Name : Devaneshan.V

Roll No : 730919104027

1.Loading Dataset into tool

```
from google.colab import files

uploaded = files.upload()
```

Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.

Saving abalone.csv to abalone.csv

```
import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

import warnings

warnings.filterwarnings('ignore')

data = pd.read_csv("abalone.csv")
```

2.Performing Visualization

Univariate Analysis

```
data.head()
```

	Sex	Length	Diameter	Height	Whole weight	Shucked weight	Viscera weight	Shell weight	Rings
0	M	0.455	0.365	0.095	0.5140	0.2245	0.1010	0.150	15
1	M	0.350	0.265	0.090	0.2255	0.0995	0.0485	0.070	7

2 F 0.530 0.420 0.135 0.6770 0.2565 0.1415 0.210 9

3 M 0.440 0.365 0.125 0.5160 0.2155 0.1140 0.155 10

4 I 0.330 0.255 0.080 0.2050 0.0895 0.0395 0.055 7

```
sns.boxplot(data['Diameter'])
```

```
plt.hist(data['Diameter'])
```

```
(array([ 13.,  66.,  180.,  344.,  513.,  812., 1017.,  934.,  275.,
         23.]),
 array([0.055 , 0.1145, 0.174 , 0.2335, 0.293 , 0.3525, 0.412 , 0.4715,
        0.531 , 0.5905, 0.65  ]),
 )
```

```
plt.plot(data['Diameter'].head(10))
```

```
[]
```

```
plt.pie(data['Diameter'].head(),autopct='%.3f')
```

```
([,
  ,
  ,
  ,
  ],
 [Text(0.8507215626110557, 0.6973326486753676, ''),
  Text(-0.32611344931648134, 1.0505474849691026, ''),
  Text(-1.0998053664078908, -0.02069193128747144, ''),
  Text(-0.08269436219656089, -1.096887251480709, ''),
```

```
 Text(0.9758446362287218, -0.5076684409569241, '')],
[Text(0.46402994324239394, 0.3803632629138369, '21.856'),
 Text(-0.17788006326353525, 0.5730259008922377, '15.868'),
 Text(-0.5998938362224858, -0.011286507974984419, '25.150'),
 Text(-0.045106015743578656, -0.5983021371712958, '21.856'),
 Text(0.5322788924883937, -0.2769100587037768, '15.269')])
```

```
sns.distplot(data['Diameter'].head(300))
```

```
plt.scatter(data['Diameter'].head(400),data['Length'].head(400))
```

```
plt.bar(data['Sex'].head(20),data['Rings'].head(20))
plt.title('Bar plot')
plt.xlabel('Diameter')
plt.ylabel('Rings')
Text(0, 0.5, 'Rings')
```

```
sns.barplot(data['Sex'], data['Rings'])
```

```
sns.jointplot(data['Diameter'].head(50),data['Rings'].head(100))
```

```
sns.barplot('Diameter','Rings',hue='Sex',data=data.head())
```

```
sns.lineplot(data['Diameter'].head(),data['Rings'].head())
```

```
sns.boxplot(data['Sex'].head(10),data['Diameter'].head(10),data['Rings'].head(10))
```

```
fig=plt.figure(figsize=(8,5))
sns.heatmap(data.head().corr(),annot=True)
```

```
sns.pairplot(data.head(),hue='Height')
```

```
sns.pairplot(data.head())
```

3.Perform Descriptive Statistics on the dataset

```
data.head()
```

Sex	Length	Diameter	Height	Whole weight	Shucked weight	Viscera weight	Shell weight	Rings
0	M	0.455	0.365	0.095	0.5140	0.2245	0.1010	0.150	15
1	M	0.350	0.265	0.090	0.2255	0.0995	0.0485	0.070	7
2	F	0.530	0.420	0.135	0.6770	0.2565	0.1415	0.210	9
3	M	0.440	0.365	0.125	0.5160	0.2155	0.1140	0.155	10
4	I	0.330	0.255	0.080	0.2050	0.0895	0.0395	0.055	7

```
data.tail()
```

