Assignment -4

Problem Statement :- SMS SPAM Classification

Assignment Date	19 October 2022
Student Name	Adithyan.M
Student Roll Number	CS19002
Maximum Marks	2 Marks

```
#import the necessary files
import numpy as np import
pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline import
seaborn as sns
from sklearn.tree import DecisionTreeRegressor from sklearn.ensemble
import RandomForestRegressorfrom sklearn.linear_model import
LinearRegression
from sklearn.metrics import mean_squared_error,r2_scorefrom sklearn.svm import
SVR
from sklearn.pipeline import Pipeline from sklearn.impute
import SimpleImputer
from sklearn.preprocessing import StandardScalerdata =
pd.read_csv("Mall_Customers.csv") data.head()
```

	CustomerID	Gender	Age	Annual Income (k$)	Spending Score (1-100)0	1
		Male	19	15	39	
1		2	Male	21	15	81
2		3	Female	20	16	6
3		4	Female	23	16	77
4		5	Female	31	17	40

```
print(data.shape)
data.info()
```

```
(200, 5)
<class 'pandas.core.frame.DataFrame'>RangeIndex:
200 entries, 0 to 199 Data columns (total 5 columns):
```

#	Column	Non-Null Count	Dtype
0	CustomerID	200 non-null	int64
1	Gender	200 non-null	object
2	Age	200 non-null	int64
3	Annual Income (k$)	200 non-null	int64

4 Spending Score (1-100) 200 non-null int64
dtypes: int64(4), object(1)memory
usage: 7.9+ KB

data.describe().T

		count	mean	std	min	25%	50%
75% \							
CustomerID		200.0	100.50	57.879185	1.0	50.75	100.5
150.25							
Age		200.0	38.85	13.969007	18.0	28.75	36.0
49.00							
Annual Income	(k$)	200.0	60.56	26.264721	15.0	41.50	61.5

78.00

Spending Score (1-100)	200.0	50.20	25.823522	1.0	34.75	50.0
73.00						

	max
CustomerID	200.0
Age	70.0
Annual Income (k$)	137.0
Spending Score (1-100)	99.0

Visualizations

Univariate Analysis

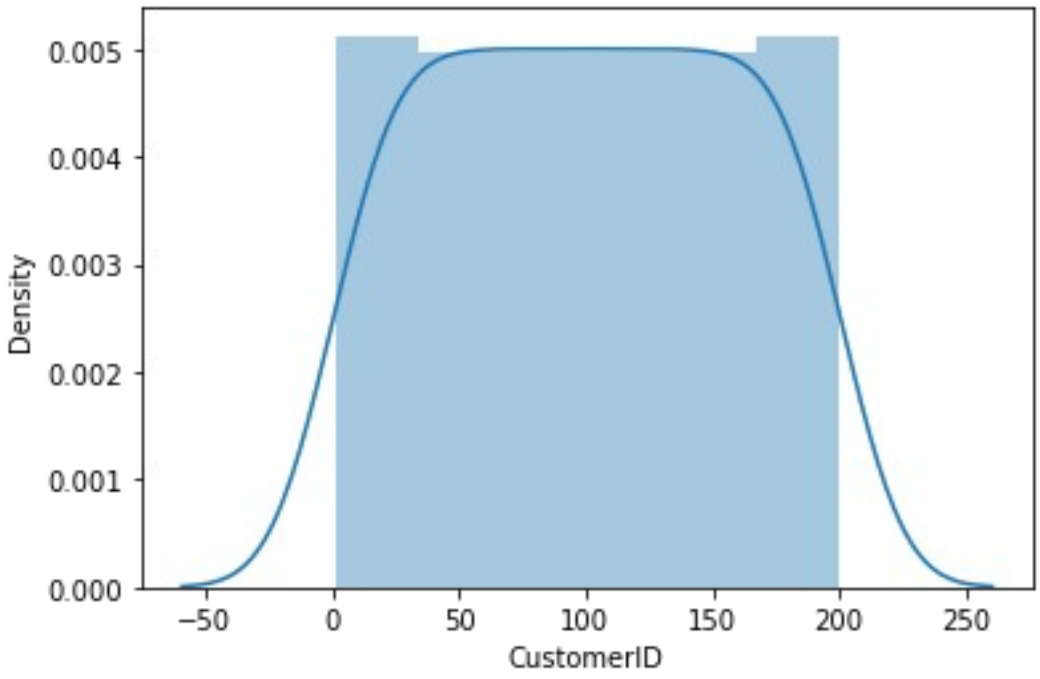
data.head()

	CustomerID	Gender	Age	Annual Income (k$)	Spending Score	(1-100)
0	1	Male	19	15		39
1	2	Male	21	15		81
2	3	Female	20	16		6
3	4	Female	23	16		77
4	5	Female	31	17		40

```
sns.distplot(data['CustomerID'])
```

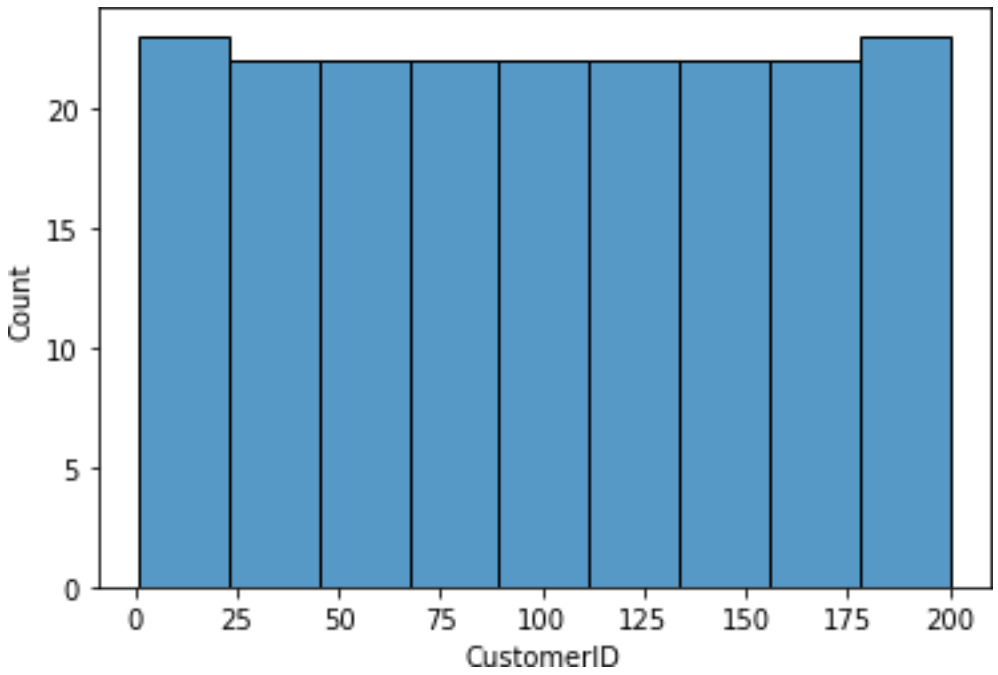
```
C:\Users\balas\anaconda3\lib\site-packages\seaborn\ distributions.py:2619: FutureWarning:
`distplot` is a deprecated function and will be removed in a future version. Please adapt yourcode to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).
    warnings.warn(msg, FutureWarning)
```

```
<AxesSubplot:xlabel='CustomerID', ylabel='Density'>
```



sns.histplot(data['CustomerID'])

<AxesSubplot:xlabel='CustomerID', ylabel='Count'>

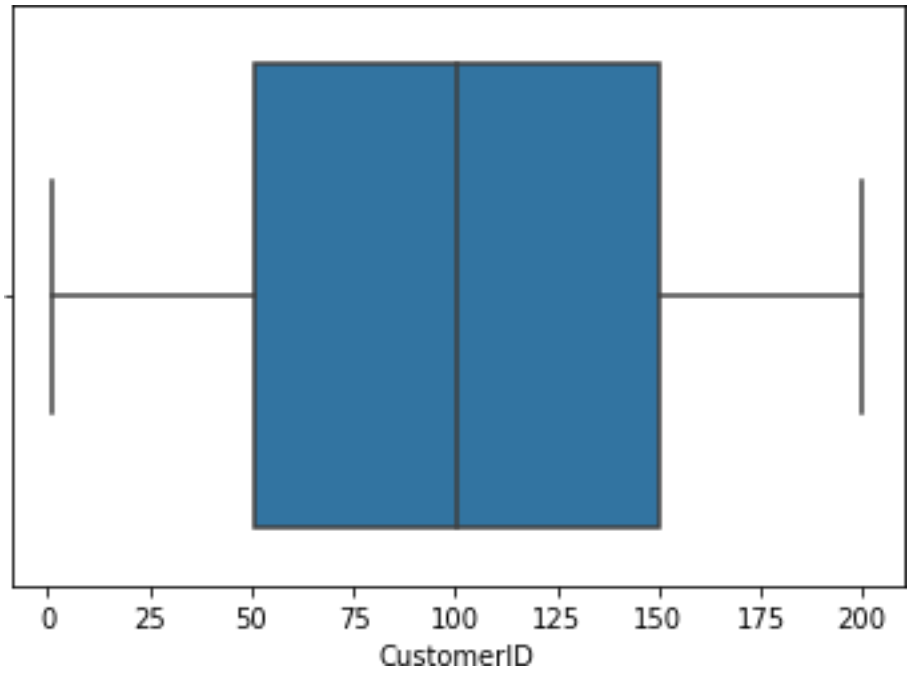


sns.boxplot(data['CustomerID'])

C:\Users\balas\anaconda3\lib\site-packages\seaborn_decorators.py:36:FutureWarning: Pass the following variable as a keyword arg: x. From

version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in anerror or misinterpretation.
    warnings.warn(

