Assignment 4

Name : Shruthee B

Roll No : 917719C096

1.Loading Dataset into tool

```
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

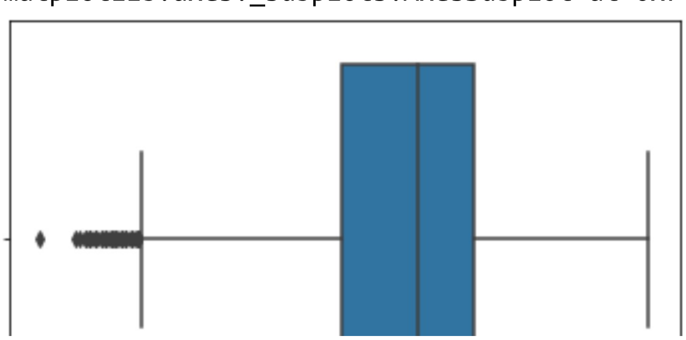
Univariate Analysis

```
data.head()
```

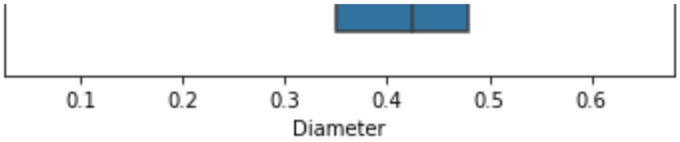
	Sex	Length	Diameter	Height	Whole weight	Shucked weight	Viscera weight	Shell weight	Rings
0	M	0.455	0.365	0.095	0.5140	0.2245	0.1010	0.150	15
1	M	0.350	0.265	0.090	0.2255	0.0995	0.0485	0.070	7
2	F	0.530	0.420	0.135	0.6770	0.2565	0.1415	0.210	9
3	M	0.440	0.365	0.125	0.5160	0.2155	0.1140	0.155	10
4	I	0.330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

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

<matplotlib.axes._subplots.AxesSubplot at 0x7f2c4d947f10>

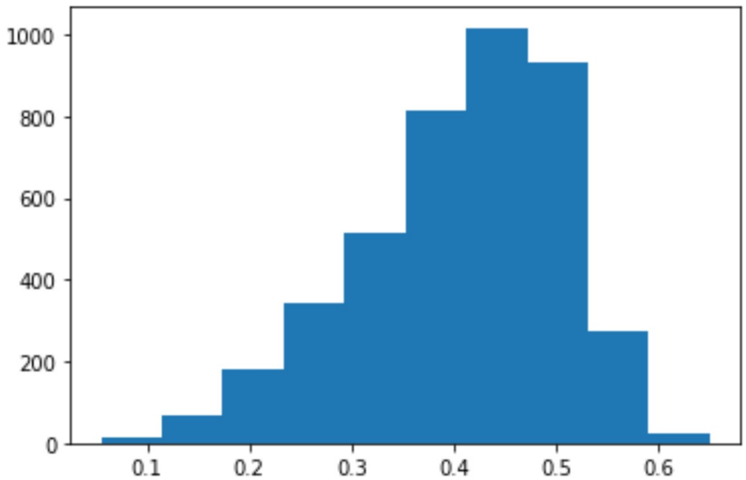


0s completed at 7:26 AM



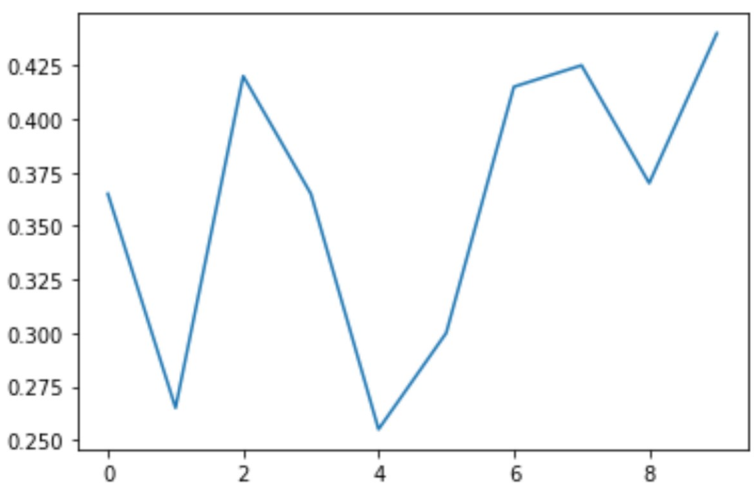

```
1 plt.hist(data['Diameter'])
```

```
(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  ]),
 <a list of 10 Patch objects>)
```




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

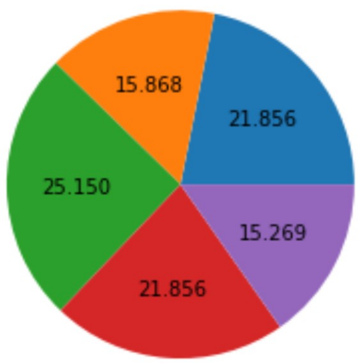
```
[<matplotlib.lines.Line2D at 0x7f2c4d340950>]
```




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

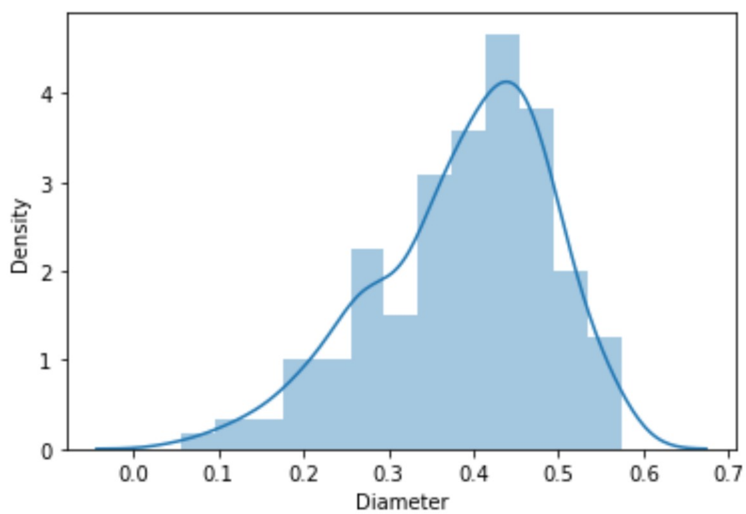
```
([<matplotlib.patches.Wedge at 0x7f2c4d2b7c90>,
  <matplotlib.patches.Wedge at 0x7f2c4d2c3490>,
  <matplotlib.patches.Wedge at 0x7f2c4d2c3d10>,
```

```
 <matplotlib.patches.Wedge at 0x7f2c4d2cd650>,
 <matplotlib.patches.Wedge at 0x7f2c4d2d71d0>],
[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')])
```



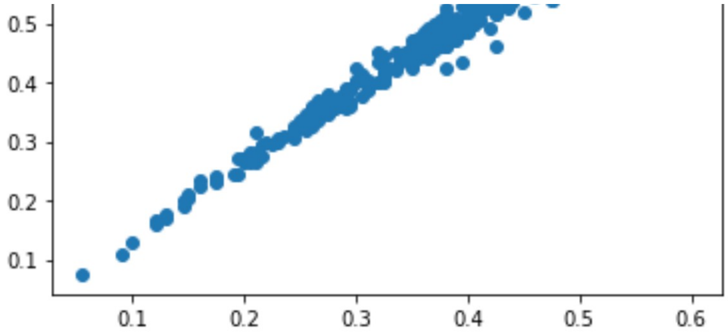

```
1 sns.distplot(data['Diameter'].head(300))
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f2c4d26d5d0>
```



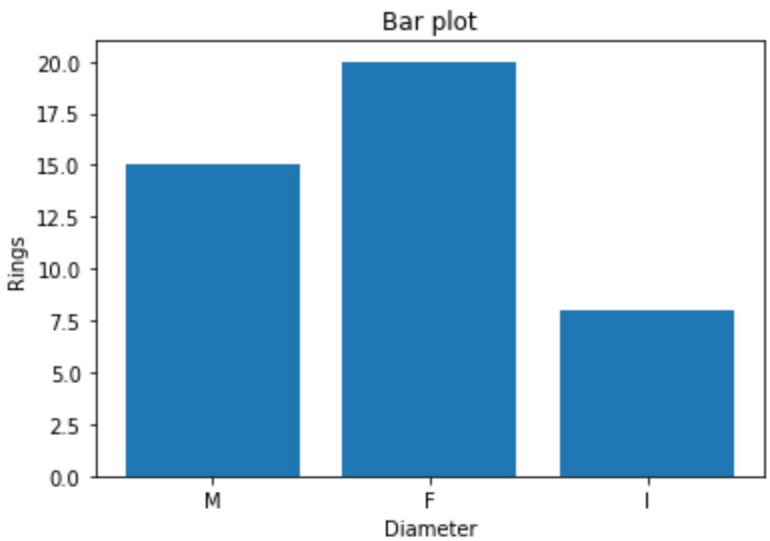

