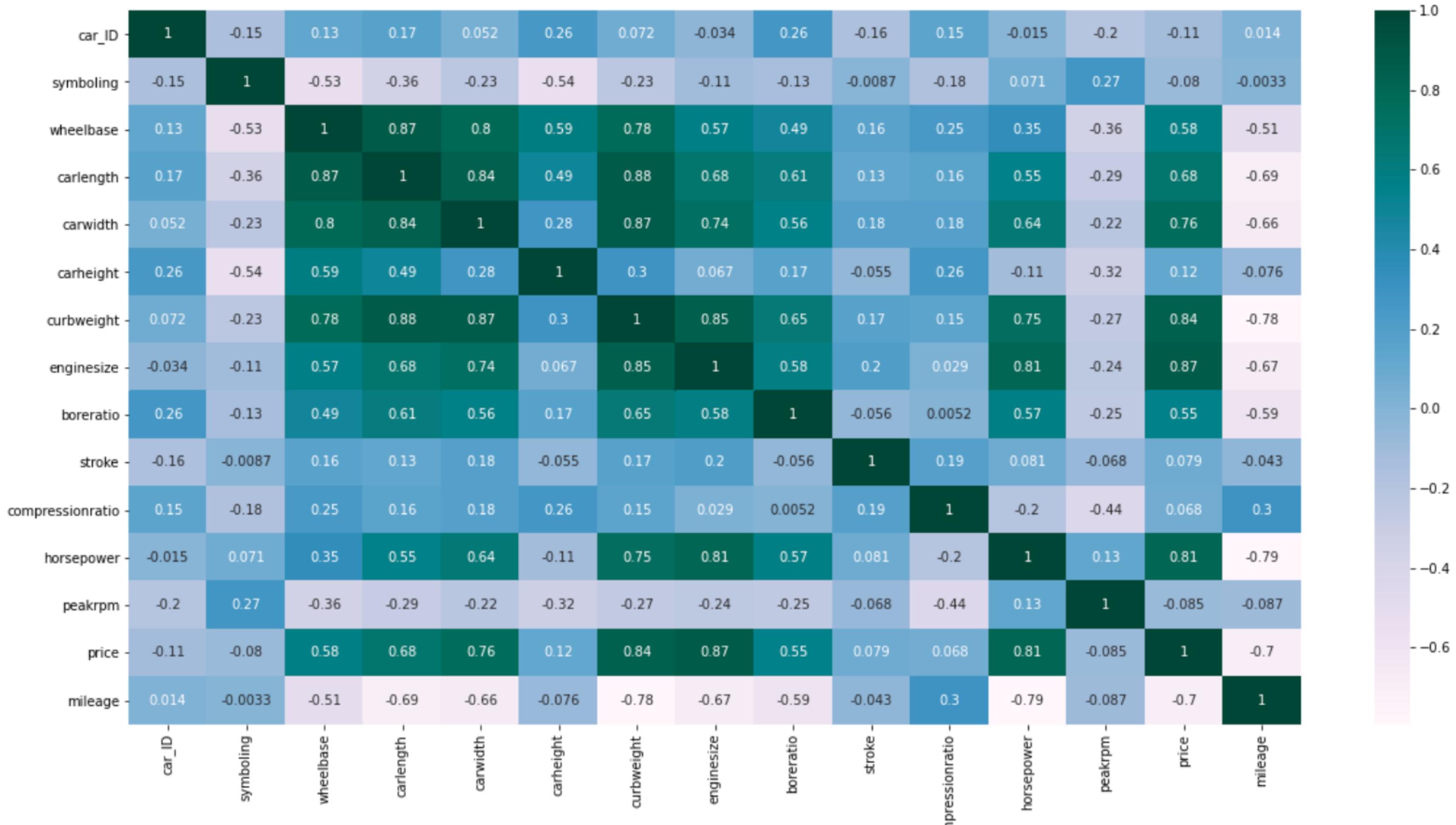
CONT: EXPOLARITY DATA ANALYSIS



CONT: EXPOLARITY DATA ANALYSIS



CONT: EXPOLARITY DATA ANALYSIS



CONT: EDA

- After visualization we can derive some features that significantly affect the price .