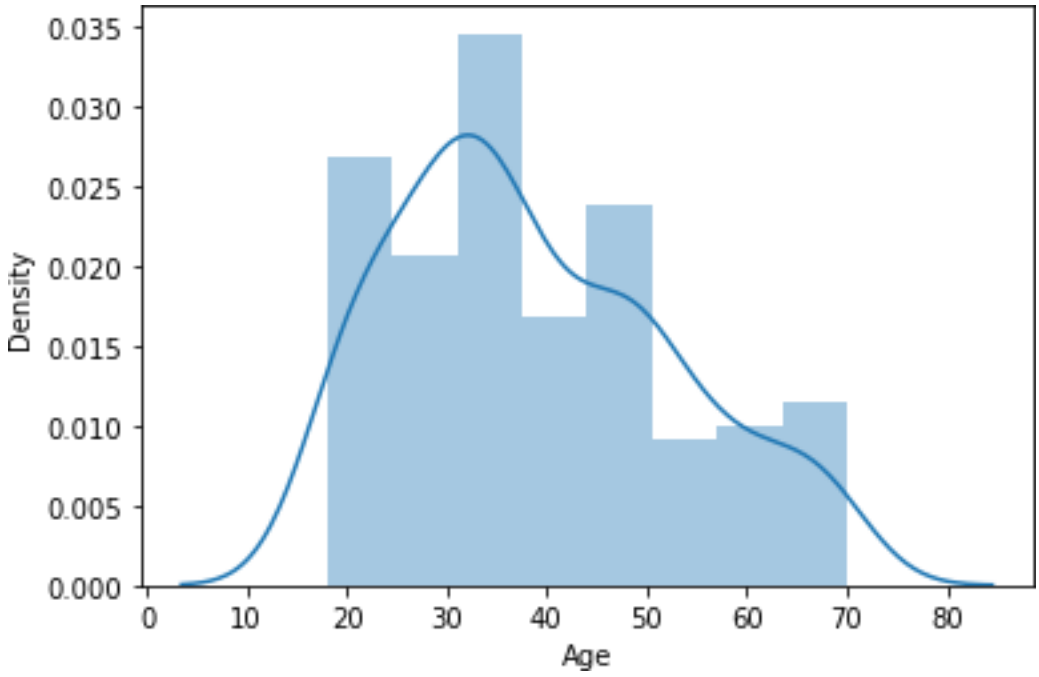
<AxesSubplot:xlabel='CustomerID'>



sns.distplot(data['Age'])

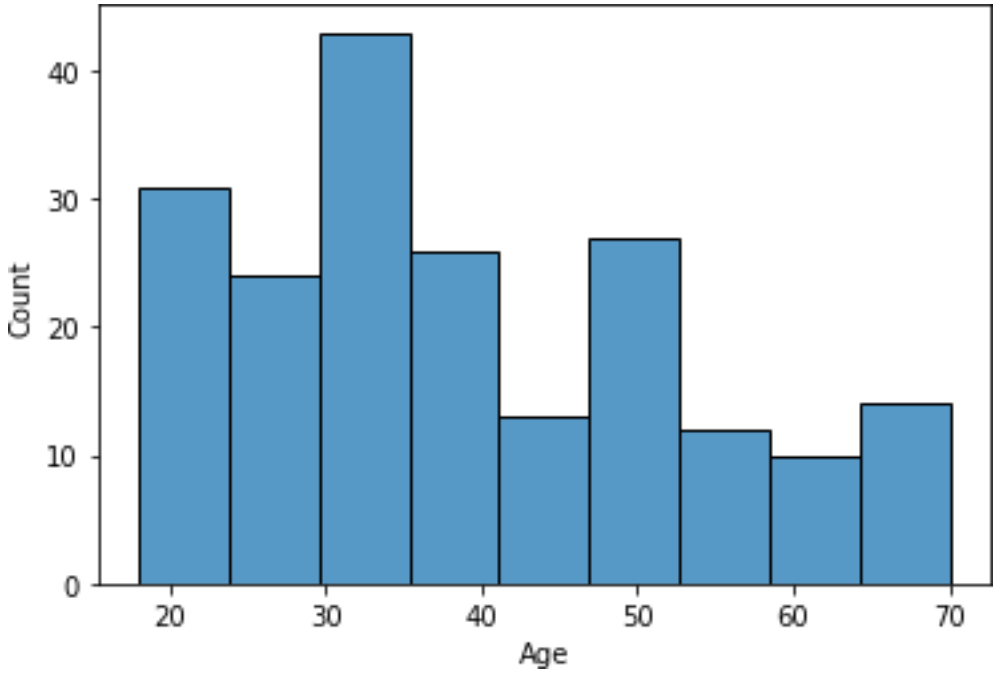
C:\Users\balas\anaconda3\lib\site-packages\seaborn\ distributions.py:2619: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt yourcode to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).
    warnings.warn(msg, FutureWarning)

<AxesSubplot:xlabel='Age', ylabel='Density'>



sns.histplot(data['Age'])

<AxesSubplot:xlabel='Age', ylabel='Count'>

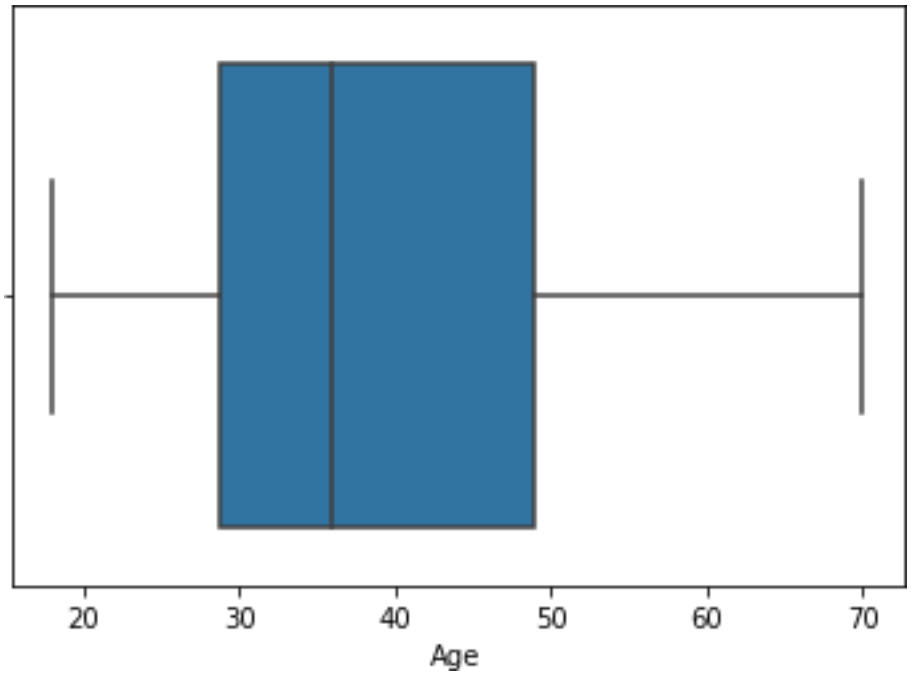


sns.boxplot(data['Age'])

C:\Users\balas\anaconda3\lib\site-packages\seaborn_decorators.py:36:FutureWarning: Pass the following variable as a keyword arg: x. From

version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in anerror or misinterpretation.
    warnings.warn(

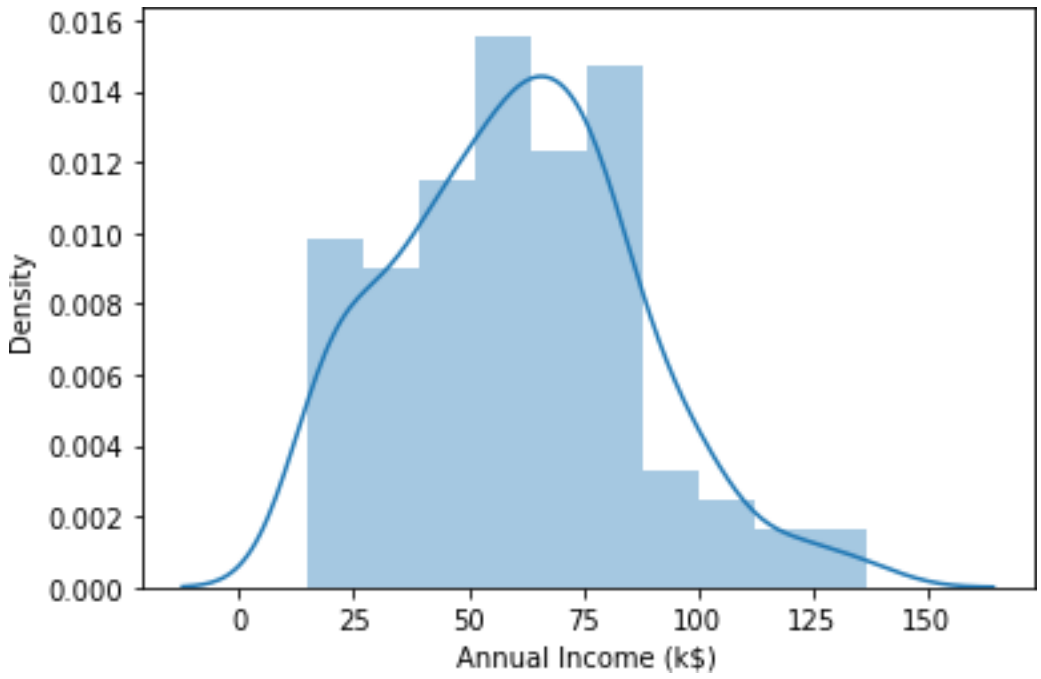
<AxesSubplot:xlabel='Age'>



sns.distplot(data['Annual Income (k$)'])

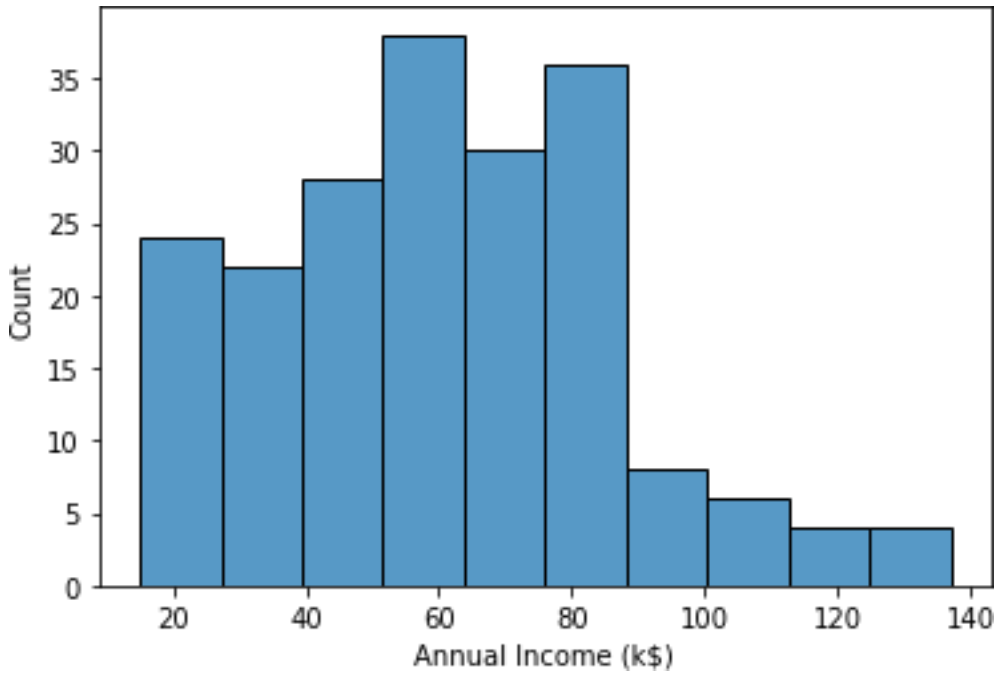
C:\Users\balas\anaconda3\lib\site-packages\seaborn\ distributions.py:2619: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt yourcode to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).
    warnings.warn(msg, FutureWarning)