```
1 plt.scatter(data['Diameter'].head(400),data['Length'].head(400))
```

```
<matplotlib.collections.PathCollection at 0x7f2c4d1db910>
```



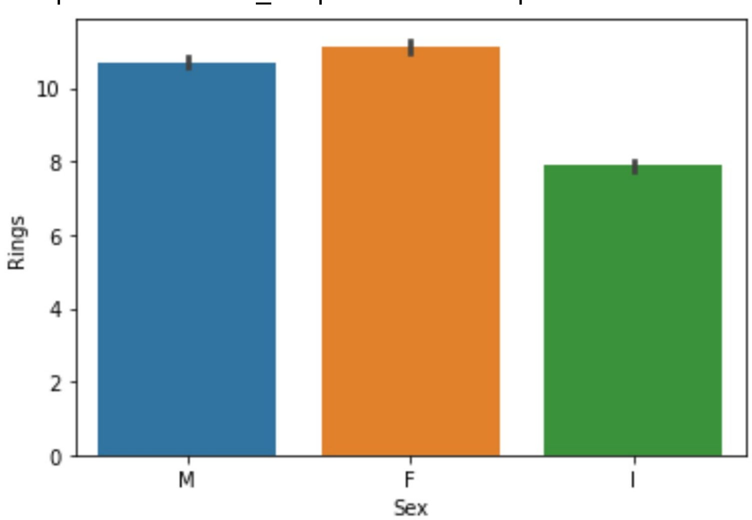

```
1 plt.bar(data['Sex'].head(20),data['Rings'].head(20))
2 plt.title('Bar plot')
3 plt.xlabel('Diameter')
4 plt.ylabel('Rings')
```

```
Text(0, 0.5, 'Rings')
```



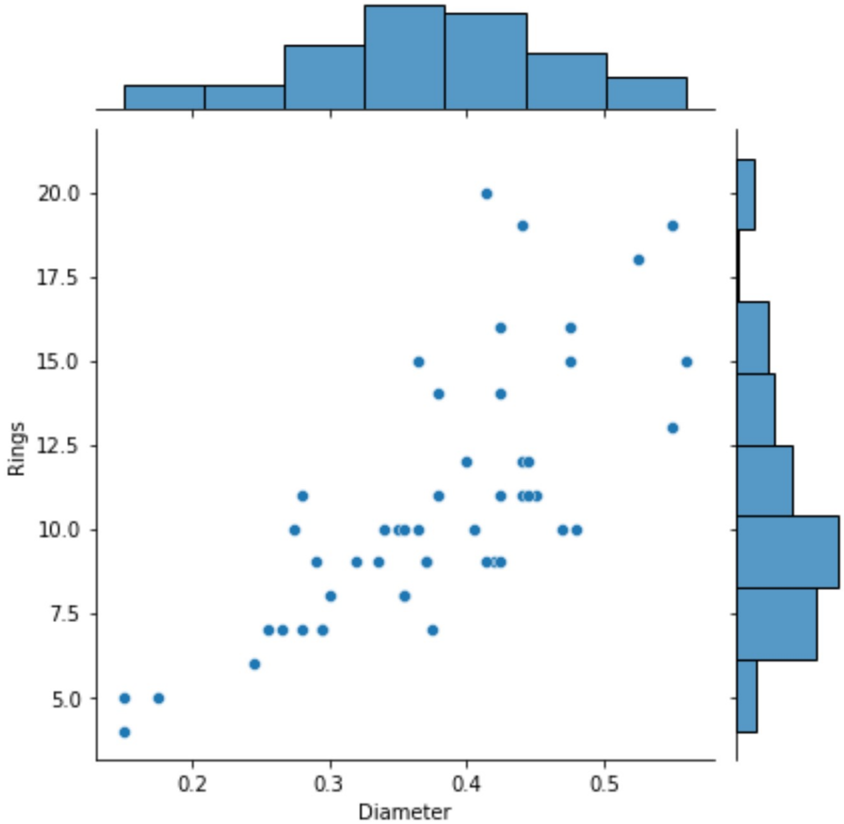

```
1 sns.barplot(data['Sex'], data['Rings'])
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f2c4d0becd0>
```



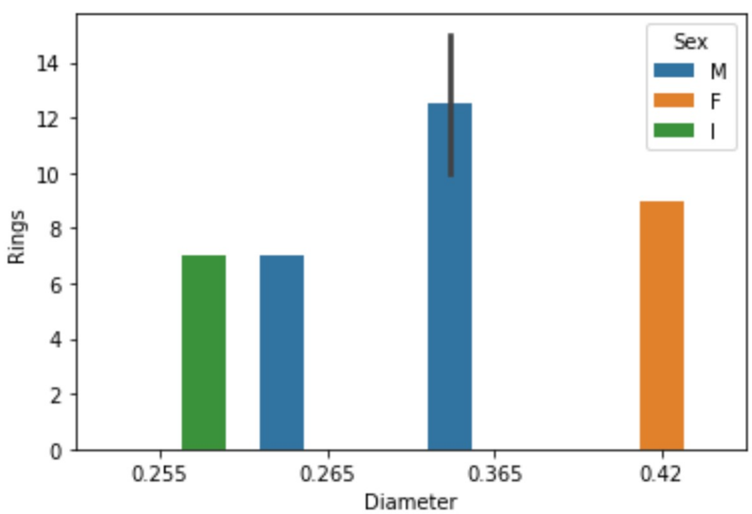
```
1 sns.jointplot(data['Diameter'].head(50),data['Rings'].head(100))
```

```
<seaborn.axisgrid.JointGrid at 0x7f2c4d023950>
```



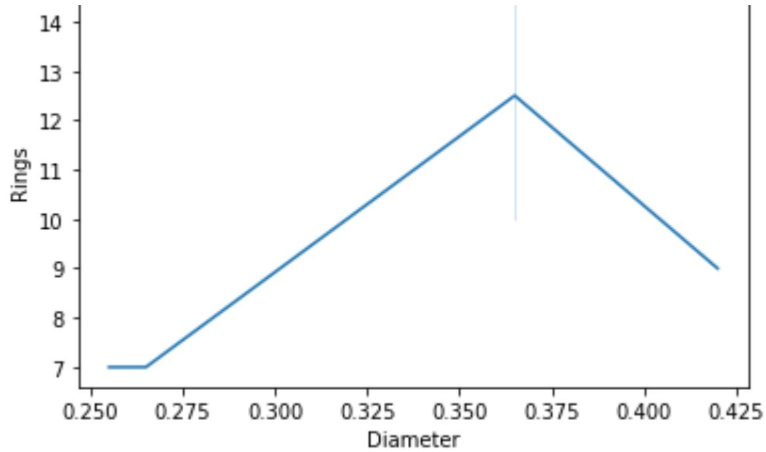

```
1 sns.barplot('Diameter','Rings',hue='Sex',data=data.head())
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f2c4a6a6c10>
```



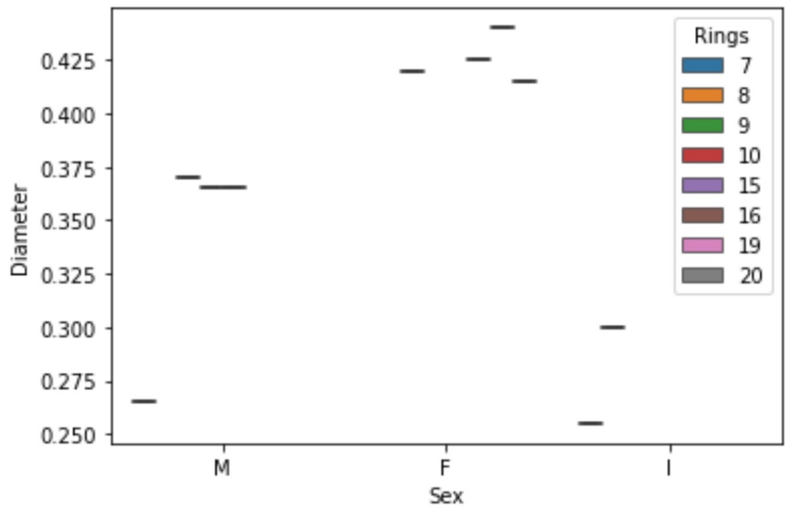

