CAR PRICE PREDICTION

NORA IBRAHEEM

BUSINESS QUESTION ?!

- We are required to model the price of cars with the available independent variables. It will be used by 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.

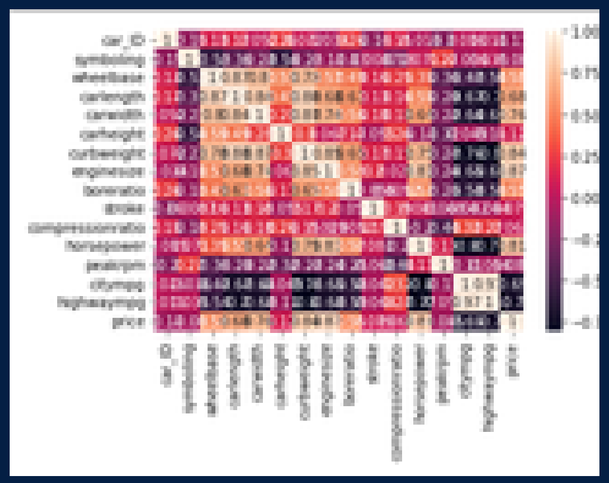
DATA PREPERATION

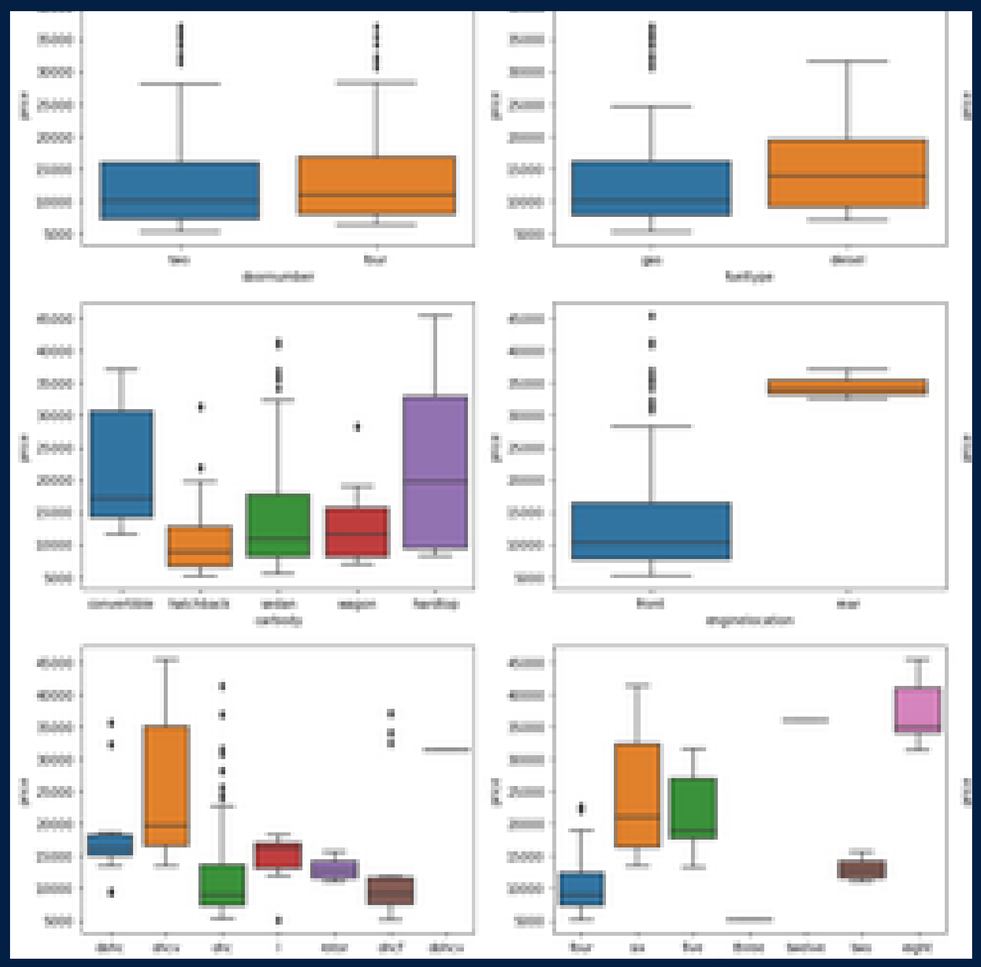
Data preparation is the process of cleaning and transforming raw data prior to processing and analysis. It is an important step prior to processing