<AxesSubplot:xlabel='Annual Income (k$)', ylabel='Density'>



sns.histplot(data['Annual Income (k$)'])

<AxesSubplot:xlabel='Annual Income (k$)', ylabel='Count'>



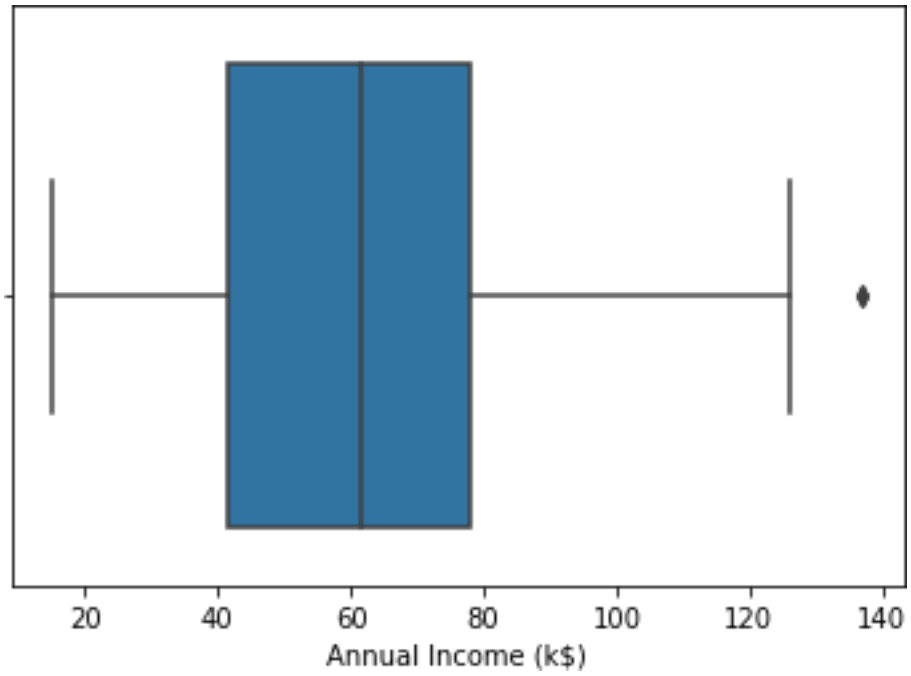
sns.boxplot(data['Annual Income (k$)'])

C:\Users\balas\anaconda3\lib\site-packages\seaborn_decorators.py:36:FutureWarning: Pass the following variable as a keyword arg: x. From

version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in anerror or misinterpretation.

warnings.warn(

<AxesSubplot:xlabel='Annual Income (k$)'>

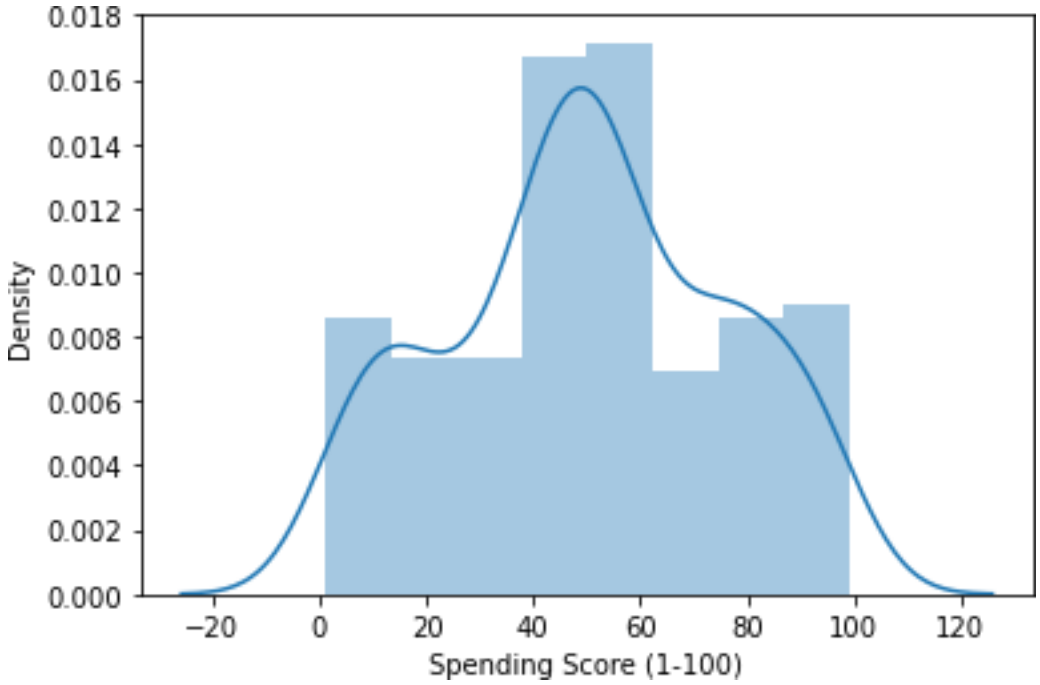


sns.distplot(data['Spending Score (1-100)'])

C:\Users\balas\anaconda3\lib\site-packages\seaborn\ distributions.py:2619: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt yourcode to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).

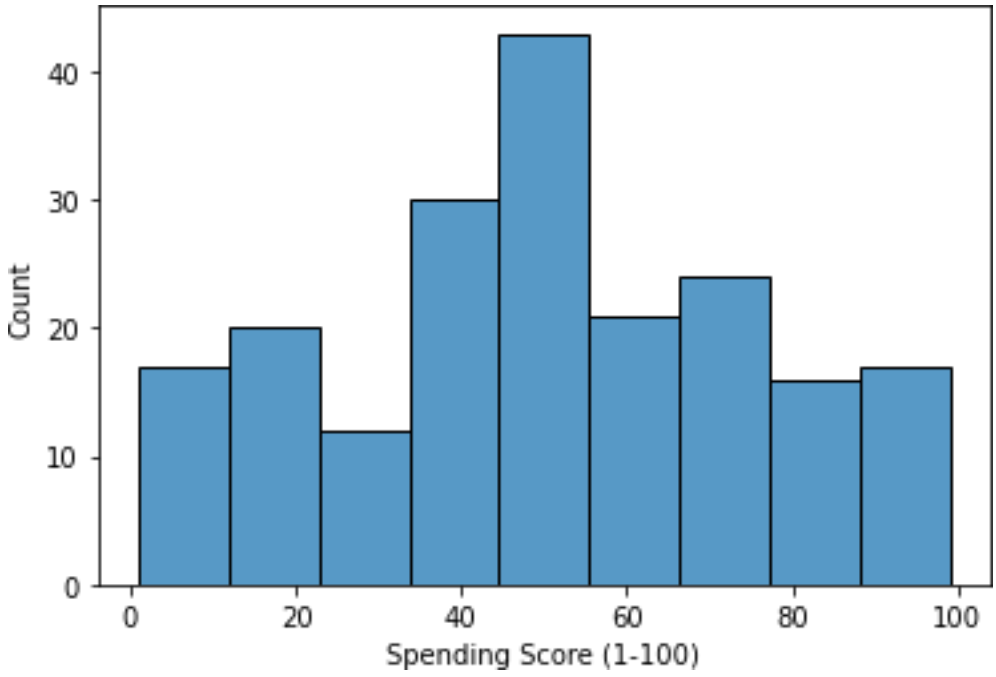
warnings.warn(msg, FutureWarning)

<AxesSubplot:xlabel='Spending Score (1-100)', ylabel='Density'>



sns.histplot(data['Spending Score (1-100)'])

<AxesSubplot:xlabel='Spending Score (1-100)', ylabel='Count'>

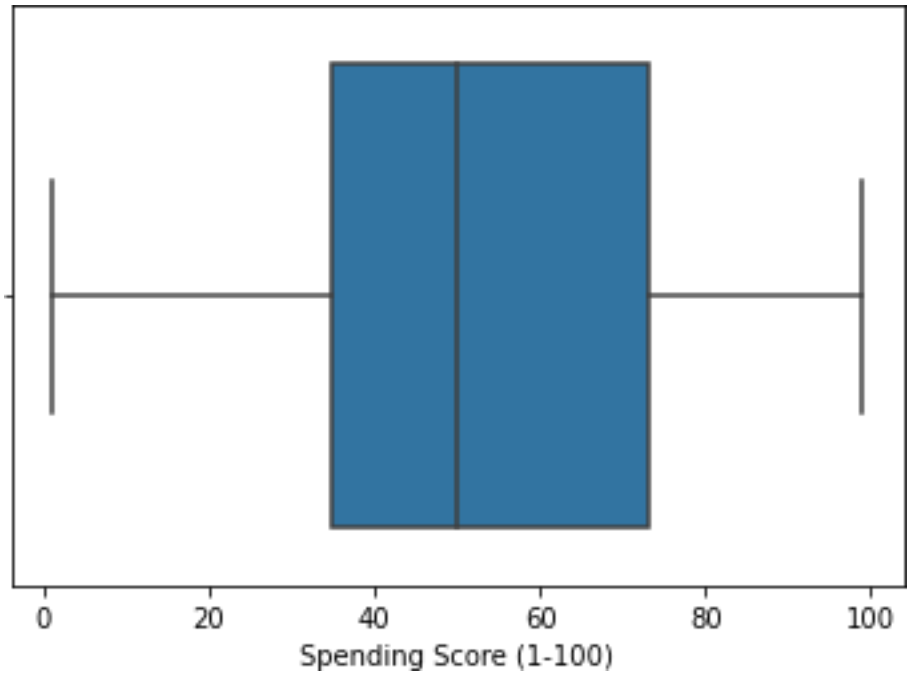


sns.boxplot(data['Spending Score (1-100)'])