```
1 sns.lineplot(data['Diameter'].head(),data['Rings'].head())
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f2c4a6541d0>
```



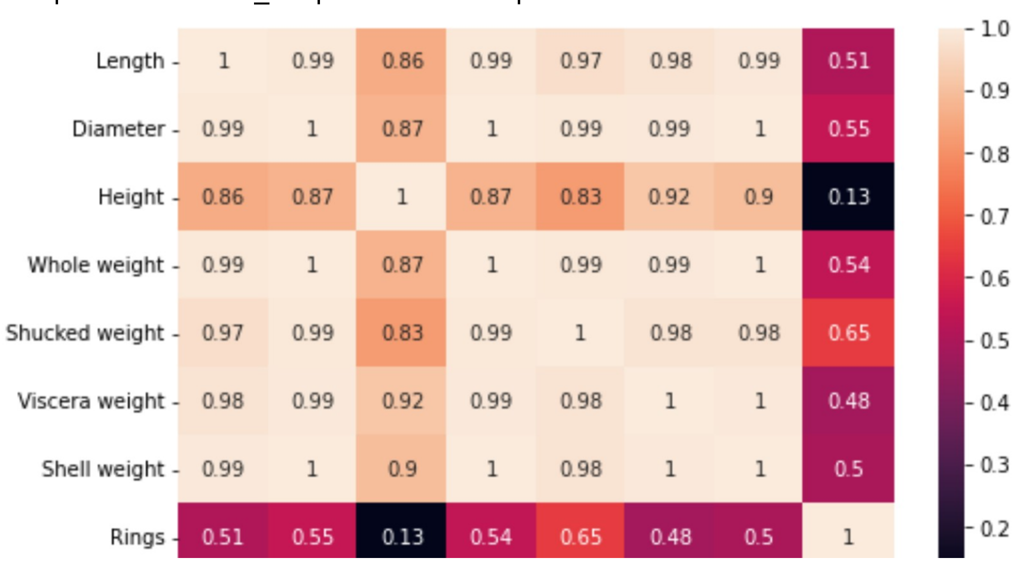

```
1 sns.boxplot(data['Sex'].head(10),data['Diameter'].head(10),data['Rings'].head(10))
```

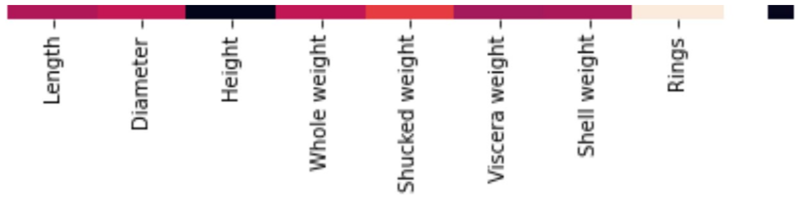
```
<matplotlib.axes._subplots.AxesSubplot at 0x7f2c4a579190>
```




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

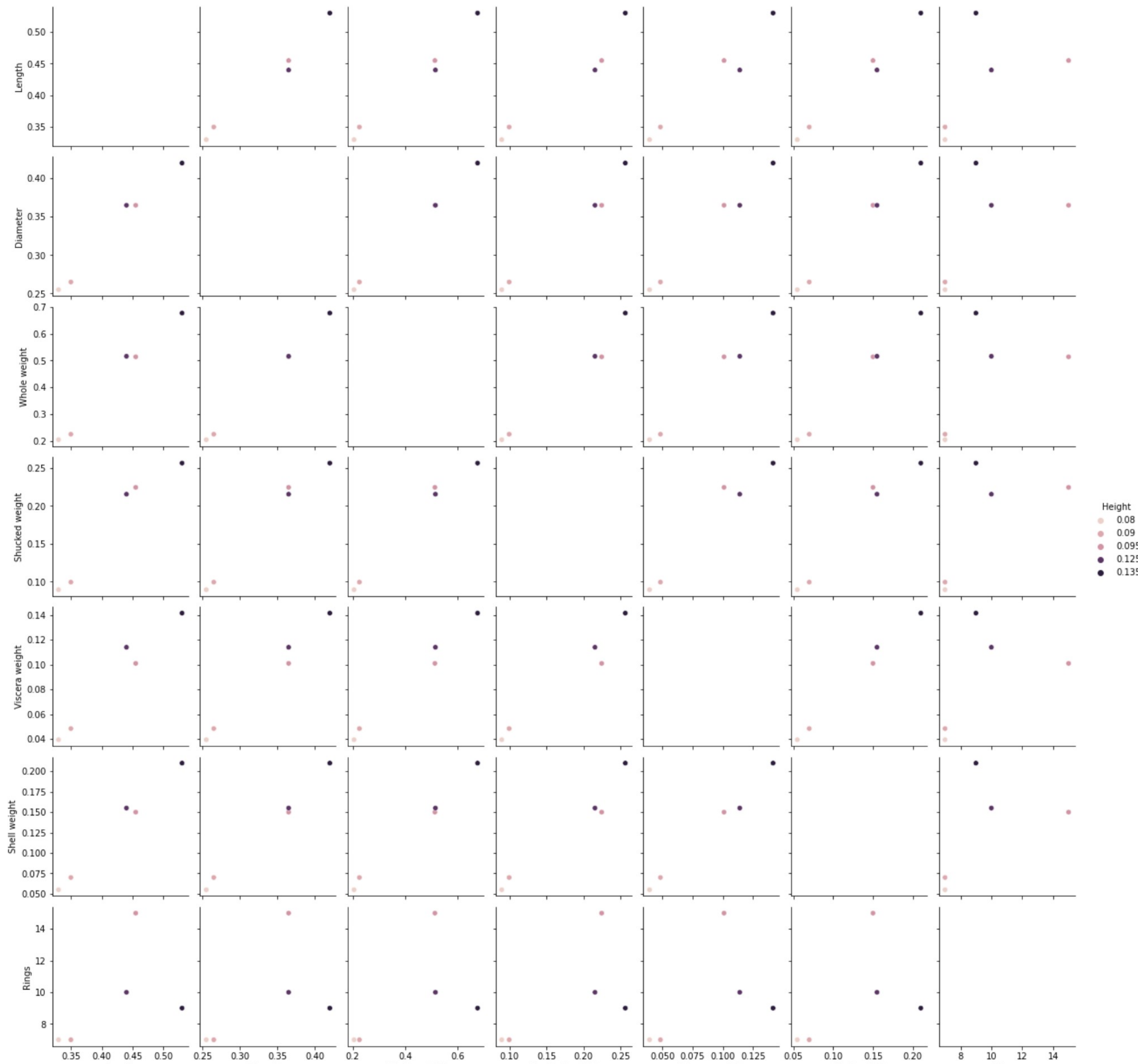
```
<matplotlib.axes._subplots.AxesSubplot at 0x7f2c4a47d250>
```





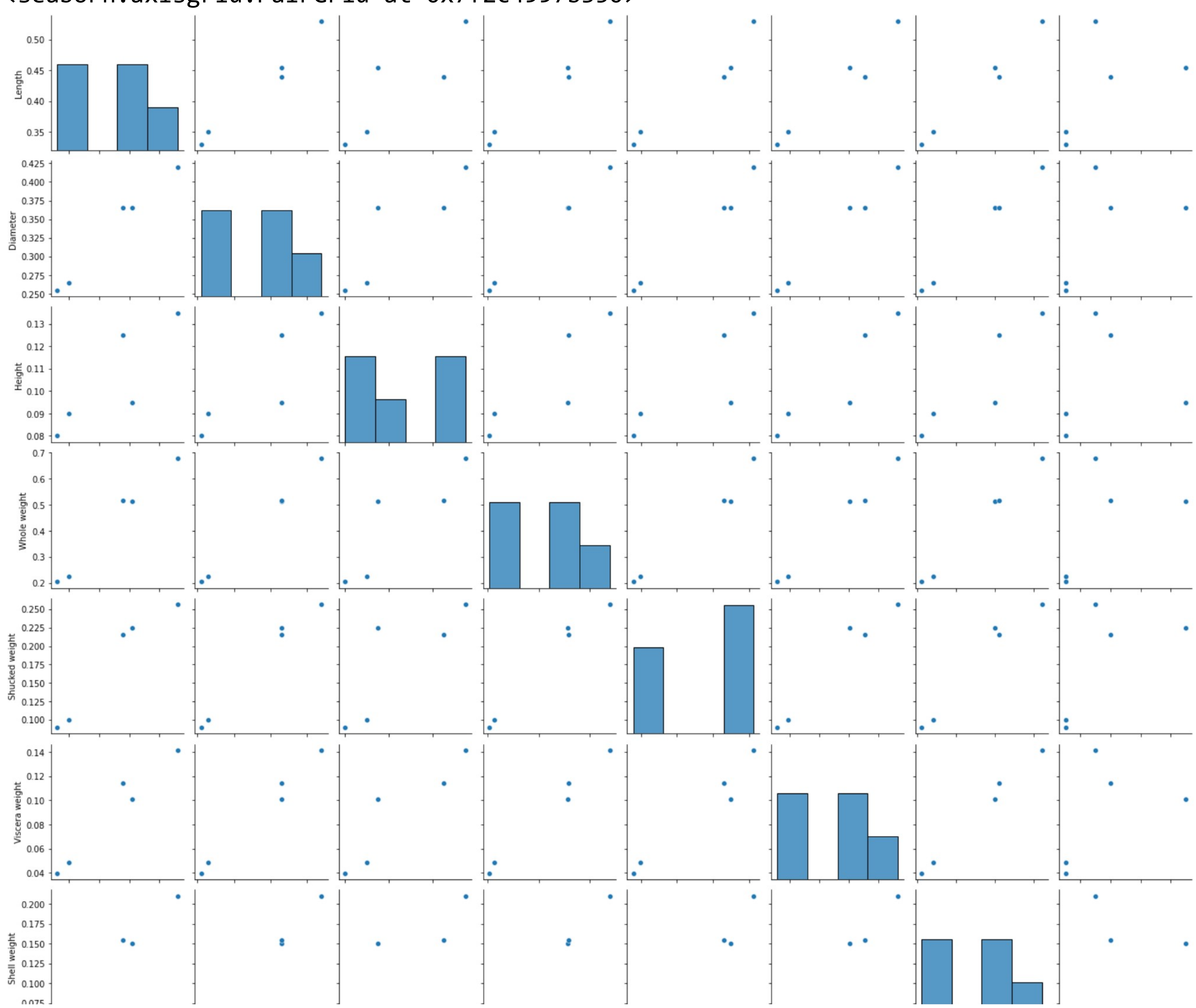

