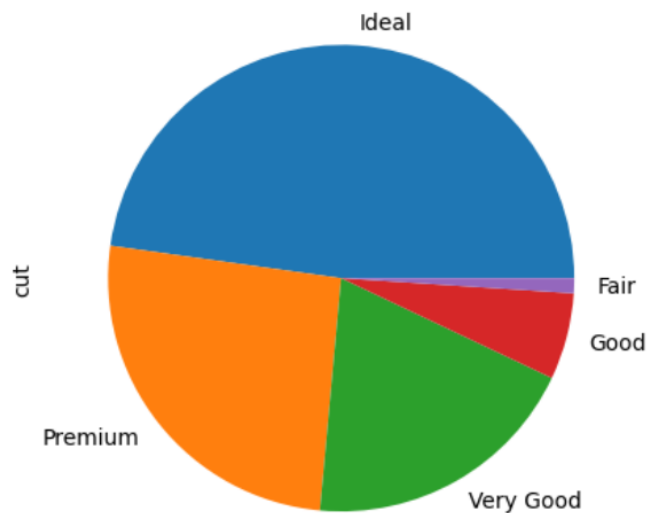
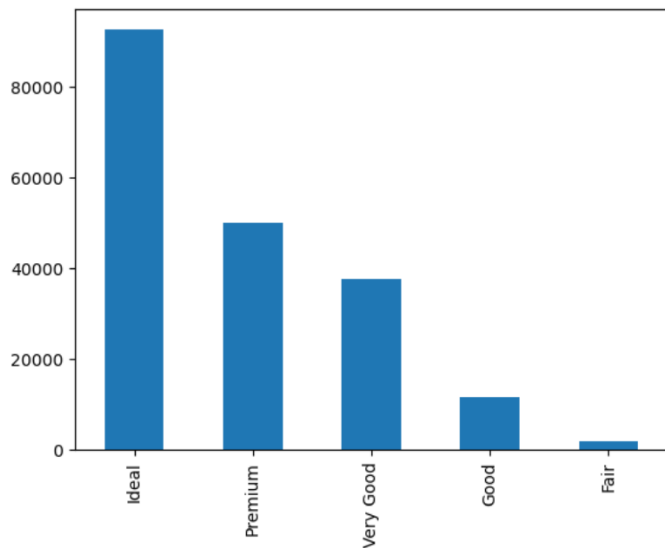


## Exploratory Data Analysis For Diamond Price Prediction:

Exploratory data analysis is an essential step in any data analysis project. EDA helps us understand the dataset, it identifies the patterns and relationships between the columns and gives us insights which helps us in model building.

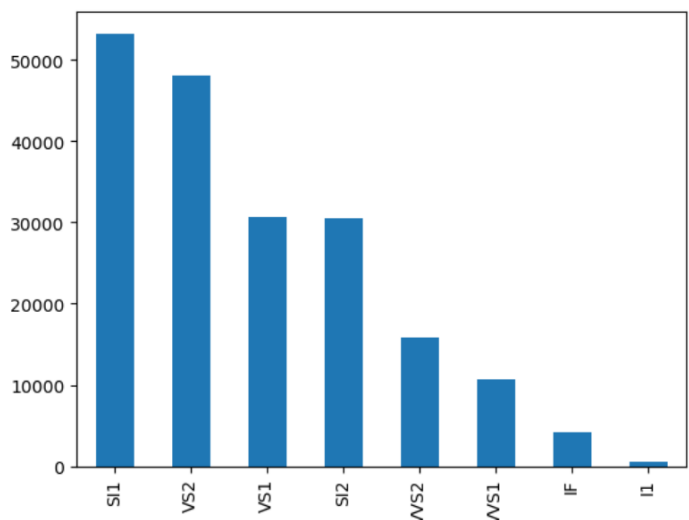
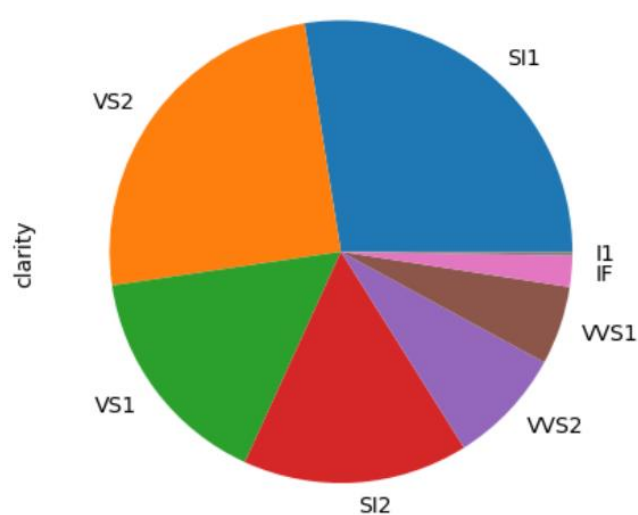
### Uni-Variate plots:

CUT:



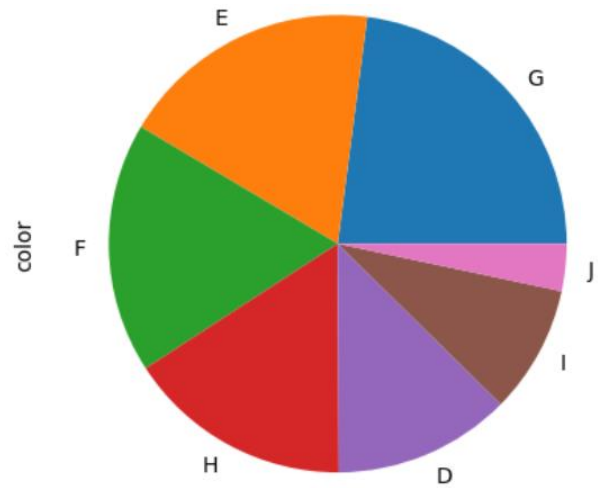
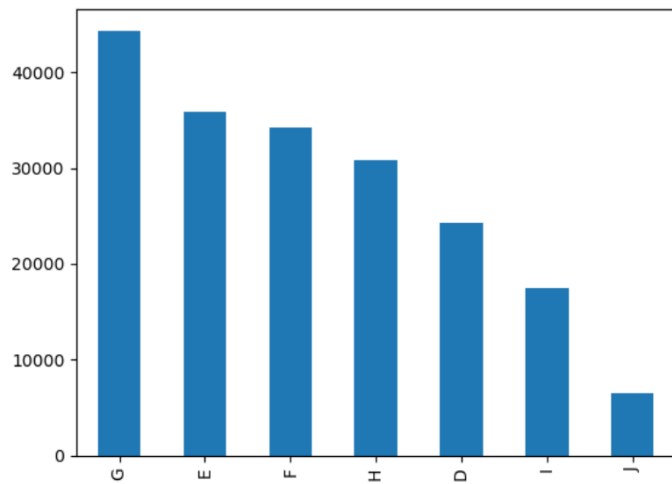
- Most of the diamonds have 'Ideal Cut'

CLARITY:

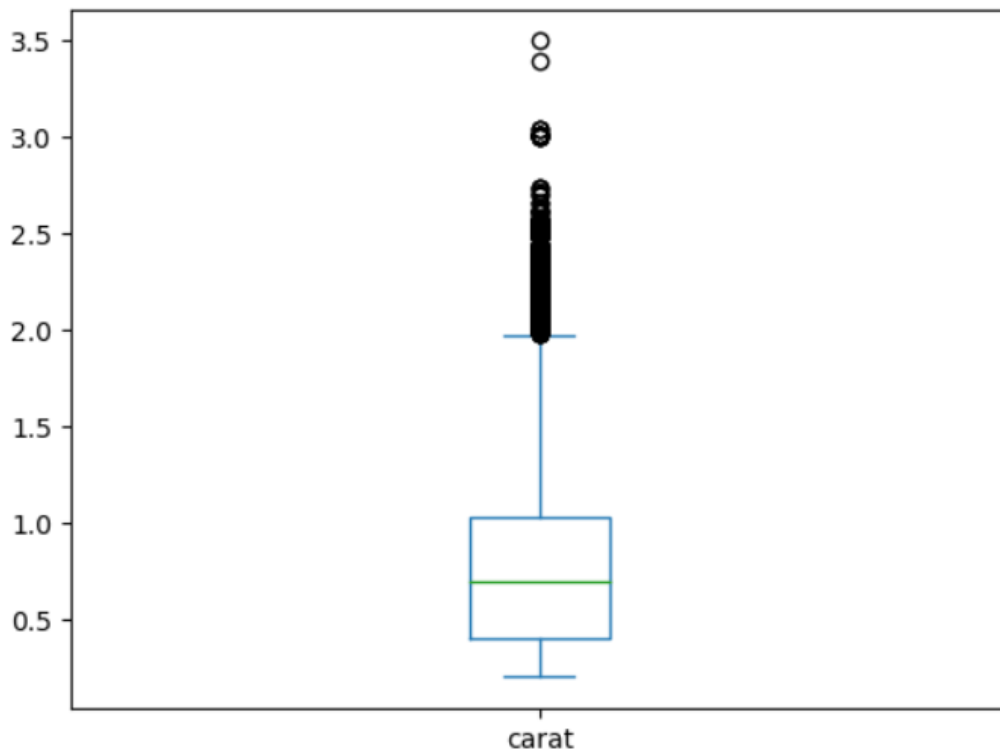


- Most of the diamonds have **SI1** clarity

## COLOR:

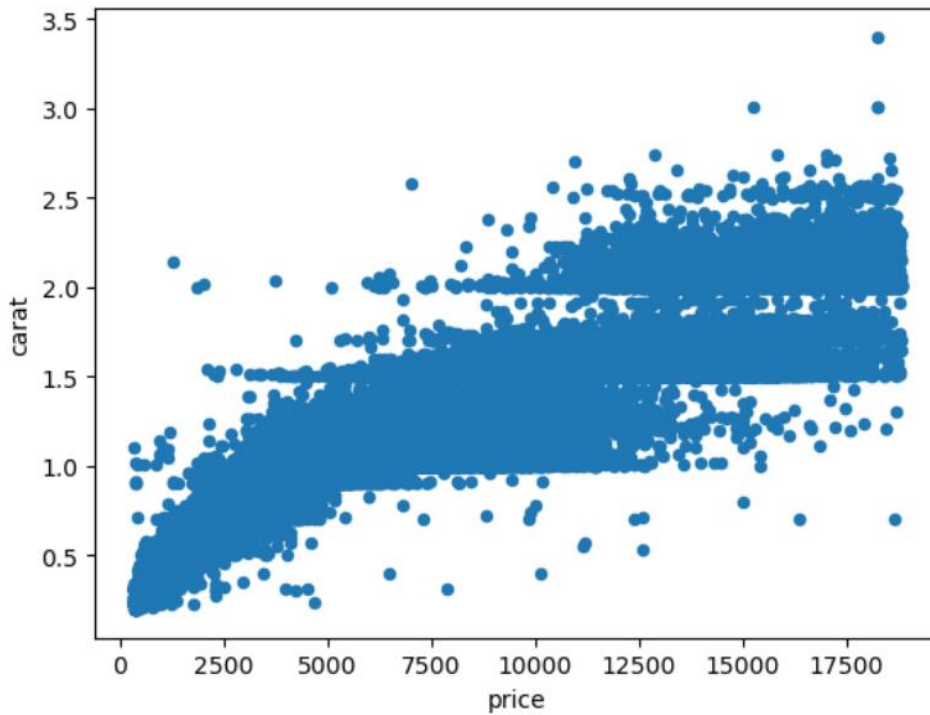


- Most of the diamonds have **G** color followed by **E** and **F**.

