**PRACTICAL:1**

**AIM:**

Perform the following using Python Pandas and Matplotlib library on given dataset:

1. Deal with missing values in the data either by deleting records or using mean/median/mode imputation.
2. Detect if Outliers exist and Plot the data distribution using Box Plots, Scatter Plots and Histograms of matplotlib library
3. Create and display the correlation matrix of all features of the data. Record and Analyze Observations.

**CODE:**

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

%matplotlib inline

import seaborn as sn

df=pd.read\_csv('fullspecs.csv')

df

new\_df=df.dropna()

new\_df

new\_df.head()

new\_df.tail()

new\_df.replace(to\_replace="Yes",value='1',inplace=True)

new\_df.replace(to\_replace="No",value='0',inplace=True)

new\_df

new\_df.drop(4,inplace=True)

new\_df

new\_df.dtypes

new\_df.rename(columns={'2019 Acura RDX Specs: FWD w/Technology Pkg':'column\_1', '2019 Acura RDX Specs: FWD w/Advance Pkg':'column\_2'},inplace=True)

new\_df = new\_df.replace(r'^\s\*$', np.nan, regex=True)

new\_df=new\_df.astype({'column\_1':float,'column\_2':float},copy=True)

new\_df.dtypes

new\_df

new\_df.notnull()

new\_df

plt.hist(new\_df['column\_1'],color='blue',bins=6)

plt.scatter(new\_df['column\_1'],new\_df['column\_2'],marker="\*")

sn.scatterplot(new\_df['column\_1'],new\_df['column\_2'],marker="\*")

df.info()

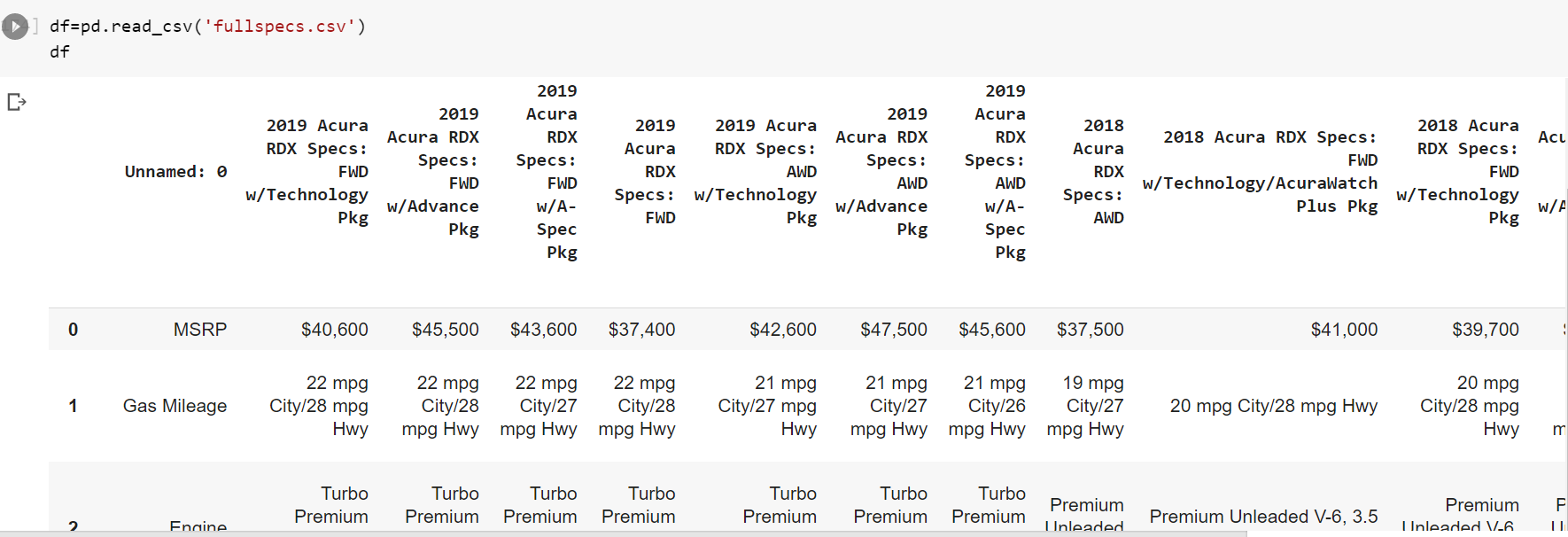
sn.boxplot(x='column\_1',data=new\_df)

new\_df.corr()

sn.heatmap(new\_df.corr(),annot=True)

sn.heatmap(new\_df.corr(),annot=False)