```
1 sns.pairplot(data.head(),hue='Height')
```

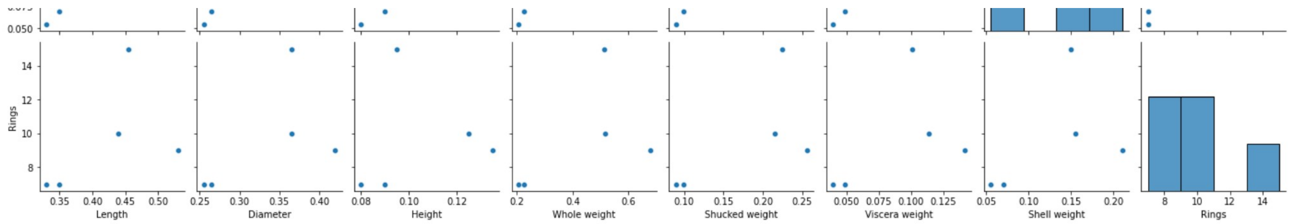
<seaborn.axisgrid.PairGrid at 0x7f2c4a3171d0>



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

<seaborn.axisgrid.PairGrid at 0x7f2c4997b350>





3.Perform Descriptive Statistics on the dataset

```
1 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

```
1 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

```
1 data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
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
```

```
1 data.describe()
```

	Length	Diameter	Height	Whole weight	Shucked weight	Viscera weight	
count	4177.000000	4177.000000	4177.000000	4177.000000	4177.000000	4177.000000	417
mean	0.523992	0.407881	0.139516	0.828742	0.359367	0.180594	
std	0.120093	0.099240	0.041827	0.490389	0.221963	0.109614	
min	0.075000	0.055000	0.000000	0.002000	0.001000	0.000500	
25%	0.450000	0.350000	0.115000	0.441500	0.186000	0.093500	
50%	0.545000	0.425000	0.140000	0.799500	0.336000	0.171000	
75%	0.615000	0.480000	0.165000	1.153000	0.502000	0.253000	
max	0.815000	0.650000	1.130000	2.825500	1.488000	0.760000	

```
1 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

```
1 data.shape
```

```
(4177, 9)
```

```
1 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
```

```
1 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
```

```
1 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
```

```
1 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

```
1 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

```
1 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
```

```
1 data.isna().sum()
```

```
Sex               0
Length            0
Diameter          0
Height            0
Whole weight      0
Shucked weight    0
Viscera weight    0
Shell weight      0
Rings             0
dtype: int64
```

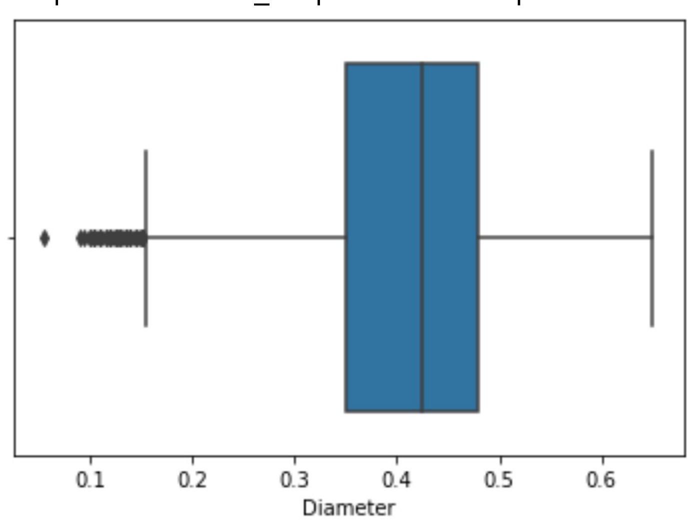
```
1 data.isna().any().sum()
```

```
0
```

5.Find the outliers and replace them outliers

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

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f2c45adead0>
```



```
quant=data.quantile(q=[0.25,0.75])
quant
```

	Length	Diameter	Height	Whole weight	Shucked weight	Viscera weight	Shell weight	Rings
0.25	0.450	0.35	0.115	0.4415	0.186	0.0935	0.130	8.0
0.75	0.615	0.48	0.165	1.1530	0.502	0.2530	0.329	11.0

```
iqr=quant.loc[0.75]-quant.loc[0.25]
iqr
```

```
Length            0.1650
Diameter          0.1300
Height            0.0500
Whole weight      0.7115
Shucked weight    0.3160
Viscera weight    0.1595
Shell weight      0.1990
Rings             3.0000
dtype: float64
```

```
low=quant.loc[0.25]-(1.5*iqr)
low
```

```
Length            0.20250
Diameter          0.15500
Height            0.04000
Whole weight     -0.62575
Shucked weight   -0.28800
Viscera weight   -0.14575
Shell weight     -0.16850
Rings             3.50000
dtype: float64
```

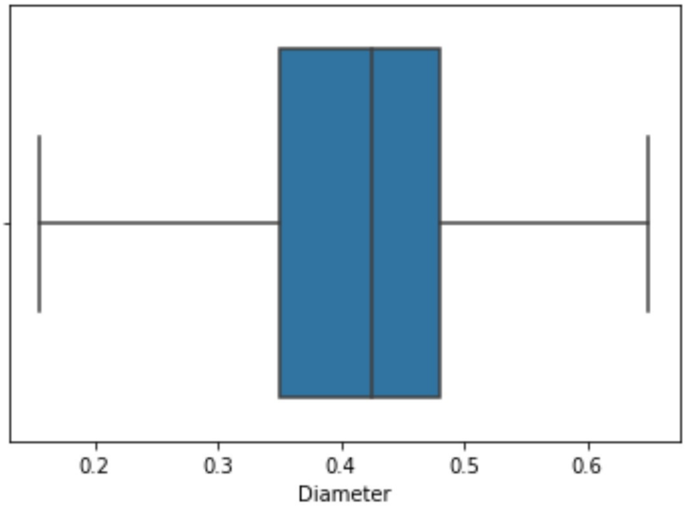
```
up=quant.loc[0.75]+(1.5*iqr)
up
```

```
Length             0.86250
Diameter           0.67500
Height             0.24000
Whole weight       2.22025
Shucked weight     0.97600
Viscera weight     0.49225
Shell weight       0.62750
Rings             15.50000
dtype: float64
```