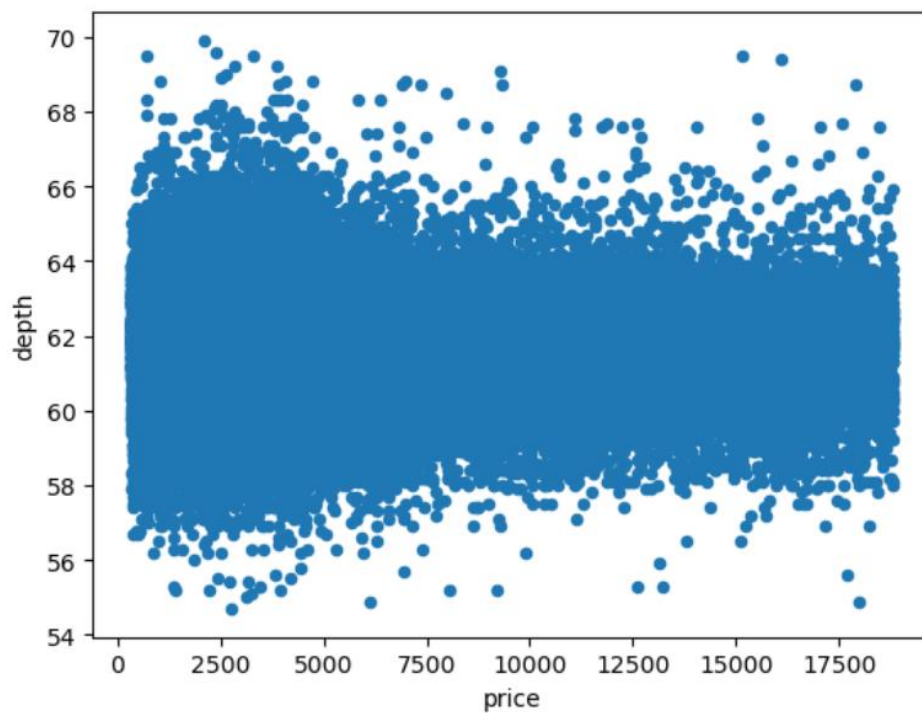


We can see the weight distribution is right skewed.

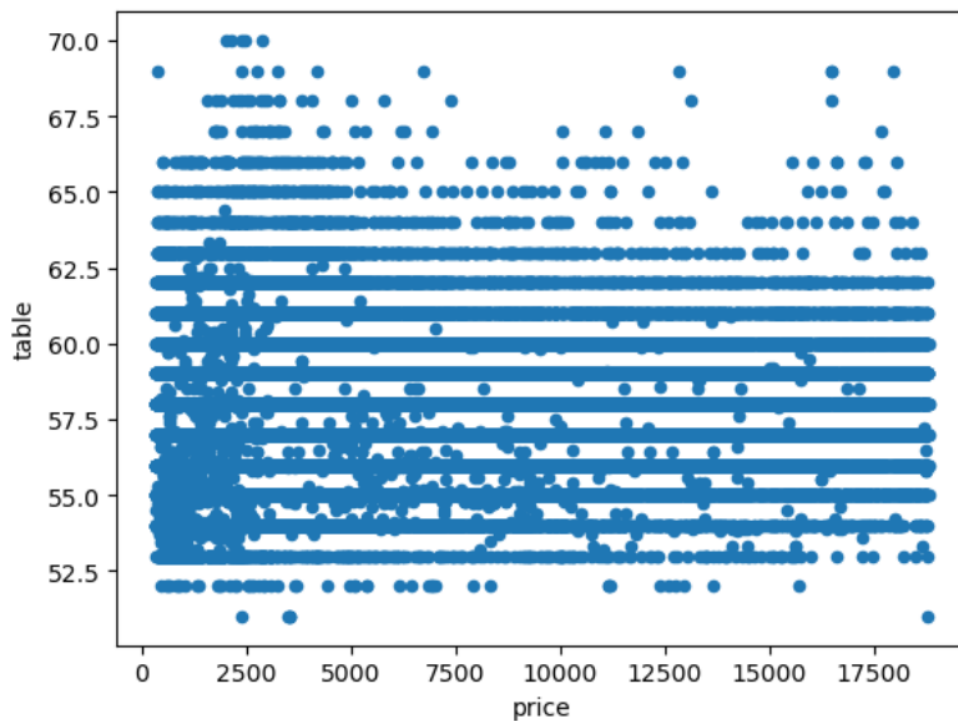
## Bi-Variate plots:



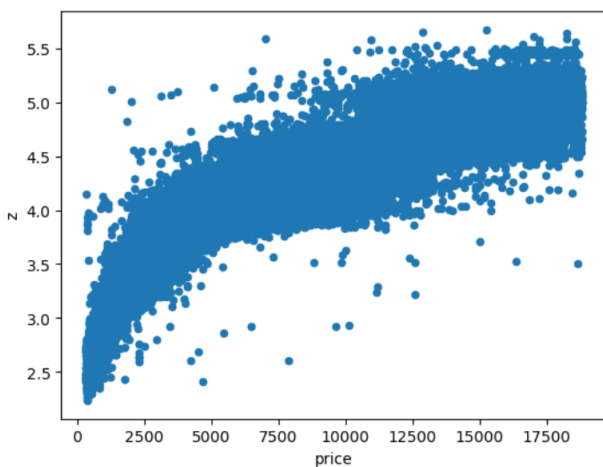
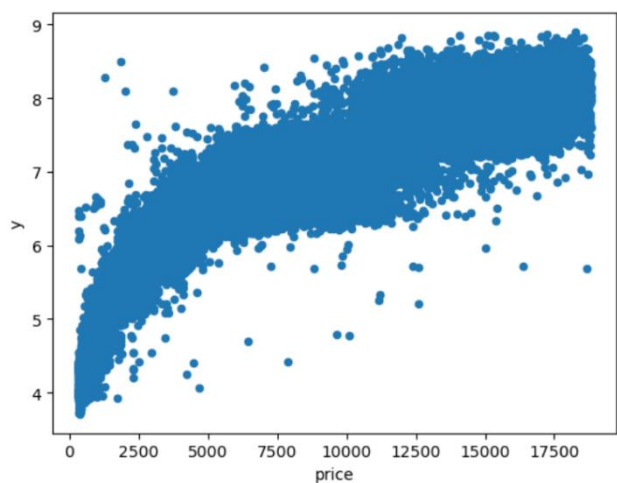
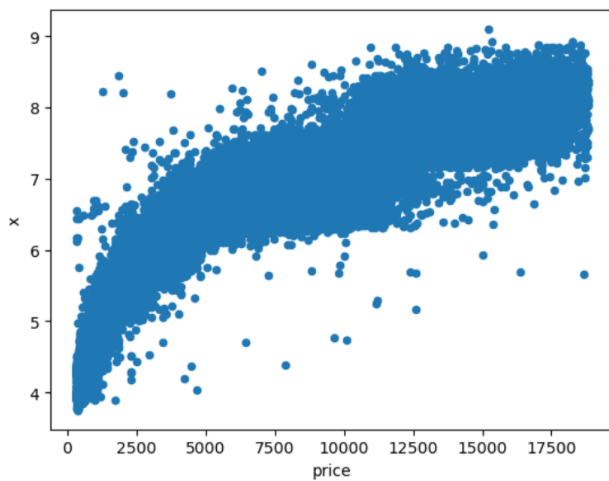
- As the carat increases, the price also increases. We can say that carat is directly proportional to price.