Sex	Length	Diameter	Height	Whole weight	Shucked weight	Viscera weight	Shell weight	Rings
4172	F	0.565	0.450	0.165	0.8870	0.3700	0.2390	0.2490	11
4173	M	0.590	0.440	0.135	0.9660	0.4390	0.2145	0.2605	10
4174	M	0.600	0.475	0.205	1.1760	0.5255	0.2875	0.3080	9
4175	F	0.625	0.485	0.150	1.0945	0.5310	0.2610	0.2960	10

4176 M 0.710 0.555 0.195 1.9485 0.9455 0.3765 0.4950 12

data.info()

RangeIndex: 4177 entries, 0 to 4176

Data columns (total 9 columns):

#	Column	Non-Null Count	Dtype
0	Sex	4177 non-null	object
1	Length	4177 non-null	float64
2	Diameter	4177 non-null	float64
3	Height	4177 non-null	float64
4	Whole weight	4177 non-null	float64
5	Shucked weight	4177 non-null	float64
6	Viscera weight	4177 non-null	float64
7	Shell weight	4177 non-null	float64
8	Rings	4177 non-null	int64

dtypes: float64(7), int64(1), object(1)

memory usage: 293.8+ KB

data.describe()

	Length	Diameter	Height	Whole weight	Shucked weight	Viscera weight	Shell weight	Rings
count	4177.000000	4177.000000	4177.000000	4177.000000	4177.000000	4177.000000	4177.000000	4177.000000
mean	0.523992	0.407881	0.139516	0.828742	0.359367	0.180594	0.238831	9.933684
std	0.120093	0.099240	0.041827	0.490389	0.221963	0.109614	0.139203	3.224169
min	0.075000	0.055000	0.000000	0.002000	0.001000	0.000500	0.001500	1.000000

25%	0.450000	0.350000	0.115000	0.441500	0.186000	0.093500
	0.130000	8.000000				
50%	0.545000	0.425000	0.140000	0.799500	0.336000	0.171000
	0.234000	9.000000				
75%	0.615000	0.480000	0.165000	1.153000	0.502000	0.253000
	0.329000	11.000000				
max	0.815000	0.650000	1.130000	2.825500	1.488000	0.760000
	1.005000	29.000000				

data.mode().T

0	1	
Sex	M	NaN
Length	0.55	0.625
Diameter	0.45	NaN
Height	0.15	NaN
Whole weight	0.2225	NaN
Shucked weight	0.175	NaN
Viscera weight	0.1715	NaN
Shell weight	0.275	NaN
Rings	9.0	NaN

data.shape

(4177, 9)

data.kurt()

Length	0.064621
Diameter	-0.045476
Height	76.025509
Whole weight	-0.023644
Shucked weight	0.595124

Viscera weight 0.084012

Shell weight 0.531926

Rings 2.330687

dtype: float64

data.skew()

Length -0.639873

Diameter -0.609198

Height 3.128817

Whole weight 0.530959

Shucked weight 0.719098

Viscera weight 0.591852

Shell weight 0.620927

Rings 1.114102

dtype: float64

data.var()

Length 0.014422

Diameter 0.009849

Height 0.001750

Whole weight 0.240481

Shucked weight 0.049268

Viscera weight 0.012015

Shell weight 0.019377

Rings 10.395266

dtype: float64

data.nunique()

Sex 3

Length 134

Diameter 111

Height 51

Whole weight 2429

Shucked weight 1515

Viscera weight 880

Shell weight 926

Rings 28

dtype: int64

4.Check for missing values and deal with them

data.isna()

Sex	Length	Diameter	Height	Whole weight	Shucked weight	Viscera weight	Shell weight	Rings
0	False	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False
...	...	...	...	...	...	...	...	...	...
4172	False	False	False	False	False	False	False	False	False
4173	False	False	False	False	False	False	False	False	False
4174	False	False	False	False	False	False	False	False	False
4175	False	False	False	False	False	False	False	False	False
4176	False	False	False	False	False	False	False	False	False

4177 rows × 9 columns

data.isna().any()

Sex False

Length False

Diameter False

Height False

Whole weight False

Shucked weight False

Viscera weight False

Shell weight False

Rings False

dtype: bool

data.isna().sum()

Sex 0

Length 0

Diameter 0

Height 0

Whole weight 0

Shucked weight 0

Viscera weight 0

Shell weight 0