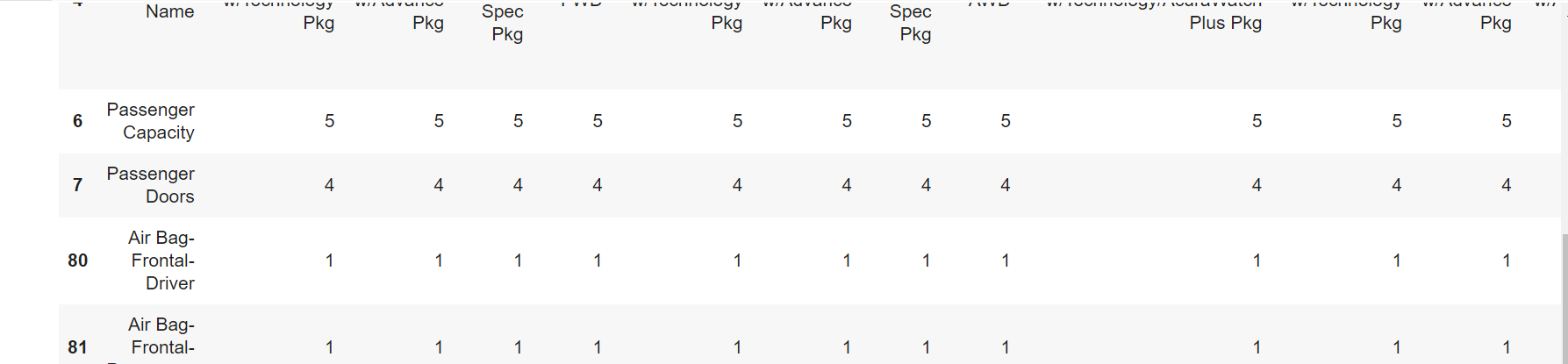
**OUTPUT:**



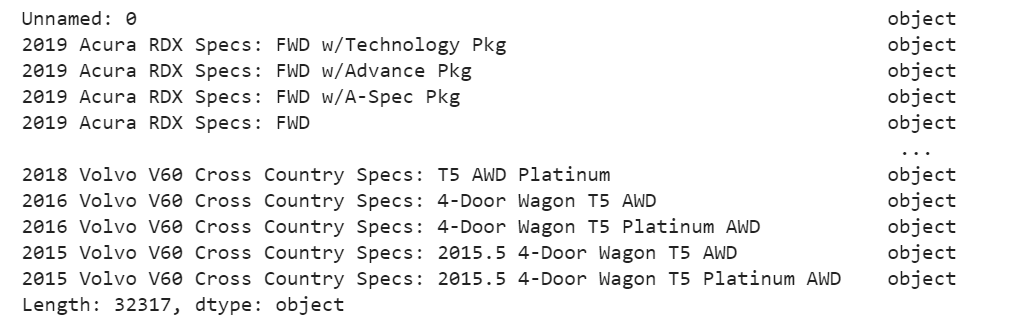
*Glimpses of df*



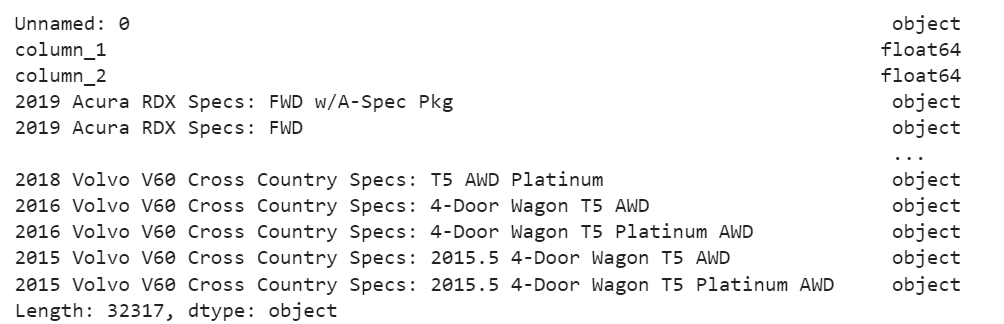
*Dataset after dropping nan values*



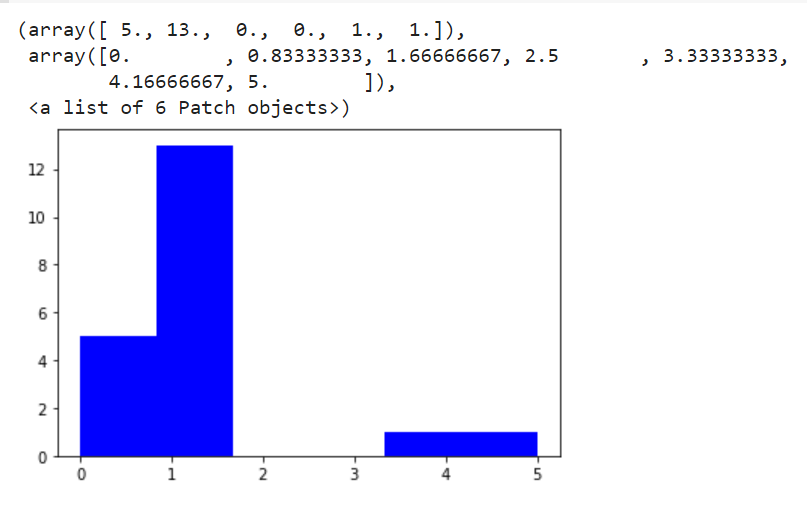
*Dataset after replacing yes/no with 1/0*