PRE-PROCESSING

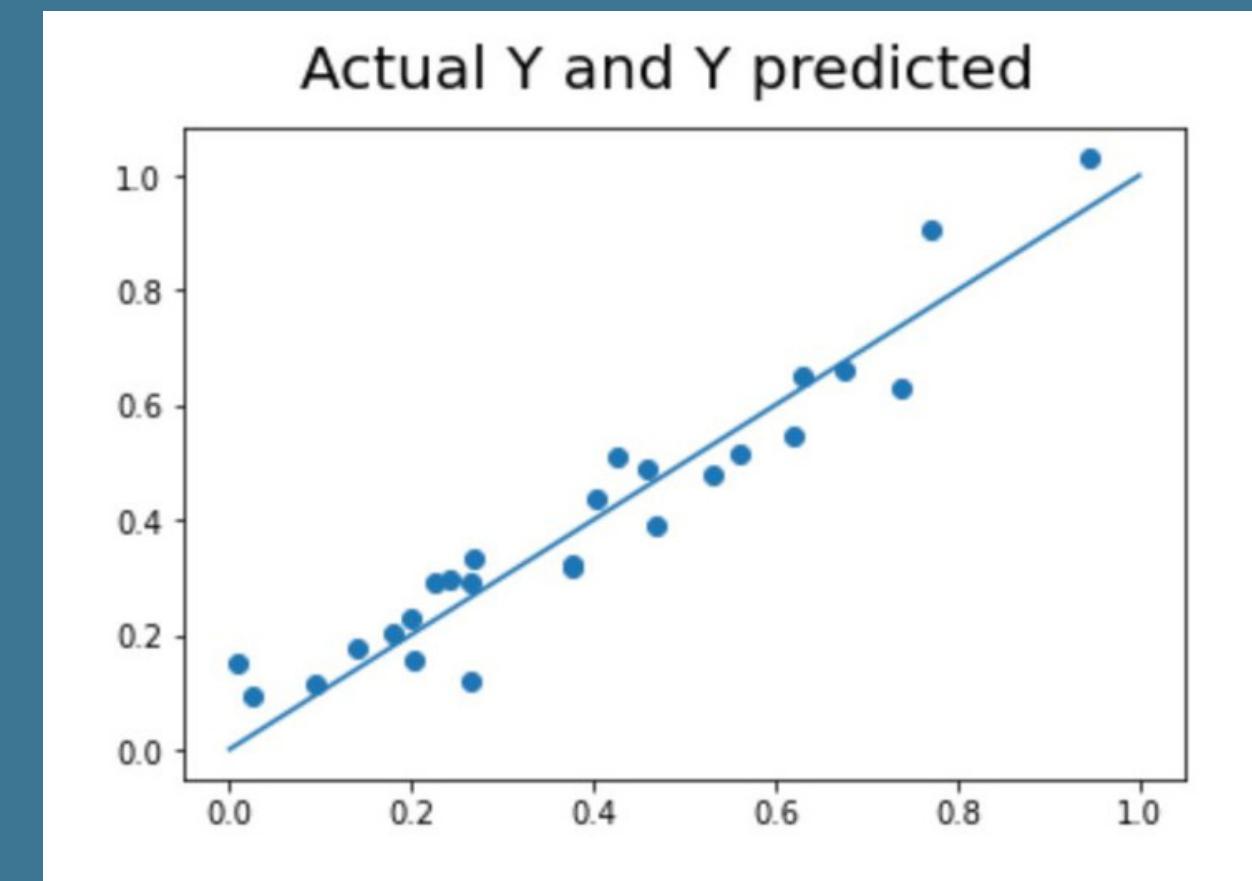
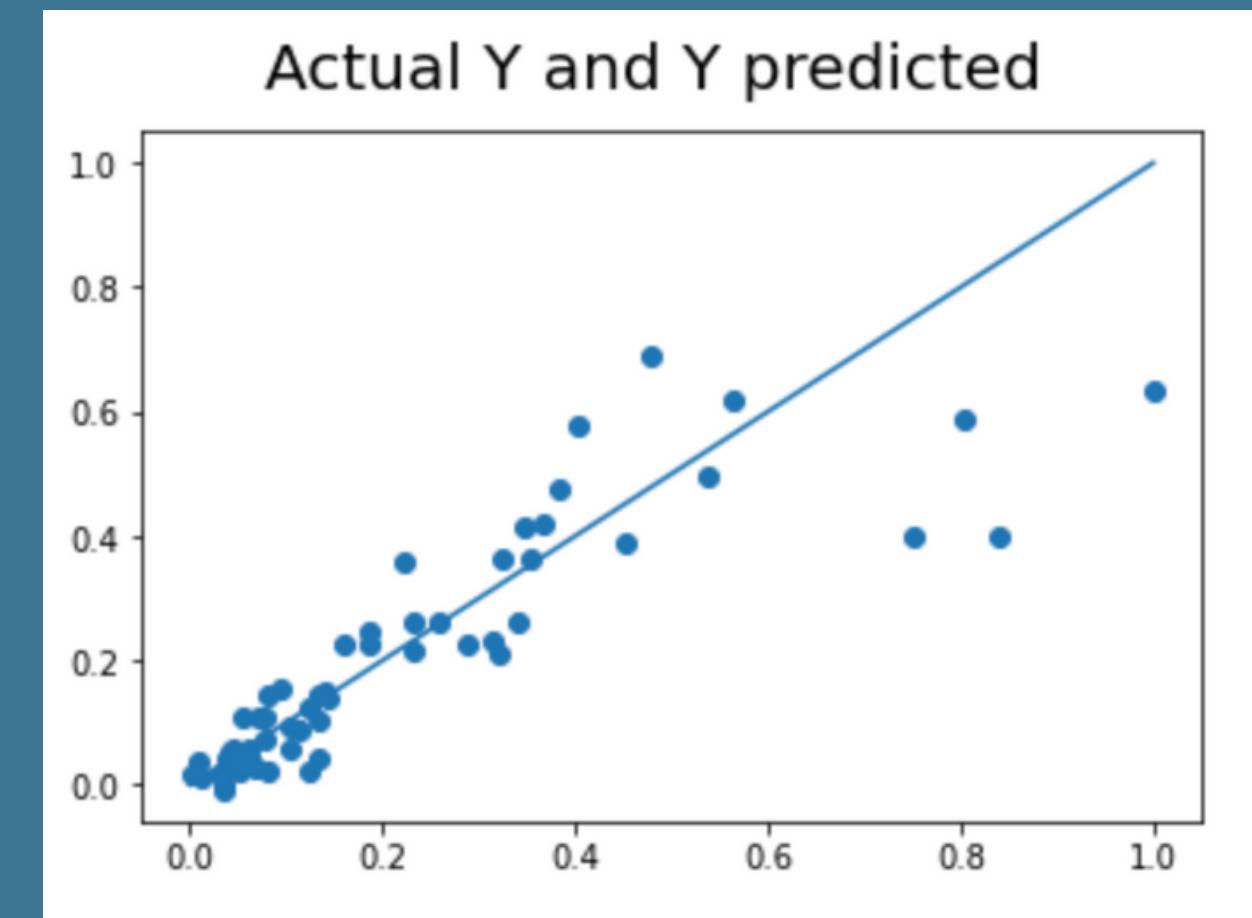
1. Converting categorical data into dummy variables
2. Derive a new features from our data
3. Scaled our data into the same scale
4. Created a new dataframe