```
data['Diameter']=np.where(data['Diameter']<0.155,0.4078,data['Diameter'])
```

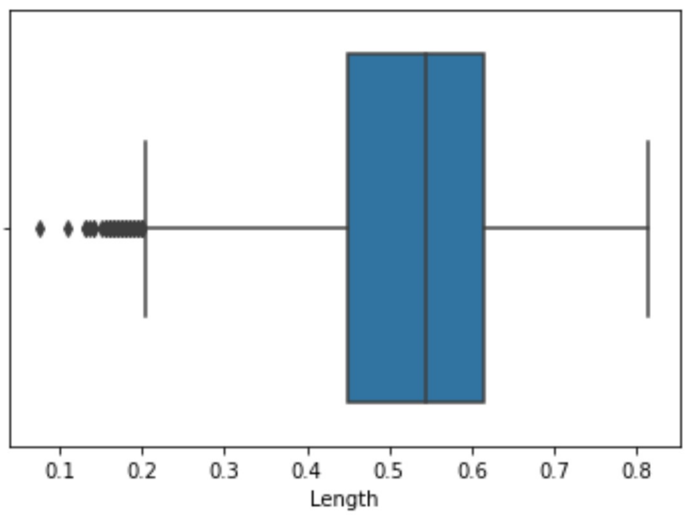
```
2 sns.boxplot(data['Diameter'])
```

<matplotlib.axes._subplots.AxesSubplot at 0x7f2c45abd450>



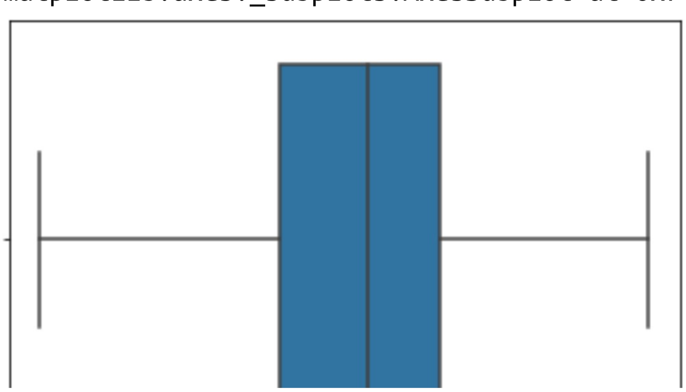

```
1 sns.boxplot(data['Length'])
```

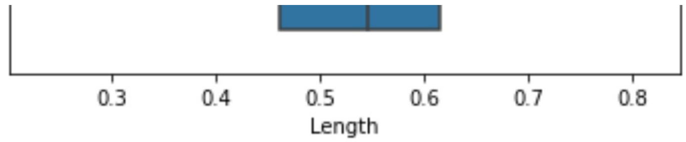
<matplotlib.axes._subplots.AxesSubplot at 0x7f2c45a95690>




```
1 data['Length']=np.where(data['Length']<0.23,0.52, data['Length'])
2 sns.boxplot(data['Length'])
```

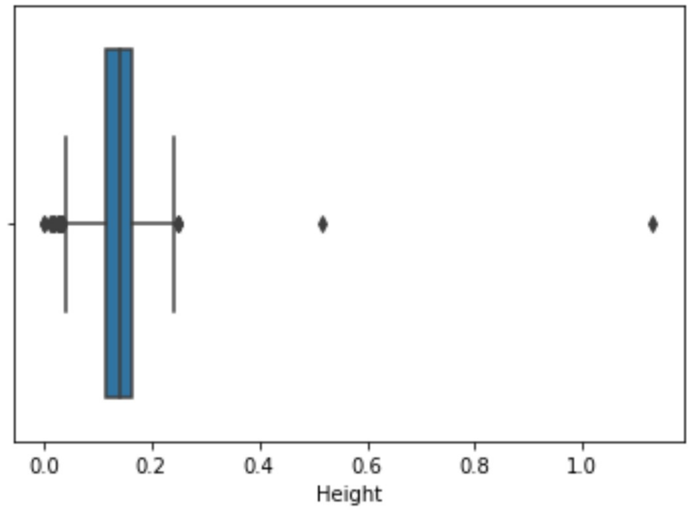
<matplotlib.axes._subplots.AxesSubplot at 0x7f2c459fe4d0>





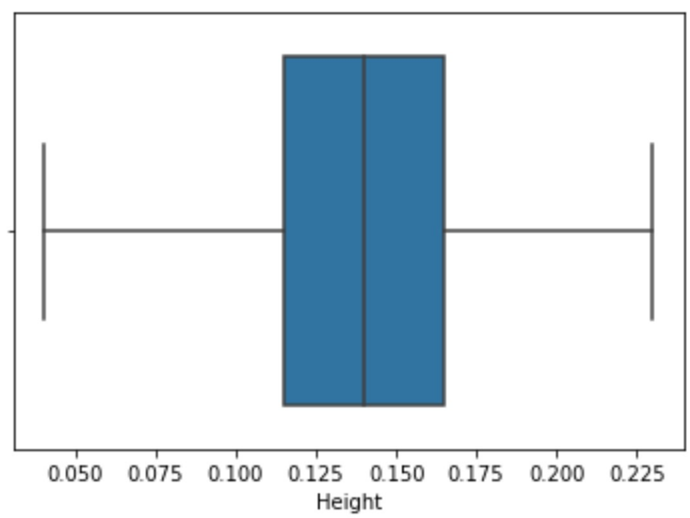

```
sns.boxplot(data['Height'])
```

<matplotlib.axes._subplots.AxesSubplot at 0x7f2c4597fc10>



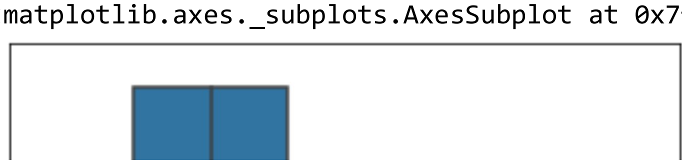

```
data['Height']=np.where(data['Height']<0.04,0.139, data['Height'])
data['Height']=np.where(data['Height']>0.23,0.139, data['Height'])
sns.boxplot(data['Height'])
```

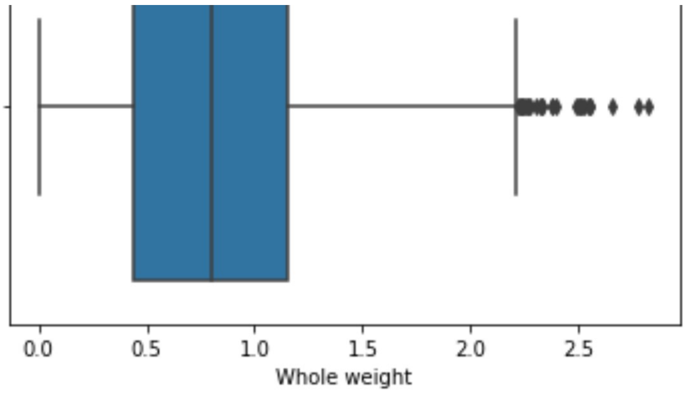
<matplotlib.axes._subplots.AxesSubplot at 0x7f2c458df1d0>




```
sns.boxplot(data['Whole weight'])
```

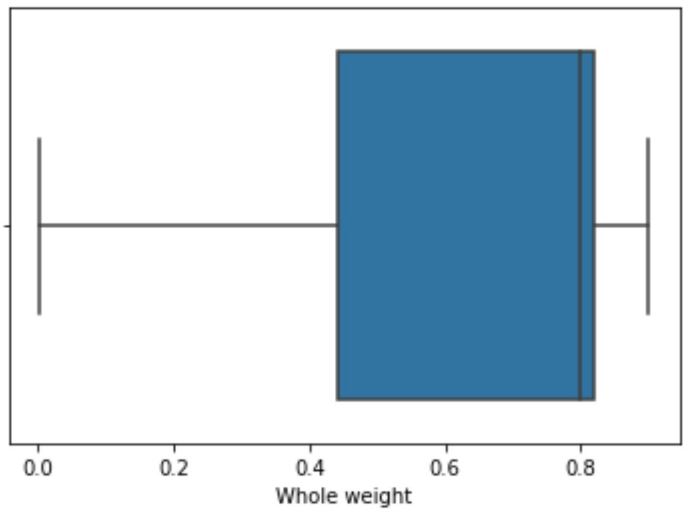
<matplotlib.axes._subplots.AxesSubplot at 0x7f2c458ec3d0>





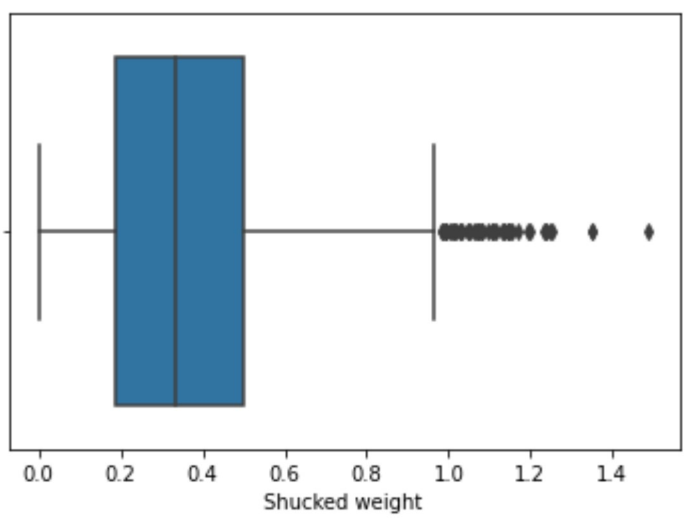