C:\Users\balas\anaconda3\lib\site-packages\seaborn_decorators.py:36:FutureWarning: Pass the following variable as a keyword arg: x. From

version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in anerror or misinterpretation.
    warnings.warn(

<AxesSubplot:xlabel='Spending Score (1-100)'>

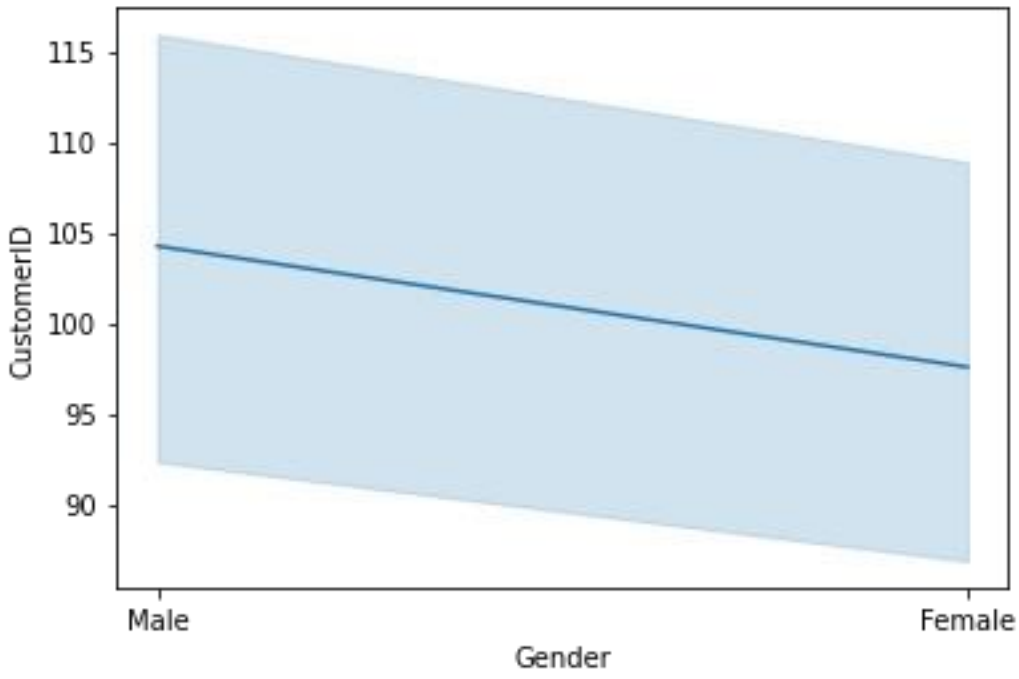


Bivariate Analysis

```
sns.lineplot(data['Gender'], data['CustomerID'])
```

C:\Users\balas\anaconda3\lib\site-packages\seaborn_decorators.py:36:FutureWarning: Pass the following variables as keyword args: x, y.
From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result inan error or misinterpretation.
    warnings.warn(

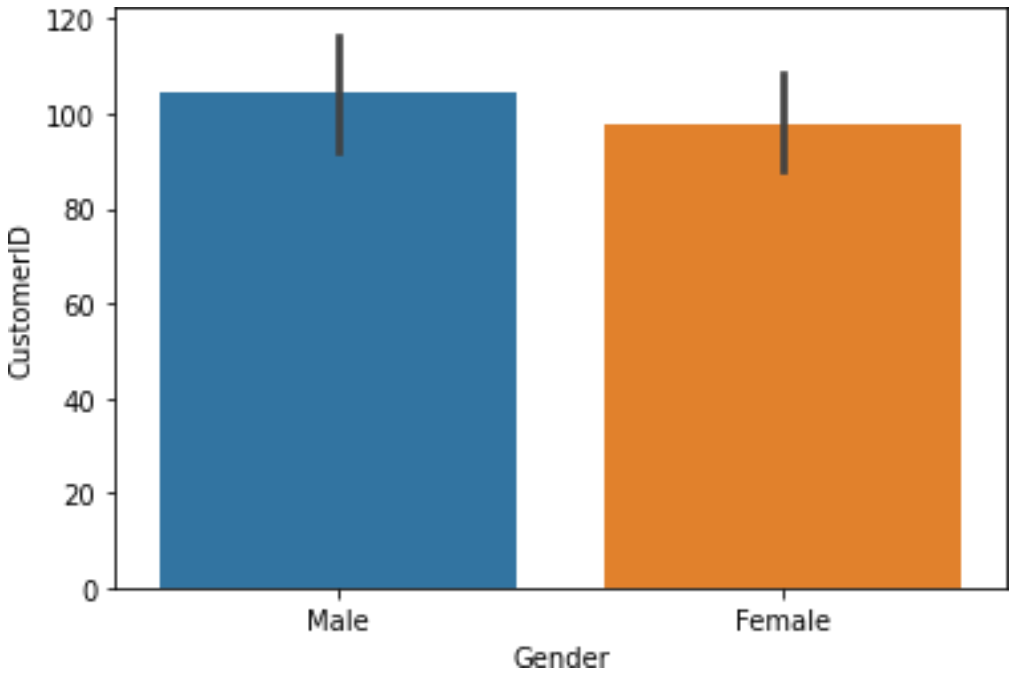
<AxesSubplot:xlabel='Gender', ylabel='CustomerID'>




```
sns.barplot(data['Gender'], data['CustomerID'])
```

```
C:\Users\balas\anaconda3\lib\site-packages\seaborn\_decorators.py:36:FutureWarning: Pass the following variables as keyword args: x, y.
From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result inan error or misinterpretation.
    warnings.warn(
```

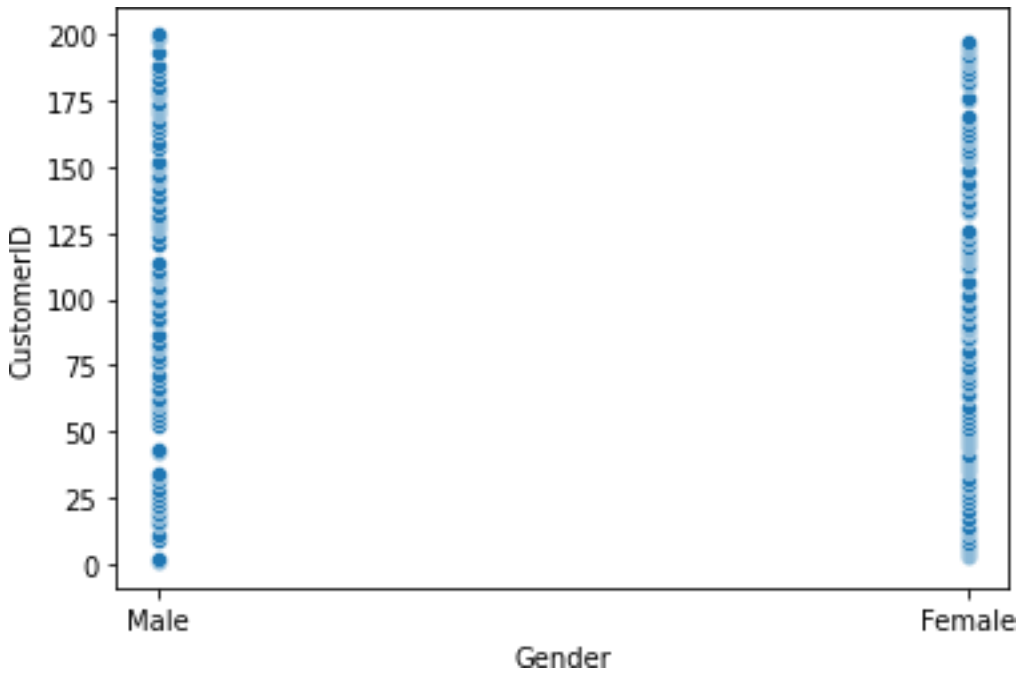
```
<AxesSubplot:xlabel='Gender', ylabel='CustomerID'>
```




```
sns.scatterplot(data['Gender'], data['CustomerID'])
```

```
C:\Users\balas\anaconda3\lib\site-packages\seaborn\_decorators.py:36:FutureWarning: Pass the following variables as keyword args: x, y.
From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result inan error or misinterpretation.
    warnings.warn(
```

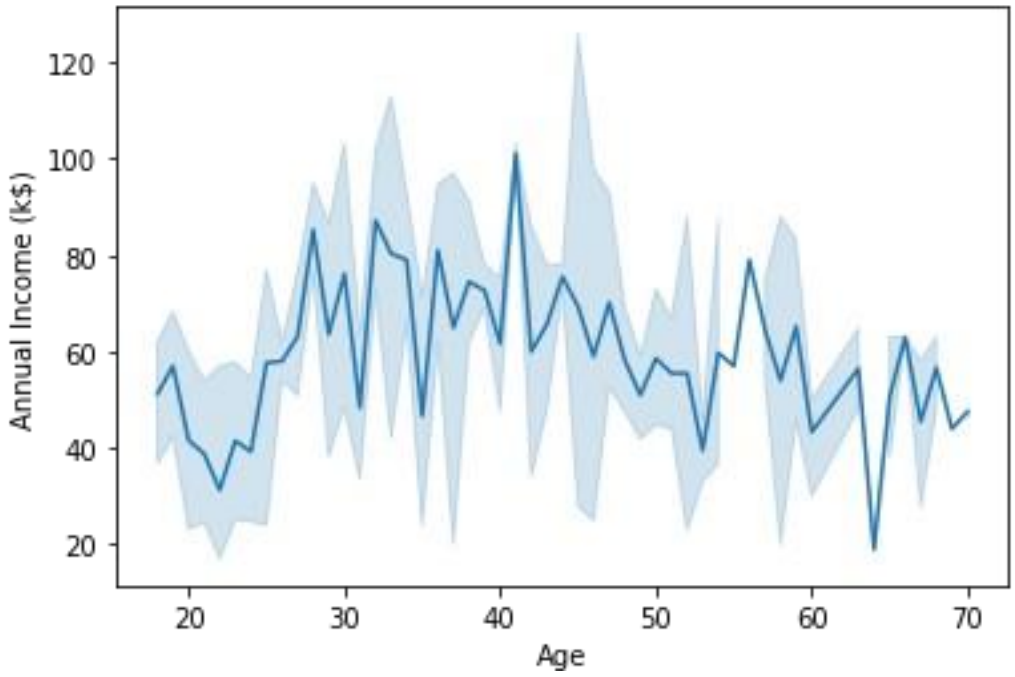
```
<AxesSubplot:xlabel='Gender', ylabel='CustomerID'>
```




```
sns.lineplot(data['Age'], data['Annual Income (k$)'])
```

```
C:\Users\balas\anaconda3\lib\site-packages\seaborn\_decorators.py:36:FutureWarning: Pass the following variables as keyword args: x, y.
From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result inan error or misinterpretation.
    warnings.warn(
```

