Code Plateau Fellowship;
Data Science Capstone Project 3

By: DAVID VICTOR LASSON

CONTEXT

- Geely Auto car prices in Nigerian Market

OBJECTIVE!

▶ The objective of this project is to model the price of cars with the available independent variables. And this is expected to help the management to understand how exactly the prices vary with the independent variables. They can accordingly manipulate the design of the cars, the business strategy etc. to meet certain price levels. Further, the model will be a good way for management to understand the pricing dynamics of a new market

VITAL QUESTIONS

▶ Which variables are significant in predicting the price of a car

▶ How well do those variables describe the price of a car Based on various market surveys, you have gathered a large data set of different types of cars across the Nigerian market.

APPROACH

By developing Machine Learning Model to explore and point out factors on which the pricing of cars depends in a typical Nigerian market.

CORRELATION

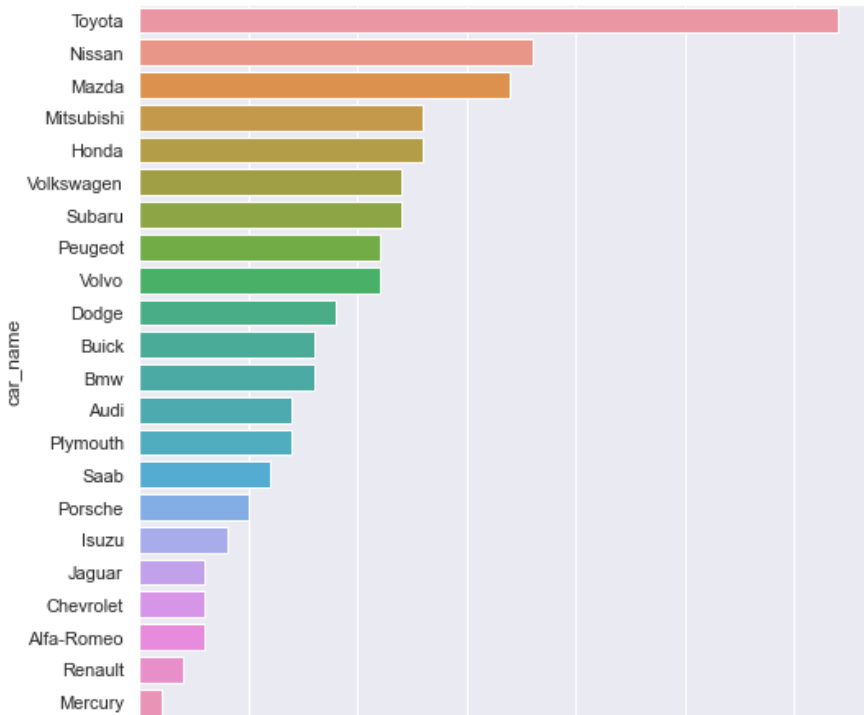▶ This correlation is majorly between numeric features of the dataset.
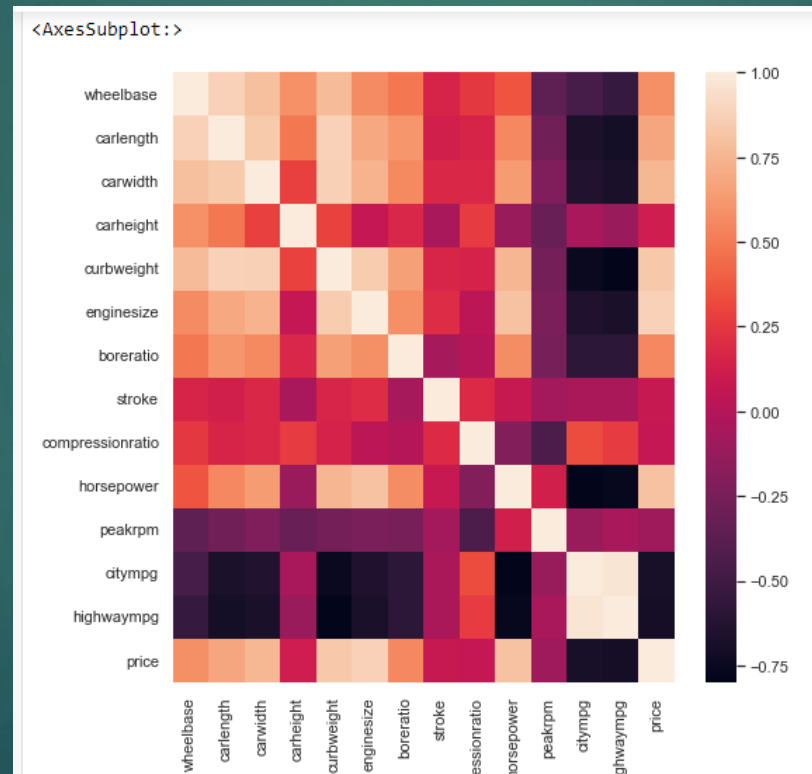
UNIVARIATE ANALYSIS