```
1 data['Whole weight']=np.where(data['Whole weight']>0.9,0.82, data['Whole weight'])
2 sns.boxplot(data['Whole weight'])
```

<matplotlib.axes._subplots.AxesSubplot at 0x7f2c45838790>




```
1 sns.boxplot(data['Shucked weight'])
```

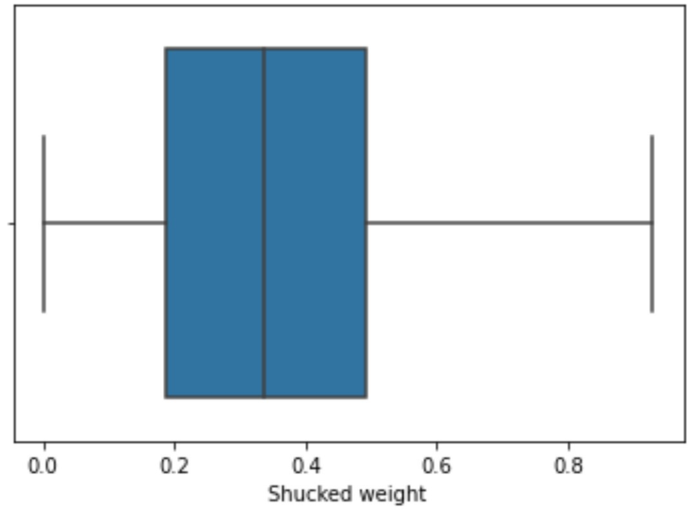
<matplotlib.axes._subplots.AxesSubplot at 0x7f2c457a6650>




```
1 data['Shucked weight']=np.where(data['Shucked weight']>0.93,0.35, data['Shucked weight'
```

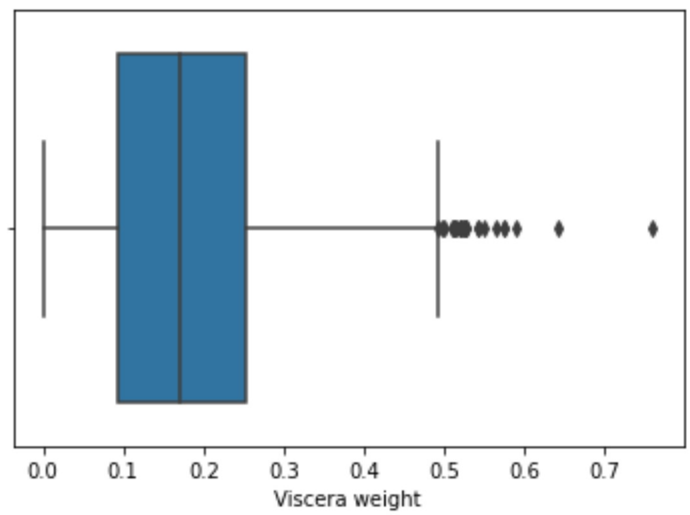
```
1 data['Shucked weight']=np.where(data['Shucked weight']>[illegible],[illegible], data['Shucked weight'
2 sns.boxplot(data['Shucked weight'])
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f2c45792fd0>
```



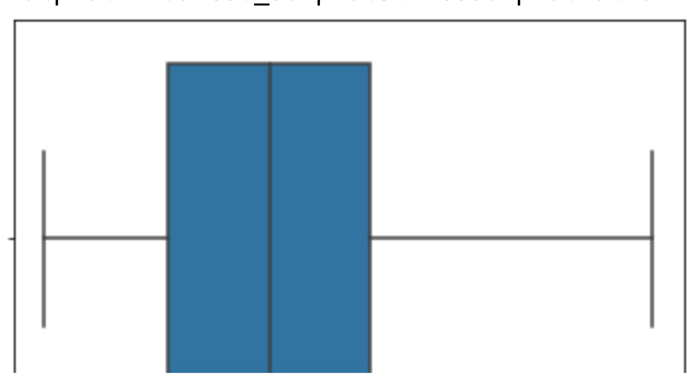
```
1 sns.boxplot(data['Viscera weight'])
```

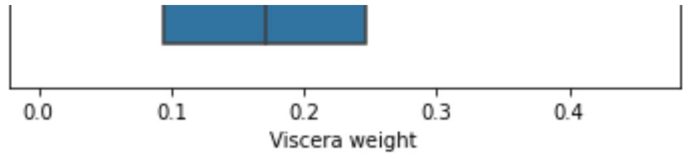
```
<matplotlib.axes._subplots.AxesSubplot at 0x7f2c457062d0>
```



```
1 data['Viscera weight']=np.where(data['Viscera weight']>0.46,0.18, data['Viscera weight']
2 sns.boxplot(data['Viscera weight'])
```

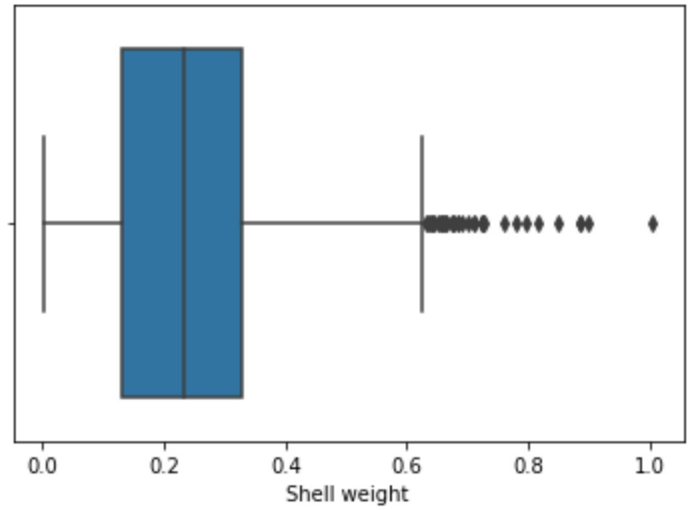
```
<matplotlib.axes._subplots.AxesSubplot at 0x7f2c45679390>
```





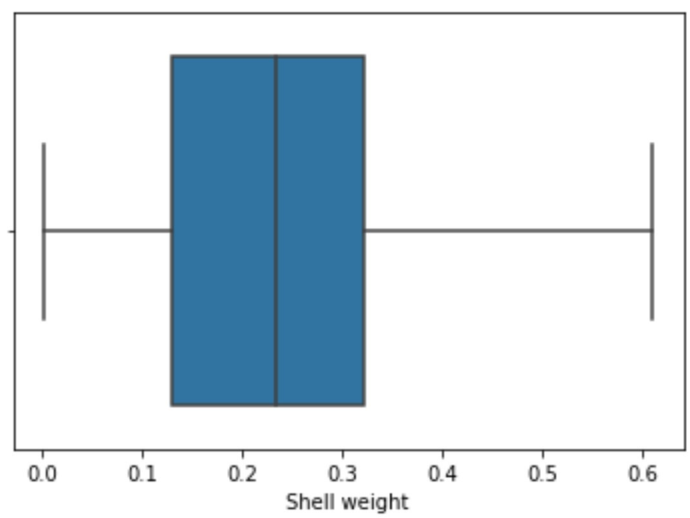

```
sns.boxplot(data['Shell weight'])
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f2c455daa50>
```




```
data['Shell weight']=np.where(data['Shell weight']>0.61,0.2388, data['Shell weight'])
sns.boxplot(data['Shell weight'])
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f2c455c9750>
```



6.Check for Categorical columns and perform encoding.