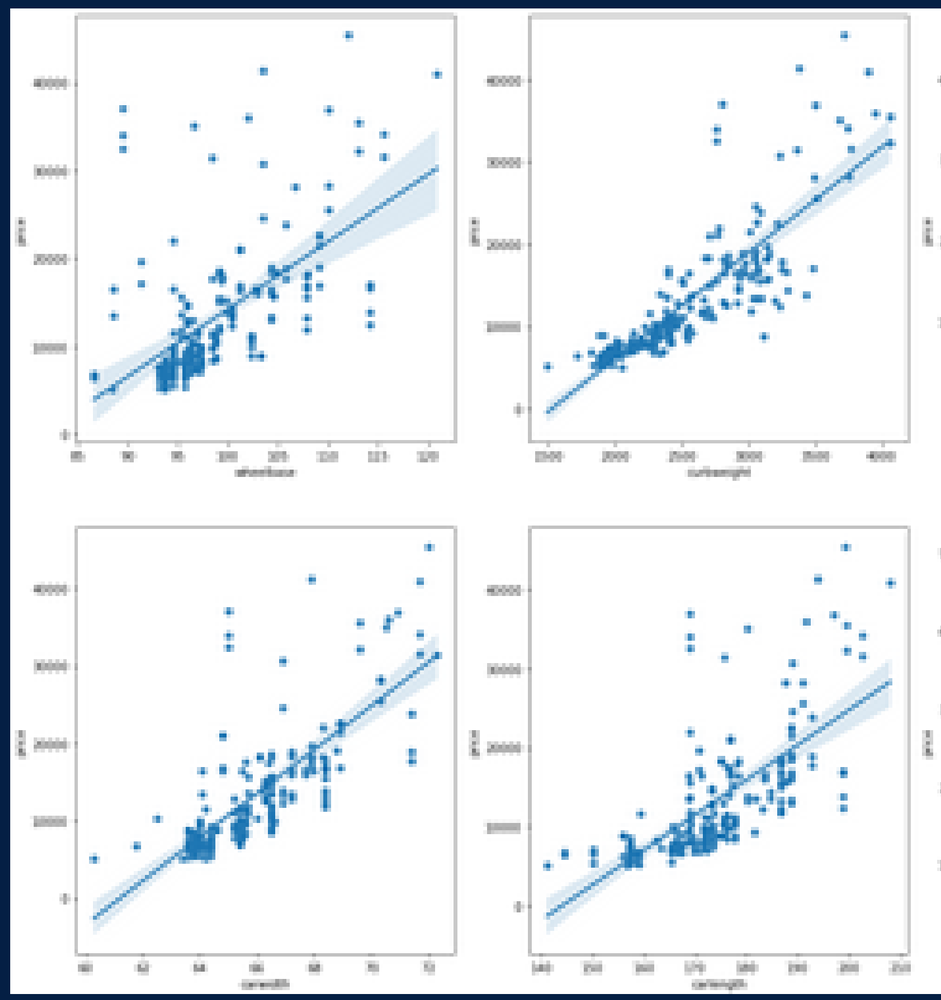
- no null values
- no duplicated data
- making all strings in the same format
- convert important categor data into numerical one

DATA UNDERSTANDING

Data understanding include explore dataandvisualize differant relations in our data.







MODELING

We based on multiple linear regression algorithm to predict car price based on multiple veriables

```
# Splitting the dataset into the Training set and Test set
X=df.drop(["car_ID","CarName","doornumber","drivewheel","enginelocation","fuelsystem","symboling",
           'compressionratio','peakrpm','citympg','highwaympg','carheight','stroke'],axis=1)
y=df['price']
X_train, X_test, y_train, y_test = train_test_split(X,y, test_size = 0.3,shuffle=True,random_state = 8)
```

```
#scaling
sc = StandardScaler()
X_train=sc.fit_transform(X_train)
X_test=sc.fit_transform(X_test)
```

```
model=LinearRegression()
model.fit(X_train, y_train)
```

```
y_pred= model.predict(X_test)
pd.DataFrame({'test':y_test,'pred':y_pred}).head()
```

EVALUATION

Model evaluation is the process of using different evaluation metrics to understand a machine learning model's performance

- mean absolute error
- mean squared errror
- SCORE

MAE: 1398.8596117729996
RMSE: 4839041.454967934
Score :0.892043151816244