Toyota is the most suitable car. Gas is the most preferd fuel type


```
15]: dat=pd.DataFrame(df['CarBrand'].value_counts()).reset_index().rename(columns=
```
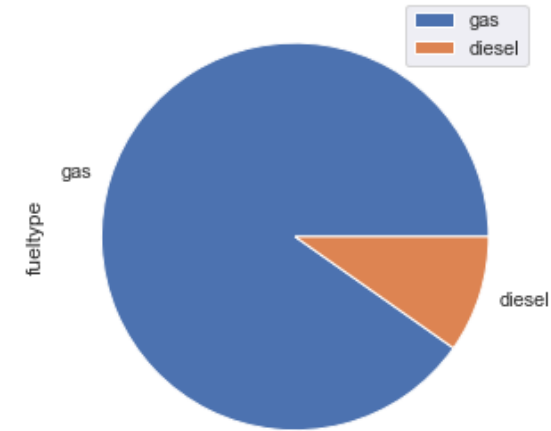
```
16]: plot = sns.barplot(y='car_name',x='count',data=dat)
     plot=plt.setp(plot.get_xticklabels(), rotation=80)
```


FUEL TYPE

```
[17]: dat=pd.DataFrame(df['fueltype'].value_counts())
```

```
[18]: plot = dat.plot.pie(y='fueltype', figsize=(5, 5))
```
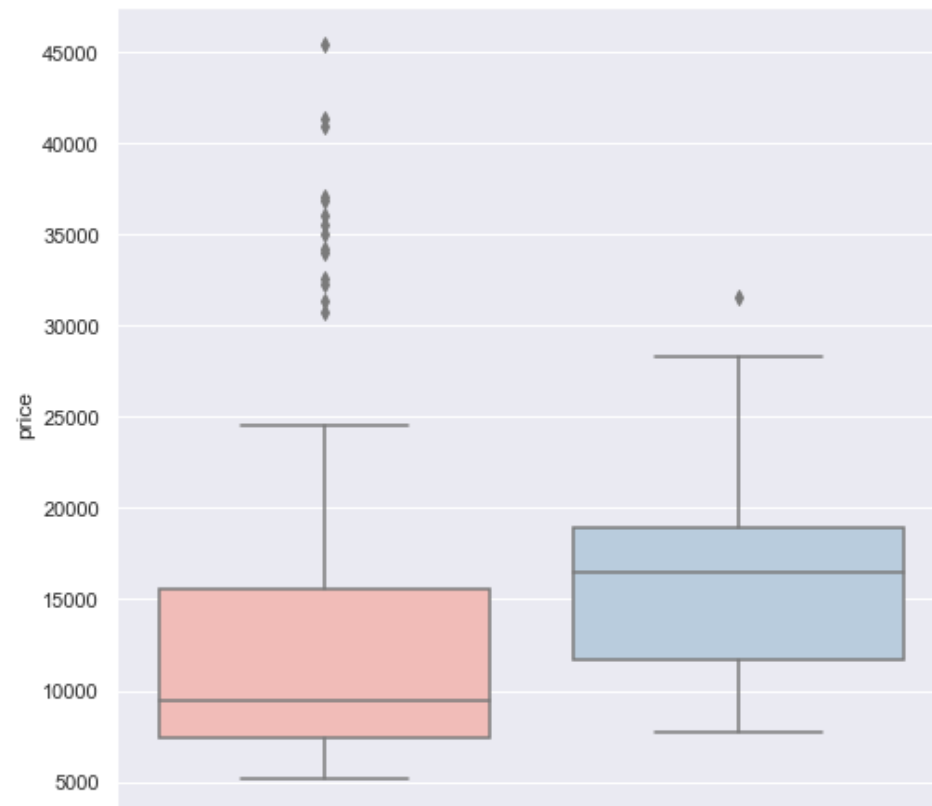
We can therefore infer from the above that most cars uses gas