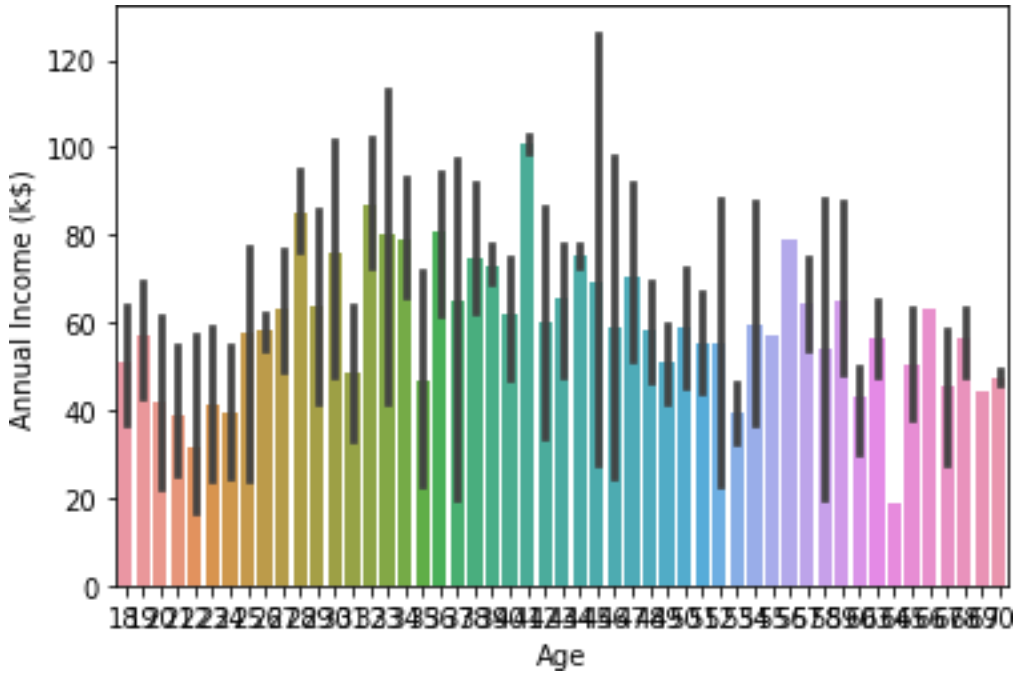
```
<AxesSubplot:xlabel='Age', ylabel='Annual Income (k$)'>
```




```
sns.barplot(data['Age'], data['Annual Income (k$)'])
```

```
C:\Users\balas\anaconda3\lib\site-packages\seaborn\_decorators.py:36:FutureWarning: Pass the following variables as keyword args: x, y.
From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result inan error or misinterpretation.
    warnings.warn(
```

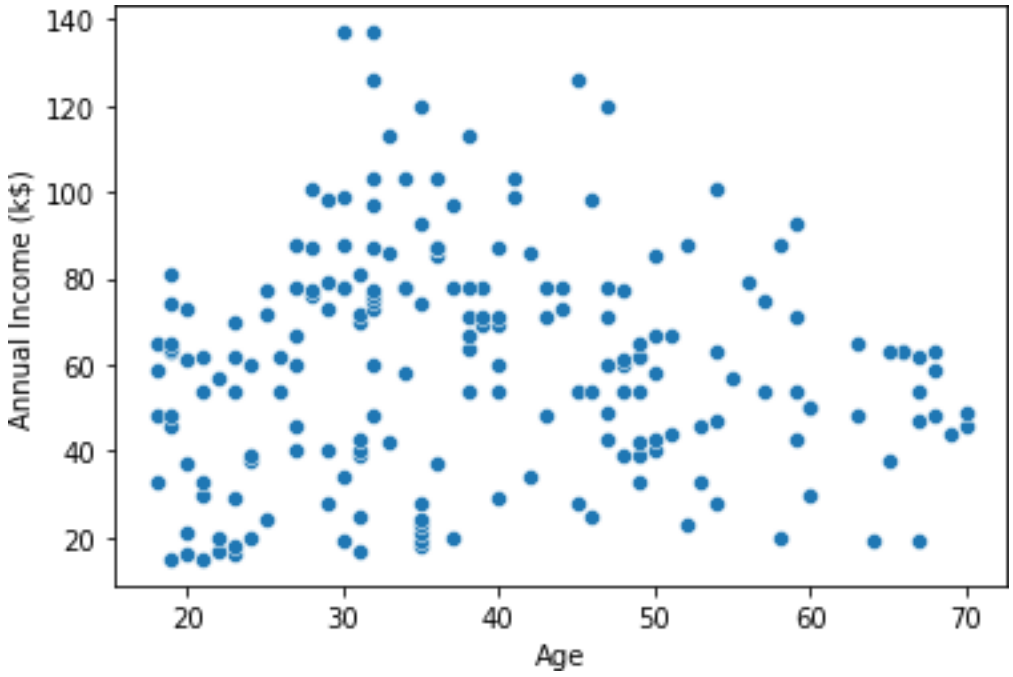
```
<AxesSubplot:xlabel='Age', ylabel='Annual Income (k$)'>
```



sns.scatterplot(data['Age'], data['Annual Income (k$)'])

C:\Users\balas\anaconda3\lib\site-packages\seaborn_decorators.py:36:FutureWarning: Pass the following variables as keyword args: x, y.
From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result inan error or misinterpretation.
  warnings.warn(

<AxesSubplot:xlabel='Age', ylabel='Annual Income (k$)'>

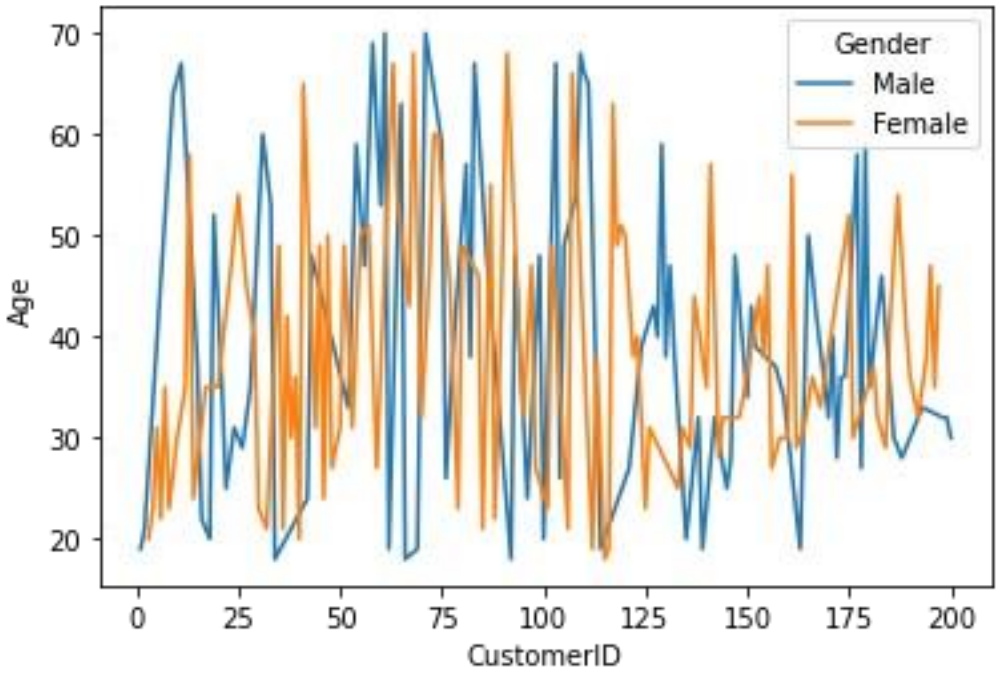


Multivariate Analysis

```
sns.lineplot(data['CustomerID'], data['Age'], hue = data['Gender'])
```

```
C:\Users\balas\anaconda3\lib\site-packages\seaborn\_decorators.py:36:FutureWarning: Pass the following variables as keyword args: x, y.
From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result inan error or misinterpretation.
    warnings.warn(
```

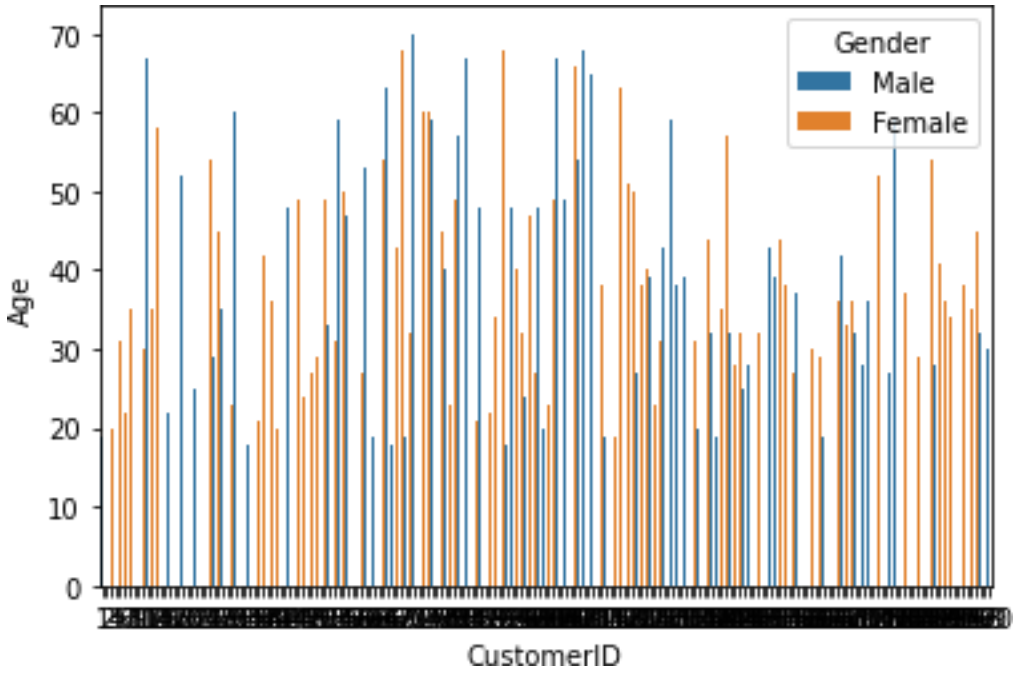
```
<AxesSubplot:xlabel='CustomerID', ylabel='Age'>
```




```
sns.barplot(data['CustomerID'], data['Age'], hue = data['Gender'])
```

```
C:\Users\balas\anaconda3\lib\site-packages\seaborn\_decorators.py:36:FutureWarning: Pass the following variables as keyword args: x, y.
From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result inan error or misinterpretation.
    warnings.warn(
```