```
data['Sex'].replace({'M':1,'F':0,'I':2},inplace=True)
data
```

Whole	Shucked	Viscera	Shell

	Sex	Length	Diameter	Height	Whole weight	Shucked weight	Viscera weight	Shell weight	Rings
0	1	0.455	0.365	0.095	0.5140	0.2245	0.1010	0.1500	15
1	1	0.350	0.265	0.090	0.2255	0.0995	0.0485	0.0700	7
2	0	0.530	0.420	0.135	0.6770	0.2565	0.1415	0.2100	9
3	1	0.440	0.365	0.125	0.5160	0.2155	0.1140	0.1550	10
4	2	0.330	0.255	0.080	0.2050	0.0895	0.0395	0.0550	7
...	...	...	...	...	...	...	...	...	...
4172	0	0.565	0.450	0.165	0.8870	0.3700	0.2390	0.2490	11
4173	1	0.590	0.440	0.135	0.8200	0.4390	0.2145	0.2605	10
4174	1	0.600	0.475	0.205	0.8200	0.5255	0.2875	0.3080	9
4175	0	0.625	0.485	0.150	0.8200	0.5310	0.2610	0.2960	10
4176	1	0.710	0.555	0.195	0.8200	0.3500	0.3765	0.4950	12

4177 rows × 9 columns

7.Split the data into dependent and independent variables.

```
x=data.drop(columns= ['Rings'])
y=data['Rings']
x
```

	Sex	Length	Diameter	Height	Whole weight	Shucked weight	Viscera weight	Shell weight
0	1	0.455	0.365	0.095	0.5140	0.2245	0.1010	0.1500
1	1	0.350	0.265	0.090	0.2255	0.0995	0.0485	0.0700
2	0	0.530	0.420	0.135	0.6770	0.2565	0.1415	0.2100
3	1	0.440	0.365	0.125	0.5160	0.2155	0.1140	0.1550
4	2	0.330	0.255	0.080	0.2050	0.0895	0.0395	0.0550
...	...	...	...	...	...	...	...	...
4172	0	0.565	0.450	0.165	0.8870	0.3700	0.2390	0.2490
4173	1	0.590	0.440	0.135	0.8200	0.4390	0.2145	0.2605
4174	1	0.600	0.475	0.205	0.8200	0.5255	0.2875	0.3080
4175	0	0.625	0.485	0.150	0.8200	0.5310	0.2610	0.2960
4176	1	0.710	0.555	0.195	0.8200	0.3500	0.3765	0.4950

4177 rows × 8 columns

4177 rows × 8 columns

```
y
```

```
0       15
1        7
2        9
3       10
4        7
        ..
4172    11
4173    10
4174     9
4175    10
4176    12
Name: Rings, Length: 4177, dtype: int64
```

8.Scale the independent variables

```
from sklearn.preprocessing import scale
x = scale(x)
x
```

```
array([[-0.0105225 , -0.67088921, -0.50179694, ..., -0.61037964,
        -0.7328165 , -0.64358742],
       [-0.0105225 , -1.61376082, -1.57304487, ..., -1.22513334,
        -1.24343929, -1.25742181],
       [-1.26630752,  0.00259051,  0.08738942, ..., -0.45300269,
        -0.33890749, -0.18321163],
       ...,
       [-0.0105225 ,  0.63117159,  0.67657577, ...,  0.86994729,
         1.08111018,  0.56873549],
       [-1.26630752,  0.85566483,  0.78370057, ...,  0.89699645,
         0.82336724,  0.47666033],
       [-0.0105225 ,  1.61894185,  1.53357412, ...,  0.00683308,
         1.94673739,  2.00357336]])
```

9.Split the data into training and testing

```
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(x,y, test_size = 0.2)
print(x_train.shape, x_test.shape)
```

```
(3341, 8) (836, 8)
```

10.Build the Model

```
from sklearn.linear_model import LinearRegression
MLR=LinearRegression()
```

```
MLR=LinearRegression()
```

11.Train the model

```
MLR.fit(x_train,y_train)
```

```
LinearRegression()
```

12.Test the model

```
y_pred=MLR.predict(x_test)
y_pred
```

```
array([ 7.57936633, 10.66502123, 10.52719064, 14.09295726,  6.52127149,
        9.03476014, 13.77122971,  5.74010562, 12.66655164, 11.66459878,
       10.09473112,  9.63504043, 10.01584761,  6.70461657,  7.91484815,
        8.58663645,  8.28884611,  8.74708756,  9.42905341, 11.98435906,
       12.62869259,  4.67682252, 10.00703764, 12.49751696, 11.08348581,
        7.23035149, 10.27790557,  8.06810675, 10.16242489, 11.73469998,
        7.59073001,  9.19015028,  8.83271977,  7.97308496, 12.91320811,
       12.2190592 ,  8.34881257, 11.16830224, 11.27936239,  8.37831566,
       11.30065939, 13.29417433,  9.6526121 ,  6.66809222, 11.38626212,
       13.19243541,  7.25213555,  8.60357975, 10.64987501,  7.55777733,
       11.09076984,  7.01387278, 11.56819081, 13.09725865, 10.42979363,
        6.86170055,  9.84030194,  9.05606835,  2.71048796, 11.69429509,
       10.13966111, 11.41545898, 10.54370222,  6.87770405,  9.8169579 ,
       11.34192259,  9.9100049 ,  7.36716167, 10.34213485, 10.91520291,
        6.92937933,  7.24793046,  9.04514215, 11.06083081, 13.28804307,
        9.97379523, 10.13117441,  6.3896473 ,  7.2659971 ,  5.81386013,
        7.91573339, 10.80966746, 15.49548138,  6.39326496, 10.05181898,
        9.60191489, 12.66410799,  6.34525911,  4.26122214,  9.8604734 ,
        9.84462576,  7.77244786,  8.33307514, 11.09183305, 10.25974885,
       10.85890961,  8.08809304,  6.64356277,  5.68488117,  8.36644472,
       11.1775508 , 17.42011526, 11.09814716, 10.47328329, 11.90270427,
       14.45368598,  9.23951091, 11.27216876,  6.70690038,  6.07127817,
        8.82853003,  8.71459996,  9.83752648,  8.82818337,  9.99998632,
        7.24156173, 12.62457284,  9.66839459,  7.48648404,  9.63858685,
       15.13955152,  8.35884383, 10.07154963,  5.67547992,  9.66036627,
        9.71627737,  9.07818153,  4.17546462,  5.91589006, 13.28197225,
       12.52596824,  6.37282624,  9.951238  , 11.88214332, 10.81432205,
       11.82901   , 11.02172869, 13.1177598 , 10.41860706, 13.00609394,
       12.21916137, 11.82773365, 10.1328873 ,  9.78489566, 10.45292158,
        6.42476207, 10.98523551,  7.03558873, 11.47974779,  8.86164356,
        4.46784311, 12.26008109,  9.4999429 ,  8.87777308,  6.11221572,
       10.49585274,  9.82761991, 10.0430825 , 11.77413242, 11.15538309,
       12.22419192,  8.9234896 ,  7.48575476, 10.11745982,  9.18012612,
       12.61260463,  6.67946755, 10.41017186, 11.59259262,  9.60856303,
        8.4304331 , 12.15176482, 11.66145004, 10.45975336,  6.92040264,
       11.27564716,  8.15786795, 13.21681414,  7.03508625, 13.5483655 ,
       13.41573299,  6.75585068, 11.95648115,  9.41076026, 10.69937624,
       12.09092071,  7.7195914 , 10.66297506,  8.75053055, 12.77356642,
```