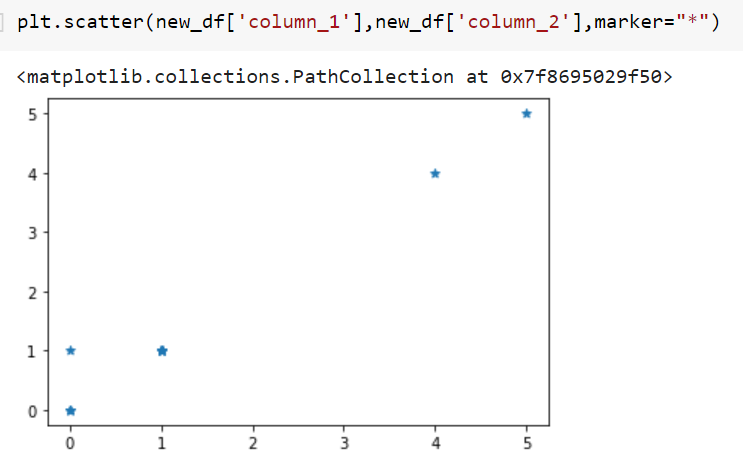
*Showcasing the use of dtypes*



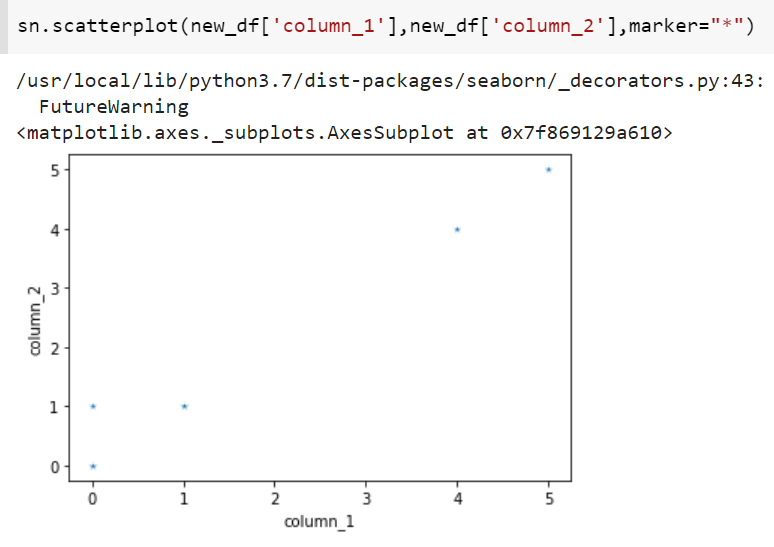
*Replacing the name and datatype of some data attributes*



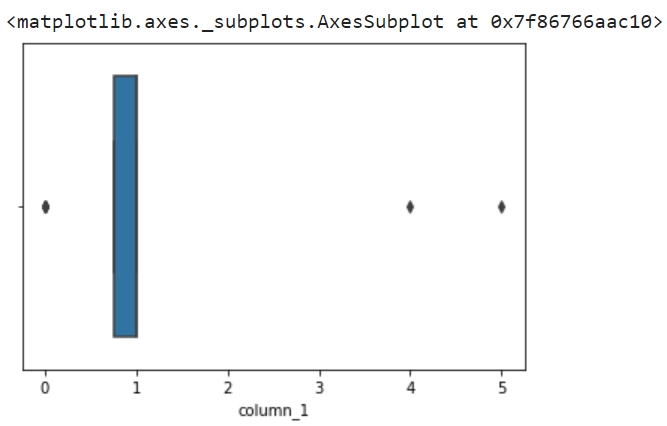
*Histogram*



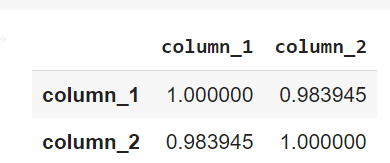
*Scatterplot using matplotlib*



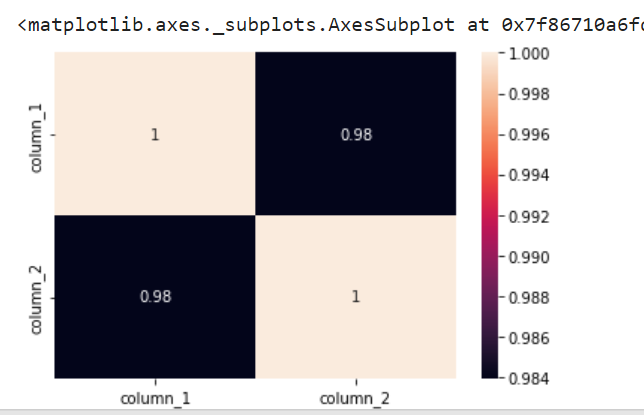
*Scatterplot using seaborn*



*Boxplot*



*Output of corr() function*



*Heatmap*