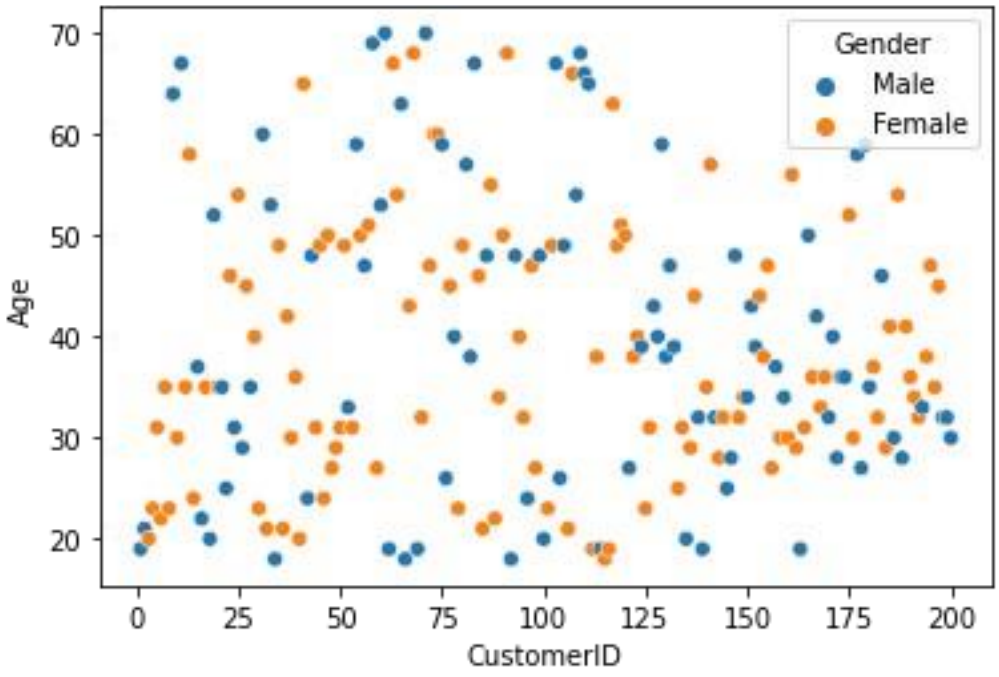
```
<AxesSubplot:xlabel='CustomerID', ylabel='Age'>
```




```
sns.scatterplot(data['CustomerID'], data['Age'], hue = data['Gender'])
```

```
C:\Users\balas\anaconda3\lib\site-packages\seaborn\_decorators.py:36:FutureWarning: Pass the following variables as keyword args: x, y.
From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result inan error or misinterpretation.
    warnings.warn(
```

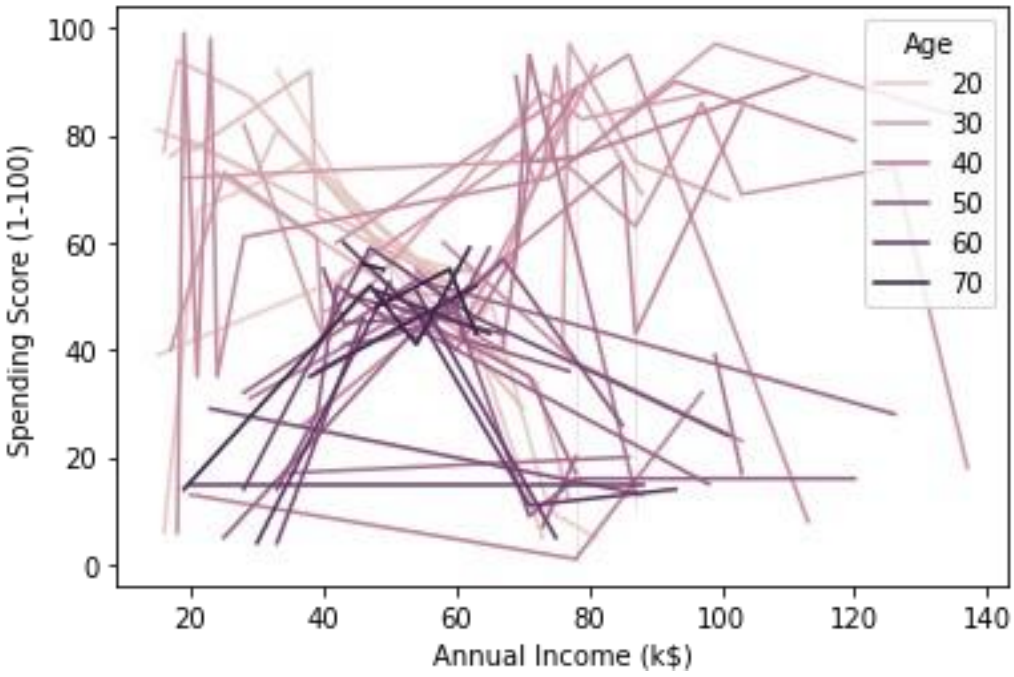
```
<AxesSubplot:xlabel='CustomerID', ylabel='Age'>
```




```
sns.lineplot(data['Annual Income (k$)'], data['Spending Score (1-100)'], hue = data['Age'])
```

```
C:\Users\balas\anaconda3\lib\site-packages\seaborn\_decorators.py:36:FutureWarning: Pass the following variables as keyword args: x, y.
From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result inan error or misinterpretation.
    warnings.warn(
```

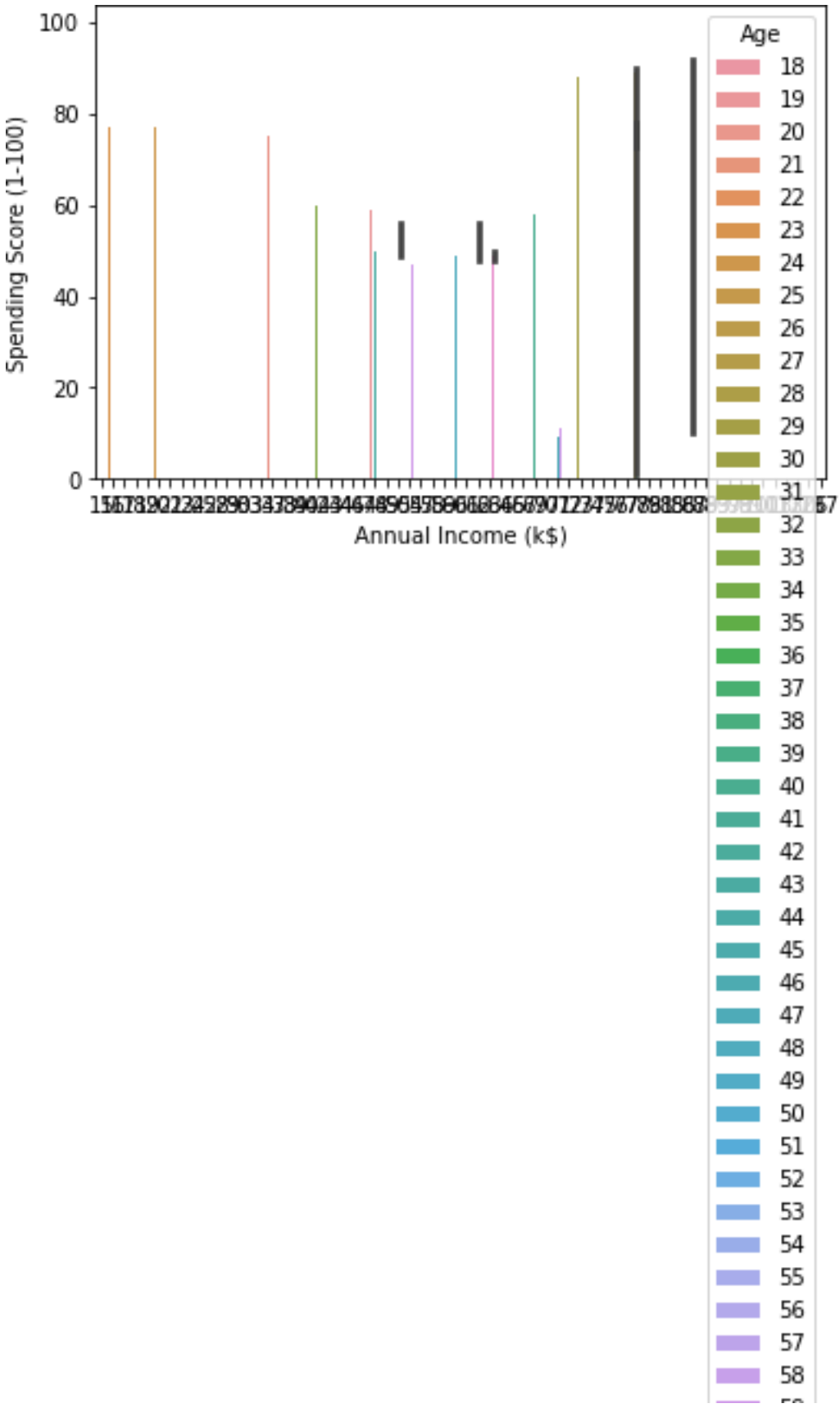
```
<AxesSubplot:xlabel='Annual Income (k$)', ylabel='Spending Score (1-100)'>
```




```
sns.barplot(data['Annual Income (k$)'], data['Spending Score (1-100)'], hue = data['Age'])
```

```
C:\Users\balas\anaconda3\lib\site-packages\seaborn\_decorators.py:36:FutureWarning: Pass the following variables as keyword args: x, y.
From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result inan error or misinterpretation.
    warnings.warn(
```