MODELING

Linear Regression

- Split the data into train and test
- Sklearn library were used

MSE Test Score = 0.012

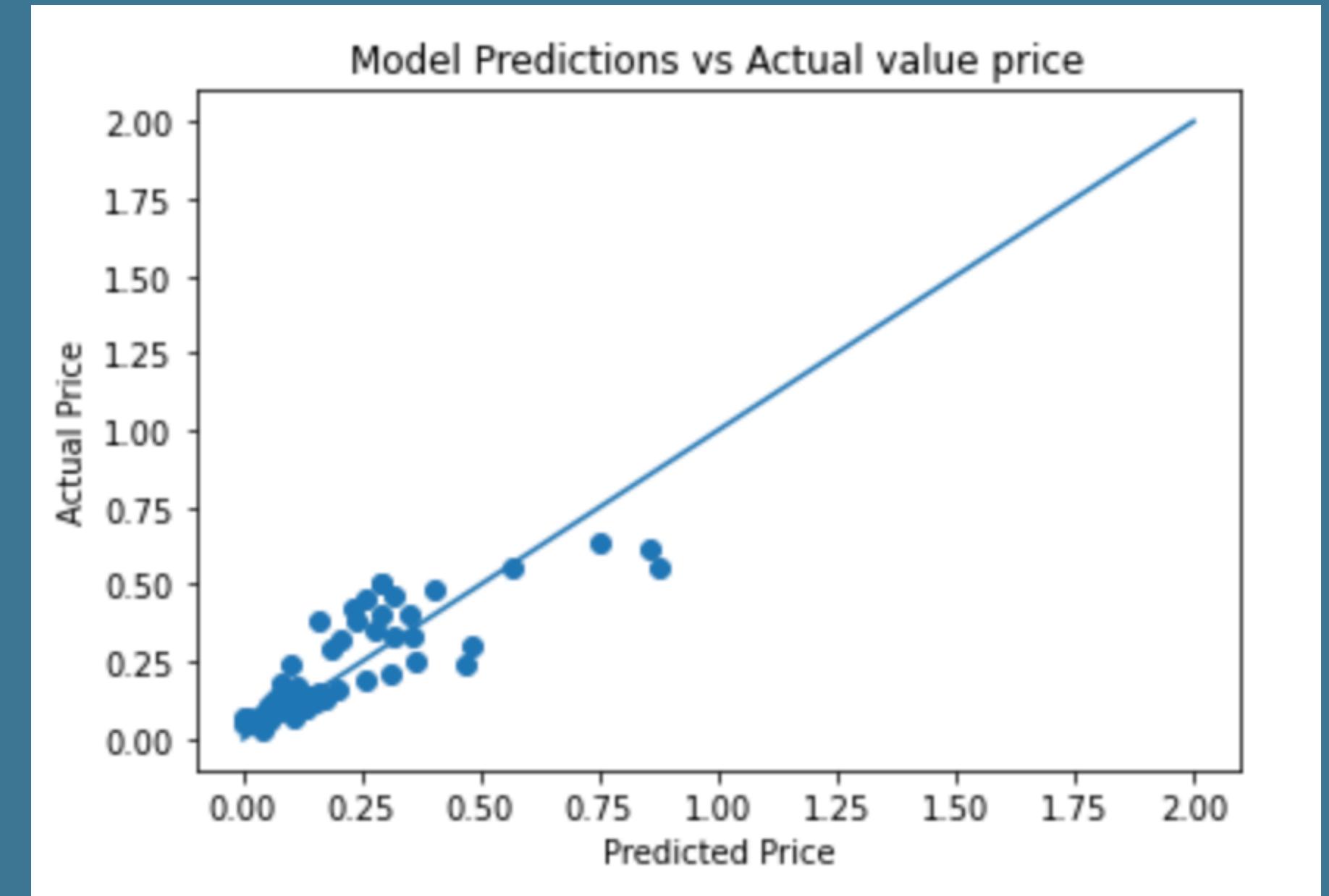


MODELING

Lasso Regression

- Split data into train and test
- The best Alpha value were 0.01

MSE Test Score = 0.012

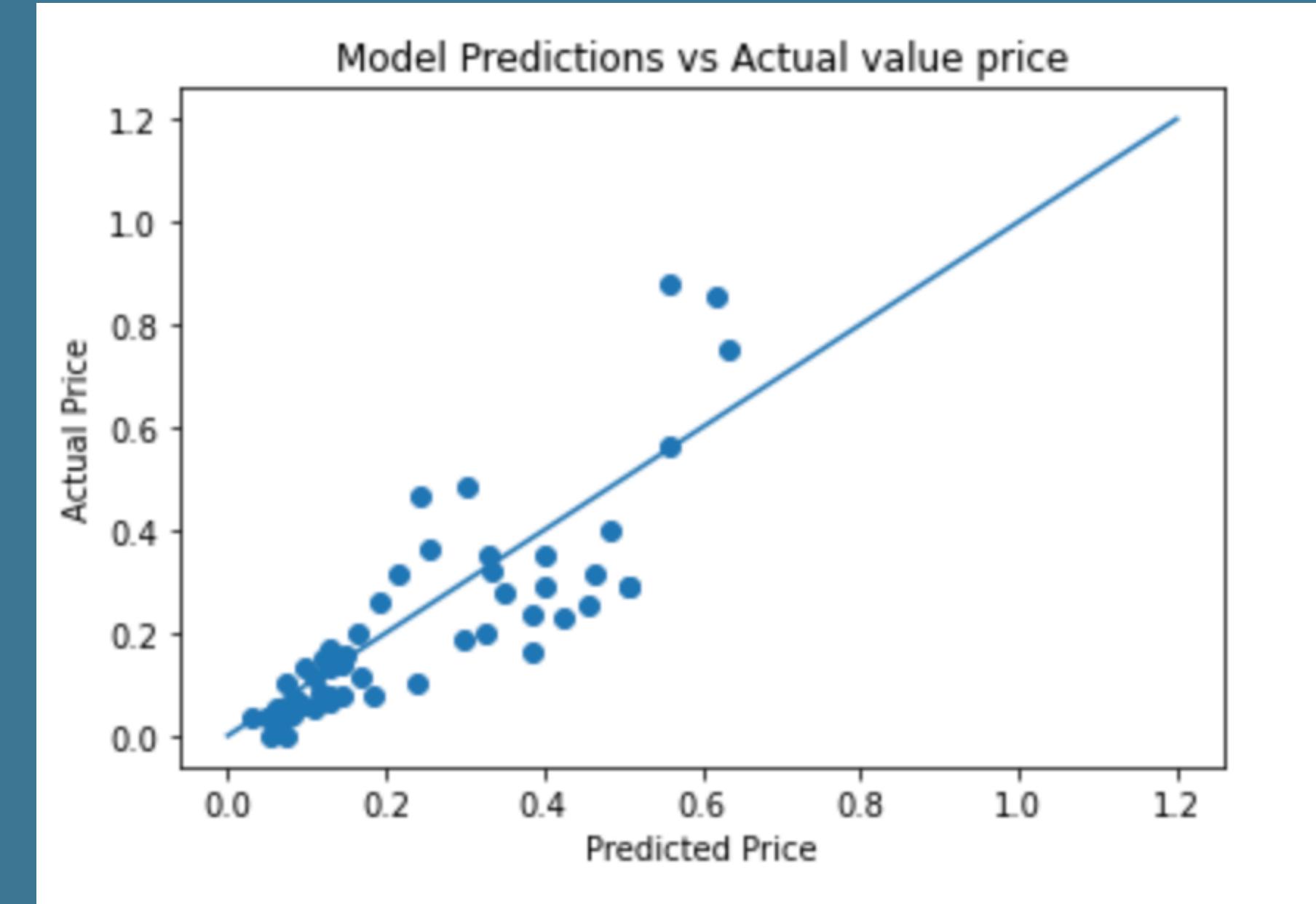