BIVARIATE ANALYSIS

BIVARIATE ANALYSIS CONTD
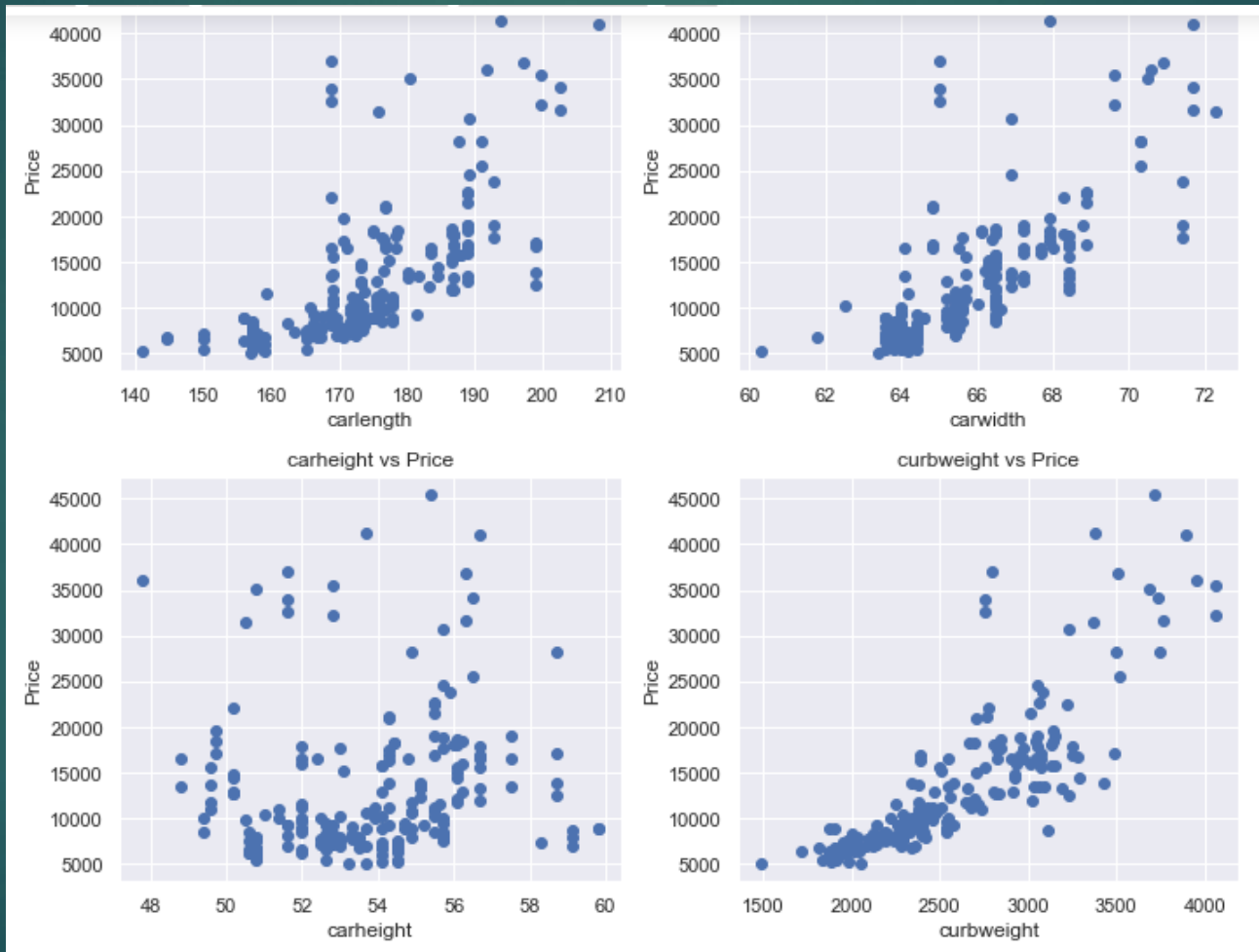
MODEL & EVALUATION

Random Forest Regressor

```
In [49]: Rf = RandomForestRegressor(n_estimators = 15,
                                     criterion = 'mse',
                                     random_state = 20,
                                     n_jobs = -1)
         Rf.fit(x_train,y_train)
         Rf_train_pred = Rf.predict(x_train)
         Rf_test_pred = Rf.predict(x_test)


         r2_score(y_test,Rf_test_pred)

Out[49]: 0.9129541993611439
```

CONCLUSION

▶ Having applied three models that Linear Regression, Decision Tree Regressor and Random Forest Regressor and subjecting each to accuracy evaluation, we can now have a satisfied model from the random forest Regressor model based on its higher accuarcy performance to predict the prices of cars by Geely Auto in Nigerian market.

- Toyota is the most suitable car whereas mercury is the least suitable car.

- Most of the cars uses gas.

- Most of the cars have standard aspiration.

- Maximum number of engine type are of 'ohc'.

- Most of the cars have fwd drive wheel.

- Sedan is the most preferred Car Body followed by hatchback.

- Top important features that affects car pricing in the Nigerian market are: Carboby, Wheelbase, Carlength, Carwidth, Carheight, Carweight, Enginesize, Carbrand etc.