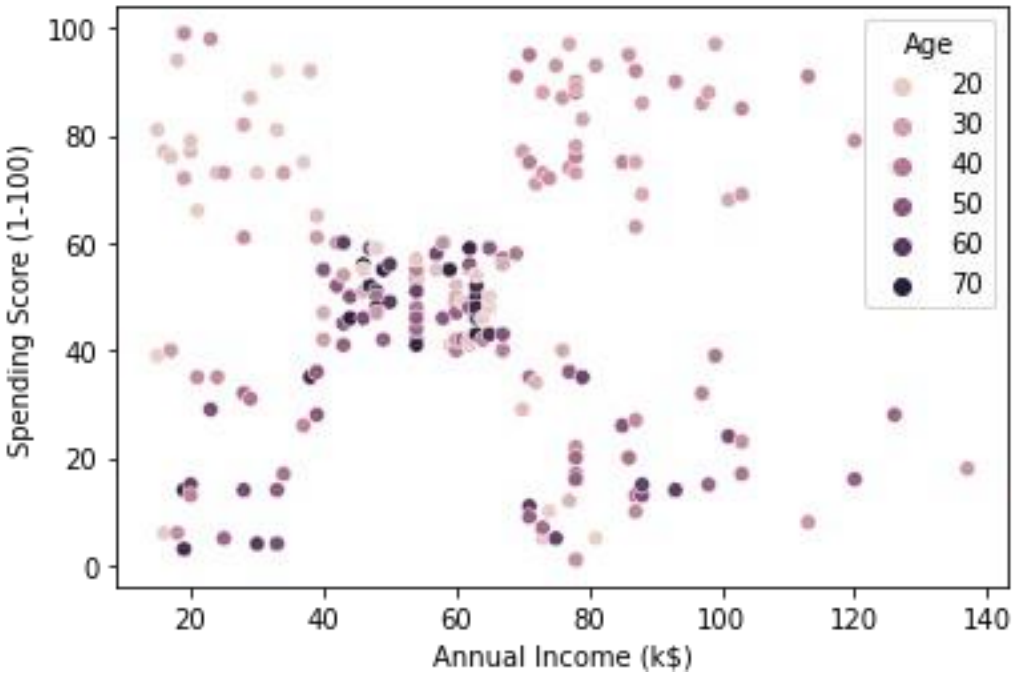
```
<AxesSubplot:xlabel='Annual Income (k$)', ylabel='Spending Score (1-100)'>
```


Spending Score (1-100)
100
80
60
40
20
0
Annual Income (k$)
Age
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58

```
sns.scatterplot(data['Annual Income (k$)'], data['Spending Score (1-100)'], hue = data['Age'])
```

```
C:\Users\balas\anaconda3\lib\site-packages\seaborn\_decorators.py:36:FutureWarning: Pass the following variables as keyword args: x, y.
From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result inan error or misinterpretation.
    warnings.warn(
```

```
<AxesSubplot:xlabel='Annual Income (k$)', ylabel='Spending Score (1-100)'>
```



Descriptive statistics

```
data.mean()
```

```
C:\Users\balas\AppData\Local\Temp\ipykernel_12512\531903386.py:1: FutureWarning: Dropping of nuisance columns in DataFrame reductions(with 'numeric_only=None') is deprecated; in a future version this will raise TypeError.  Select only valid columns before calling thereduction.
    data.mean()
```

```
CustomerID                100.50
Age                        38.85
Annual Income (k$)         60.56
Spending Score (1-100)     50.20
dtype: float64
```

```
data.mode()
```

	CustomerID	Gender	Age	Annual Income (k$)	Spending Score (1-100)
0	1	Female	32.0	54.0	42.0
1	2	NaN	NaN	78.0	NaN
2	3	NaN	NaN	NaN	NaN
3	4	NaN	NaN	NaN	NaN
4	5	NaN	NaN	NaN	NaN
..	...	...	...	...	...
195	196	NaN	NaN	NaN	NaN
196	197	NaN	NaN	NaN	NaN
197	198	NaN	NaN	NaN	NaN
198	199	NaN	NaN	NaN	NaN
199	200	NaN	NaN	NaN	NaN

[200 rows x 5 columns]

```
data.median()
```

C:\Users\balas\AppData\Local\Temp\ipykernel_12512\4184645713.py:1: FutureWarning: Dropping of nuisance columns in DataFrame reductions(with 'numeric_only=None') is deprecated; in a future version this will raise TypeError. Select only valid columns before calling thereduction.

```
    data.median()
```

```
CustomerID                100.5
Age                        36.0
Annual Income (k$)         61.5
Spending Score (1-100)     50.0
dtype: float64
data.var()
```

C:\Users\balas\AppData\Local\Temp\ipykernel_12512\445316826.py:1: FutureWarning: Dropping of nuisance columns in DataFrame reductions(with 'numeric_only=None') is deprecated; in a future version this will raise TypeError. Select only valid columns before calling thereduction.

```
    data.var()
```

```
CustomerID                    3350.000000
Age                            195.133166
Annual Income (k$)             689.835578
Spending Score (1-100)         666.854271
dtype: float64
```

```
data.std()
```

```
C:\Users\balas\AppData\Local\Temp\ipykernel_12512\2723740006.py:1: FutureWarning: Dropping of nuisance columns in DataFrame reductions(with 'numeric_only=None') is deprecated; in a future version this will raise TypeError.  Select only valid columns before calling thereduction.
  data.std()
```

```
CustomerID                    57.879185
Age                           13.969007
Annual Income (k$)            26.264721
Spending Score (1-100)        25.823522
dtype: float64
```

```
data.describe()
```

	CustomerID	Age	Annual Income (k$)	Spending Score (1-100)
count	200.000000	200.000000	200.000000	200.000000
mean	100.500000	38.850000	60.560000	50.200000
std	57.879185	13.969007	26.264721	25.823522
min	1.000000	18.000000	15.000000	1.000000
25%	50.750000	28.750000	41.500000	34.750000
50%	100.500000	36.000000	61.500000	50.000000
75%	150.250000	49.000000	78.000000	73.000000
max	200.000000	70.000000	137.000000	99.000000

```
data['Age'].unique ()array([19,
54,
       29, 21, 20, 23, 31, 22, 35, 64, 30, 67, 58, 24, 37, 52, 25, 46,
51,
       69, 45, 40, 60, 53, 18, 49, 42, 36, 65, 48, 50, 27, 33, 59, 47,
41],       70, 63, 43, 68, 32, 26, 57, 38, 55, 34, 66, 39, 44, 28, 56,
      dtype=int64)
```

```
data['Gender'].unique ()
```

```
array(['Male', 'Female'], dtype=object)
data['CustomerID'].unique ()
array([  1,   2,   3,   4,   5,   6,   7,   8,   9,  10,  11,  12,
13,
        14,  15,  16,  17,  18,  19,  20,  21,  22,  23,  24,  25,
26,
        27,  28,  29,  30,  31,  32,  33,  34,  35,  36,  37,  38,
39,
        40,  41,  42,  43,  44,  45,  46,  47,  48,  49,  50,  51,
52,
        53,  54,  55,  56,  57,  58,  59,  60,  61,  62,  63,  64,
65,
        66,  67,  68,  69,  70,  71,  72,  73,  74,  75,  76,  77,
78,
        79,  80,  81,  82,  83,  84,  85,  86,  87,  88,  89,  90,
91,
        92,  93,  94,  95,  96,  97,  98,  99, 100, 101, 102, 103,
104,
       105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116,
117,
       118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129,
130,
       131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142,
143,
       144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155,
156,
       157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168,
169,
       170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181,
182,
       183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194,
195,
       196, 197, 198, 199, 200], dtype=int64)
```

```
data['Annual Income (k$)'].unique ()
```

```
array([ 15,  16,  17,  18,  19,  20,  21,  23,  24,  25,  28,  29,
30,
        33,  34,  37,  38,  39,  40,  42,  43,  44,  46,  47,  48,
49,
        50,  54,  57,  58,  59,  60,  61,  62,  63,  64,  65,  67,
69,
        70,  71,  72,  73,  74,  75,  76,  77,  78,  79,  81,  85,
86,
```

87, 88, 93, 97, 98, 99, 101, 103, 113, 120, 126, 137], dtype=int64)

```
data['Spending Score (1-100)'].unique ()
```

```
array([39, 81,  6, 77, 40, 76, 94,  3, 72, 14, 99, 15, 13, 79, 35, 66,
29,
       98, 73,  5, 82, 32, 61, 31, 87,  4, 92, 17, 26, 75, 36, 28, 65,
55,
       47, 42, 52, 60, 54, 45, 41, 50, 46, 51, 56, 59, 48, 49, 53, 44,
57,
       58, 43, 91, 95, 11,  9, 34, 71, 88,  7, 10, 93, 12, 97, 74, 22,
90,
       20, 16, 89,  1, 78, 83, 27, 63, 86, 69, 24, 68, 85, 23,  8,

18],
      dtype=int64)
```

Check for Missing values and deal with them.

```
data.isnull().any()
```

```
CustomerID                      False
Gender                          False
Age                             False
Annual Income (k$)              False
Spending Score (1-100)dtype:    False
bool
```

```
data.isnull().sum()
```

```
CustomerID                      0
Gender                          0
Age                             0
Annual Income (k$)              0
Spending Score (1-100)dtype:    0
int64
```

```
data.mode()
```

	CustomerID	Gender	Age	Annual	Income	(k$)	Spending Score (1-100)
0	1	Female	32.0			54.0	42.0
1	2	NaN	NaN			78.0	NaN
2	3	NaN	NaN			NaN	NaN
3	4	NaN	NaN			NaN	NaN
4	5	NaN	NaN			NaN	

NaN ...
.. 196 ... ...
...
195 NaN NaN NaN
NaN
196 197 NaN NaN NaN

```
NaN
197             198       NaN     NaN                     NaN
NaN
198             199       NaN     NaN                     NaN
NaN
199             200       NaN     NaN                     NaN
NaN

[200 rows x 5 columns]
data['Age'].unique()
array([19,54,
        29,  21, 20, 23, 31, 22, 35, 64, 30, 67, 58, 24, 37, 52, 25, 46,
51,
        69,  45, 40, 60, 53, 18, 49, 42, 36, 65, 48, 50, 27, 33, 59, 47,
41],        70, 63, 43, 68, 32, 26, 57, 38, 55, 34, 66, 39, 44, 28, 56,
      dtype=int64)

data['Age'].value_counts

<bound method IndexOpsMixin.value_counts of 0       19
1       21
2       20
3       23
4       31
       ..
195     35
196     45
197     32
198     32
199     30


Name: Age, Length: 200, dtype: int64>

# replacing the null values
data['Age'] = data['Age'].fillna(data['Age'].median())data['Age']
0       19
1       21
2       20
3       23
4       31
         ..
195     35
196     45
197     32
198     32
199     30
```