MODELING

Lasso Regression

- After Feature Selection

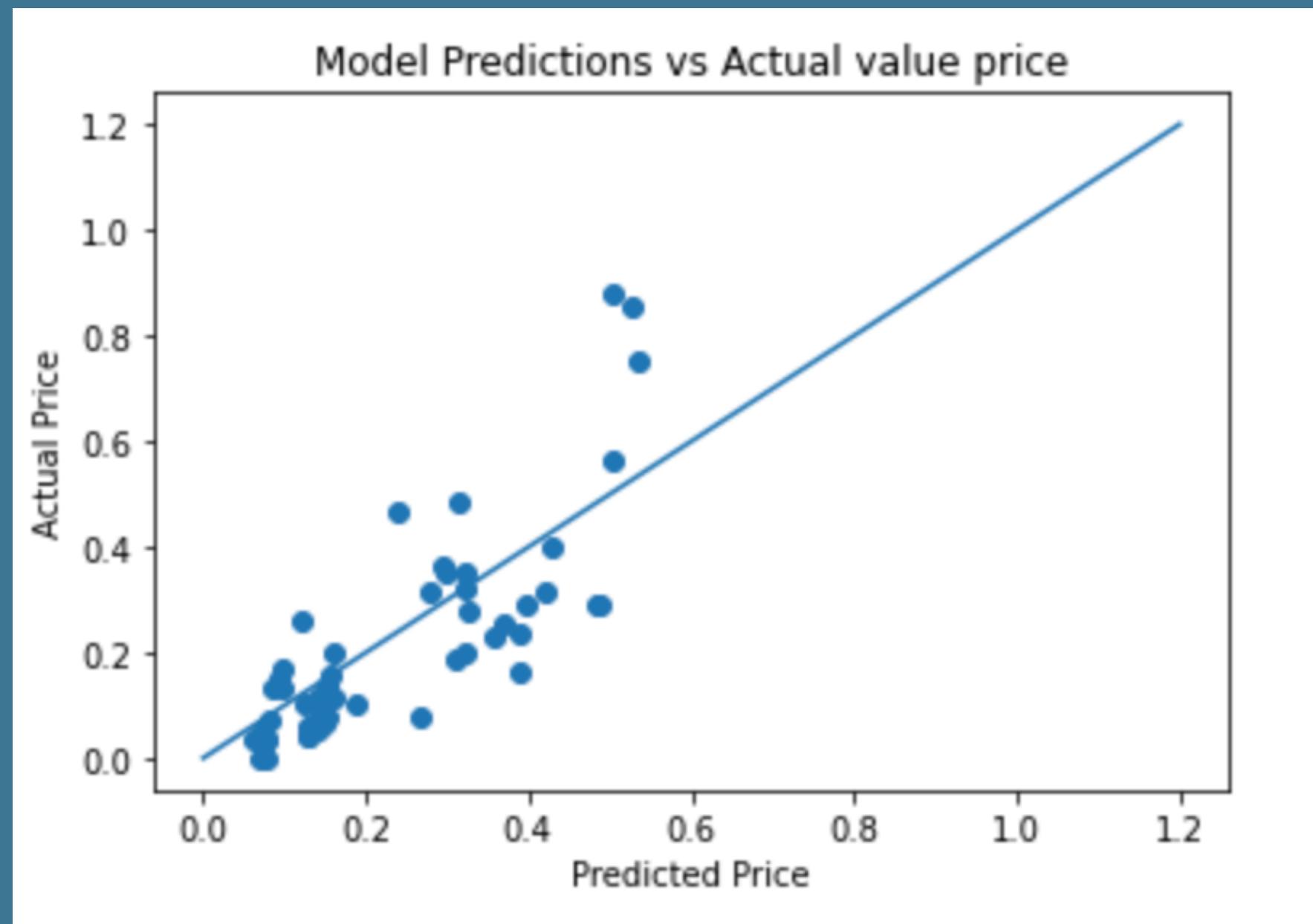
MSE Test Score = 0.011



MODELING

Ridge Regression

MSE Test Score = 0.013



CONCLUSION

The background features a dark teal color with abstract white shapes. On the left, there are two large, semi-transparent circles: one light blue circle on top and one darker blue circle below it. Both circles overlap, creating a layered effect. To the right of the circles, several thin, black, intersecting lines form a grid-like pattern that tapers towards the right edge of the frame.

THANK YOU!