```
        7.93749495,  5.70809096,  6.82453062,  9.85826253, 11.76716135,
        6.1946293 ,  8.44980548,  9.69252471,  7.33731374, 11.26150009,
       11.81990981,  7.67084366, 16.30283727,  9.88593778, 10.25098848,
        8.22850686, 10.22450449,  8.46549692,  7.05507765, 11.34387628,
        9.78449033,  7.9874792 , 10.97755319,  6.08076973,  7.96221271,
        6.71786793,  9.85301932, 10.01848815,  7.84012578,  3.86340068,
        3.64012234,  6.3655795 , 11.41108633,  8.50772906,  8.17060765,
        6.4939989 , 10.89634719, 10.70538486,  8.55566593, 10.79566101,
        9.76347225, 12.37963222,  7.98162046,  8.29452996,  6.31002476,
        7.98144799,  8.42954492, 10.99342452, 12.54840472, 11.23986084,
       12.59755806,  7.89288359,  9.22805437,  7.47595523,  8.38549863,
        9.94058816,  8.56957082, 11.39690181, 11.83653914,  9.09171188,
        4.18835525, 12.42852965, 11.58580001, 11.08459862,  9.64528352,
       11.82043069,  9.65357609,  8.02633683, 10.11804613,  7.04179666,
        7.62287137,  9.33916959,  8.79895637,  7.7718333 , 11.78319352,
        9.39070736, 10.89029071, 12.20378255, 11.69193062, 13.06276939,
       10.50094061,  7.45419762, 12.08905415, 10.87812902, 10.6042312 ,
       10.8931744 , 11.53241428,  6.78231063,  9.95062071,  8.79770235,
       12.62433387, 10.3251268 , 10.04461074,  9.71793788, 12.83450396,
       12.48710701,  8.49295962,  7.63407404,  9.90122164, 10.42711729,
```

```
pred=MLR.predict(x_train)
pred
```

```
array([11.83129524, 10.36878858, 10.81597167, ...,  6.97338674,
       15.99517289, 10.10158868])
```

```
from sklearn.metrics import r2_score
accuracy=r2_score(y_test,y_pred)
accuracy
```

```
0.47396917577569664
```

```
MLR.predict([[1,0.455,0.365,0.095,0.5140,0.2245,0.1010,0.150]])
```

```
array([9.89061655])
```

13.Measure the performance using Metrics

```
from sklearn import metrics
from sklearn.metrics import mean_squared_error
np.sqrt(mean_squared_error(y_test,y_pred))
```

```
2.2892068787382294
```

LASSO

```
from sklearn.linear_model import Lasso, Ridge
#intialising model
```

```
#intialising model
lso=Lasso(alpha=0.01,normalize=True)
#fit the model
lso.fit(x_train,y_train)
Lasso(alpha=0.01, normalize=True)
#prediction on test data
lso_pred=lso.predict(x_test)
#coef
coef=lso.coef_
coef
```

```
array([-0.        ,  0.        ,  0.        ,  0.49754536,  0.13409893,
        0.        ,  0.        ,  0.81772052])
```

```
from sklearn import metrics
from sklearn.metrics import mean_squared_error
metrics.r2_score(y_test,lso_pred)
```

```
0.37717702228637373
```

```
np.sqrt(mean_squared_error(y_test,lso_pred))
```

```
2.4909313509148516
```

RIDGE

```
#initialising model
rg=Ridge(alpha=0.01,normalize=True)
#fit the model
rg.fit(x_train,y_train)
Ridge(alpha=0.01, normalize=True)
#prediction
rg_pred=rg.predict(x_test)
rg_pred
```

```
array([ 7.62121499, 10.6443235 , 10.46446378, 13.93585982,  6.65527712,
        8.93796128, 13.60043244,  5.71344941, 12.73093218, 11.63435105,
       10.08736018,  9.8112678 ,  9.99991056,  6.68753702,  7.95750968,
        8.6037378 ,  8.33931464,  8.886473  ,  9.43192298, 11.95234039,
       12.57154019,  4.83025431, 10.01751159, 12.45594306, 11.06183705,
        7.22138319, 10.32510163,  7.98435092, 10.16120159, 11.6908897 ,
        7.58122775,  9.22837956,  8.83100028,  7.97992147, 12.81725716,
       12.19609656,  8.34416145, 11.11108119, 11.21772266,  8.37833812,
       11.28702945, 13.2859872 ,  9.60598624,  6.69408579, 11.41491795,
       13.14029802,  7.26785573,  8.57017   , 10.6518417 ,  7.56841222,
       11.19228417,  7.03558079, 11.52322241, 13.0027726 , 10.49778702,
        6.85693179,  9.8355421 ,  9.15913735,  2.93295095, 11.5989431 ,
       10.11566461, 11.41932012, 10.59303973,  6.88553927,  9.82287673,
       11.34277873,  9.91505015,  7.33652603, 10.40432793, 10.8444582 ,
        6.91057454,  7.27226987,  9.07154551, 10.96051363, 13.17367646,
```

```
       10.05187336, 10.30800225,  6.37466671,  7.28018054,  5.8121994 ,
        7.9195244 , 10.82983089, 15.38288103,  6.43288873, 10.05120677,
        9.63078325, 12.5937629 ,  6.34391917,  4.40563881,  9.88341135,
        9.87412649,  7.75609893,  8.31424866, 11.04355197, 10.1862527 ,
       10.73277044,  8.10485704,  6.66741905,  5.67906445,  8.35448985,
       11.18466136, 17.03262128, 11.08252138, 10.54407412, 11.89028589,
       14.3315458 ,  9.24653746, 11.27902251,  6.70271799,  6.0471902 ,
        8.88372211,  8.73002086,  9.83942534,  9.04135357,  9.96305452,
        7.25389147, 12.47924664,  9.76866353,  7.5017228 ,  9.62986923,
       14.96517323,  8.38927674, 10.06000395,  5.65232962,  9.651335  ,
        9.85295997,  9.30196415,  4.33573538,  5.89981879, 13.11786636,
       12.42344984,  6.37610943,  9.90322124, 11.75538903, 10.84505857,
       11.70652114, 11.05210645, 13.01488237, 10.4507523 , 12.98209839,
       12.21669643, 11.81210138, 10.17766619,  9.83854577, 10.51857004,
        6.48273121, 11.03898851,  7.03787966, 11.56771889,  8.82848192,
        4.6076826 , 12.27709271,  9.46887056,  8.87017495,  6.11487577,
       10.52567743,  9.86533274, 10.05876542, 11.7041847 , 11.11522135,
       12.18445481,  8.91375643,  7.45880889, 10.08552472,  9.18002615,
       12.49920307,  6.73582965, 10.33788768, 11.54786663,  9.71506904,
        8.43297678, 12.2075221 , 11.64781313, 10.52090148,  6.90757324,
       11.19314325,  8.14487022, 13.1849842 ,  7.03508294, 13.3704564 ,
       13.30760984,  6.75183101, 11.82497051,  9.51809872, 10.67327954,
       12.04282518,  7.6976537 , 10.6568258 ,  8.75888305, 12.72740643,
        7.95879175,  5.69490724,  6.87393297,  9.90728648, 11.68259408,
        6.20330353,  8.73007157,  9.8072228 ,  7.33835382, 11.25925132,
       11.74078149,  7.66432063, 16.07693408,  9.96670063, 10.31261133,
        8.24976323, 10.27288996,  8.49641432,  7.09684347, 11.40517361,
        9.76334845,  7.98663939, 10.89084724,  6.08198496,  7.92645778,
        6.75323523,  9.88759136, 10.09877923,  7.82799109,  4.00952334,
        4.06034141,  6.3603101 , 11.42501292,  8.69673594,  8.15112169,
        6.4971995 , 10.88723374, 10.61315147,  8.61641777, 10.7963036 ,
        9.76693341, 12.31632803,  8.00864941,  8.32873939,  6.32206682,
        8.01937372,  8.49851   , 11.0703969 , 12.49907703, 11.24153928,
       12.55907248,  7.91145771,  9.20671665,  7.47694066,  8.41410748,
       10.07276621,  8.55496191, 11.36144111, 11.80408382,  9.0956383 ,
        4.33016868, 12.44041728, 11.75783181, 11.04021359,  9.63056605,
       11.80039941,  9.70422452,  8.02610498, 10.13973282,  7.0221195 ,
        7.63284225,  9.39167944,  8.97161542,  7.78854803, 11.7130127 ,
        9.44491279, 10.98074043, 12.23942484, 11.68111844, 12.94120897,
       10.39331962,  7.44982184, 12.03965461, 10.90808205, 10.63151037,
       10.91542183, 11.49571698,  6.79676636,  9.94118808,  8.81501079,
       12.52318497, 10.41638664, 10.18244953,  9.65744053, 12.7584745 ,
       12.50333536,  8.47319341,  7.63649913,  9.87717652, 10.48263845,
```

```
1 rg.coef_
```

```
array([-0.31430508, -0.66861058,  0.21952353,  1.01917483,  0.96446379,
       -1.46767724, -0.09674935,  1.80795097])
```

```
1 metrics.r2_score(y_test,rg_pred)
```

```
0.4751366340259142
```

```
1 np.sqrt(mean_squared_error(y_test,rg_pred))
```

```
2.2866651665138993
```