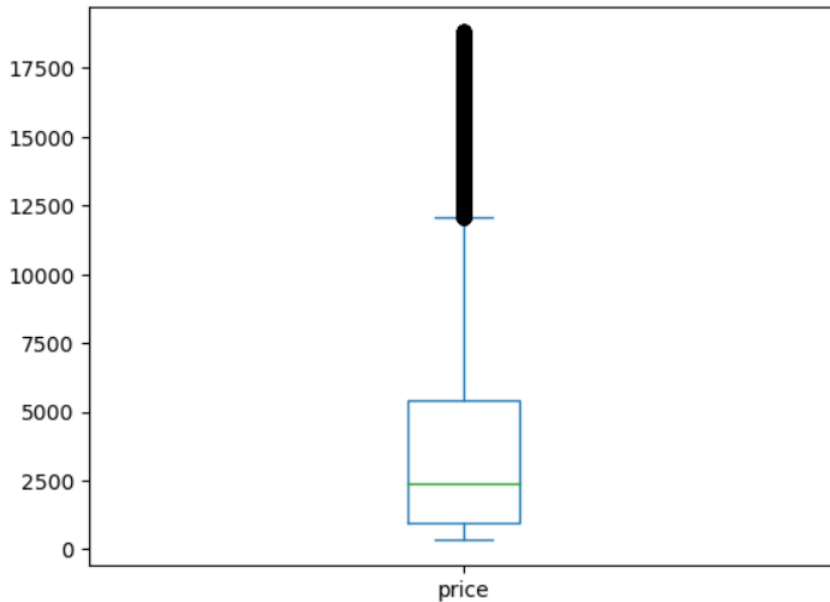
- Depth shows low correlation with the price.



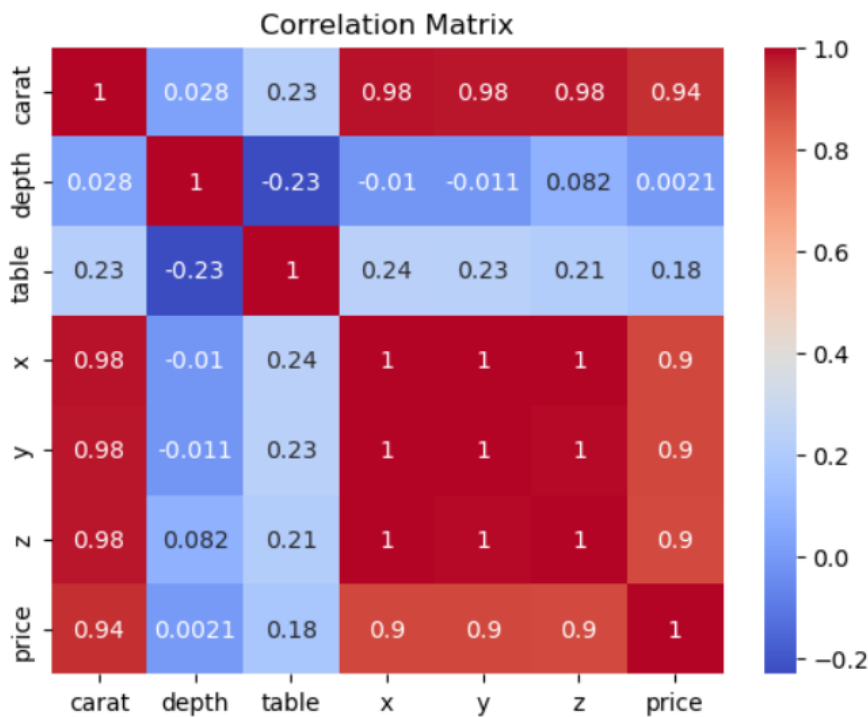
- Table shows low correlation with the price.



"x", "y", "z" and "carat" show a high correlation to the target column. while "depth" and "table" show low correlation.



We can see the price distribution is right skewed.



## **INSIGHTS:**

- High correlation between price and carat. Also high correlation between price and length(x), price and width(y), price and depth(z).
- High correlation between carat and length(x), carat and width(y), carat and depth(z).
- length(x) is fully correlated with width(y).
- Length(x) is almost fully correlated with depth(z) as like width(y) and depth(z)