Name: Age, Length: 200, dtype: int64

Find and Replace Outliers

data.describe()

```
       CustomerID         Age  Annual Income (k$)  Spending Score (1-
100)
count  200.000000  200.000000          200.000000
200.000000
mean   100.500000   38.850000           60.560000
50.200000
std     57.879185   13.969007           26.264721
25.823522
min      1.000000   18.000000           15.000000
1.000000
25%     50.750000   28.750000           41.500000
34.750000
50%    100.500000   36.000000           61.500000
50.000000
75%    150.250000   49.000000           78.000000
73.000000
max    200.000000   70.000000          137.000000
99.000000
```

```
data['Age'],data['Annual Income (k$)'].unique()
```

```
(0      19
 1      21
 2      20
 3      23
 4      31
       ..
 195    35
 196    45
 197    32
 198    32
 199    30
 Name: Age, Length: 200, dtype: int64,
 array([ 15,  16,  17,  18,  19,  20,  21,  23,  24,  25,  28,  29,
30,
        33,  34,  37,  38,  39,  40,  42,  43,  44,  46,  47,  48,
49,
        50,  54,  57,  58,  59,  60,  61,  62,  63,  64,  65,  67,
69,
        70,  71,  72,  73,  74,  75,  76,  77,  78,  79,  81,  85,
86,
        87,  88,  93,  97,  98,  99, 101, 103, 113, 120, 126, 137],
```

```
                dtype=int64))

q = data.quantile([0.75,0.50])q
```

	CustomerID	Age	Annual Income (k$)	Spending	Score	(1-100)
0.75	150.25	49.0	78.0			73.0
0.50	100.50	36.0	61.5			50.0

```
iqr = q.iloc[0] - q.iloc[1]
```

```
iqr
```

```
CustomerID                      49.75
Age                             13.00
Annual Income          (k$)     16.50
Spending Score (1-100)dtype:    23.00
float64
```

```
u = q.iloc[0] + (1.5*iqr)       # q3 + 1.5*iqr
```

```
u
```

```
CustomerID                      224.875
Age                              68.500
Annual Income (k$)              102.750
Spending Score (1-100)dtype:    107.500
float64
```

```
l = q.iloc[1] - (1.5*iqr)
```

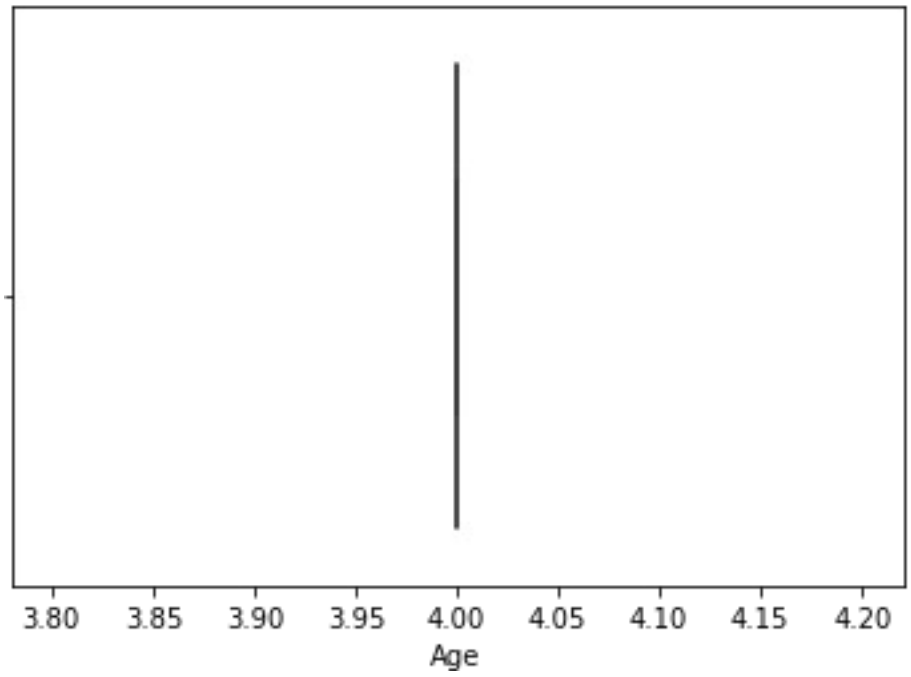
```
l
```

```
CustomerID                      25.875
Age                             16.500
Annual Income (k$)              36.750
Spending Score (1-100)dtype:    15.500
float64
```

```
## handling outliers
```

```
data['Age'] = np.where(data['Age']>5,4,data['Age'])sns.boxplot(data['Age'])
C:\Users\balas\anaconda3\lib\site-packages\seaborn\_decorators.py:36:FutureWarning: Pass the
following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be
`data`, and passing other arguments without an explicit keyword will result in anerror or
misinterpretation.
    warnings.warn(
```

```
<AxesSubplot:xlabel='Age'>
```



Check for Categorical columns and perform encoding.

```
# converting categorical values into numericalvalues - Encodingfrom sklearn.preprocessing
import    LabelEncoder,OneHotEncoderle = LabelEncoder()
oneh = OneHotEncoder()
data['Gender'] = le.fit_transform(data['Gender'])data.head()
```

	CustomerID	Gender	Age	Annual Income (k$)	Spending Score (1-100)
0	1	1	4	15	39
1	2	1	4	15	81
2	3	0	4	16	6
3	4	0	4	16	77
4	5	0	4	17	40

Scaling the data

```
from sklearn.preprocessing import MinMaxScalermm =
MinMaxScaler()
x_scaled = mm.fit_transform(data)x_scaled
        array([[0.                , 1.                , 0.                , 0.                , 0.3877551 ],
               [0.00502513, 1.                , 0.                , 0.                , 0.81632653],
               [0.01005025,  0.                ,  0.                ,  0.00819672,  0.05102041],
               [0.01507538,  0.                ,  0.                ,  0.00819672,  0.7755102 ],
               [0.0201005 ,  0.                ,  0.                ,  0.01639344,  0.39795918],
               [0.02512563,  0.                ,  0.                ,  0.01639344,  0.76530612],
               [0.03015075,  0.                ,  0.                ,  0.02459016,  0.05102041],
```

[0.03517588, 0. , 0. , 0.02459016, 0.94897959],
[0.04020101, 1. , 0. , 0.03278689, 0.02040816],
[0.04522613, 0. , 0. , 0.03278689, 0.7244898],
[0.05025126, 1. , 0. , 0.03278689, 0.13265306],
[0.05527638, 0. , 0. , 0.03278689, 1.],
[0.06030151, 0. , 0. , 0.04098361, 0.14285714],
[0.06532663, 0. , 0. , 0.04098361, 0.7755102],
[0.07035176, 1. , 0. , 0.04098361, 0.12244898],
[0.07537688, 1. , 0. , 0.04098361, 0.79591837],
[0.08040201, 0. , 0. , 0.04918033, 0.34693878],
[0.08542714, 1. , 0. , 0.04918033, 0.66326531],
[0.09045226, 1. , 0. , 0.06557377, 0.28571429],
[0.09547739, 0. , 0. , 0.06557377, 0.98979592],
[0.10050251, 1. , 0. , 0.07377049, 0.34693878],
[0.10552764, 1. , 0. , 0.07377049, 0.73469388],
[0.11055276, 0. , 0. , 0.08196721, 0.04081633],
[0.11557789, 1. , 0. , 0.08196721, 0.73469388],
[0.12060302, 0. , 0. , 0.10655738, 0.13265306],
[0.12562814, 1. , 0. , 0.10655738, 0.82653061],
[0.13065327, 0. , 0. , 0.10655738, 0.31632653],
[0.13567839, 1. , 0. , 0.10655738, 0.6122449],
[0.14070352, 0. , 0. , 0.1147541 , 0.30612245],
[0.14572864, 0. , 0. , 0.1147541 , 0.87755102],
[0.15075377, 1. , 0. , 0.12295082, 0.03061224],
[0.15577889, 0. , 0. , 0.12295082, 0.73469388],
[0.16080402, 1. , 0. , 0.14754098, 0.03061224],
[0.16582915, 1. , 0. , 0.14754098, 0.92857143],
[0.17085427, 0. , 0. , 0.14754098, 0.13265306],
[0.1758794 , 0. , 0. , 0.14754098, 0.81632653],
[0.18090452, 0. , 0. , 0.1557377 , 0.16326531],
[0.18592965, 0. , 0. , 0.1557377 , 0.73469388],
[0.19095477, 0. , 0. , 0.18032787, 0.25510204],
[0.1959799 , 0. , 0. , 0.18032787, 0.75510204],
[0.20100503, 0. , 0. , 0.18852459, 0.34693878],
[0.20603015, 1. , 0. , 0.18852459, 0.92857143],
[0.21105528, 1. , 0. , 0.19672131, 0.35714286],
[0.2160804 , 0. , 0. , 0.19672131, 0.6122449],
[0.22110553, 0. , 0. , 0.19672131, 0.2755102],
[0.22613065, 0. , 0. , 0.19672131, 0.65306122],
[0.23115578, 0. , 0. , 0.20491803, 0.55102041],
[0.2361809 , 0. , 0. , 0.20491803, 0.46938776],
[0.24120603, 0. , 0. , 0.20491803, 0.41836735],

```
[0.24623116, 0.        , 0.        , 0.20491803, 0.41836735],
[0.25125628, 0.        , 0.        , 0.22131148, 0.52040816],
[0.25628141, 1.        , 0.        , 0.22131148, 0.60204082],
[0.26130653, 0.        , 0.        , 0.2295082 , 0.54081633],
[0.26633166, 1.        , 0.        , 0.2295082 , 0.60204082],
[0.27135678, 0.        , 0.        , 0.2295082 , 0.44897959],
[0.27638191, 1.        , 0.        , 0.2295082 , 0.40816327],
[0.28140704, 0.        , 0.        , 0.23770492, 0.5       ],
```

```
[0.28643216,  1.  ,  0.  ,  0.23770492,  0.45918367],
[0.29145729,  0.  ,  0.  ,  0.25409836,  0.51020408],
[0.29648241,  1.  ,  0.  ,  0.25409836,  0.45918367],
[0.30150754,  1.  ,  0.  ,  0.25409836,  0.56122449],
[0.30653266,  1.  ,  0.  ,  0.25409836,  0.55102041],
[0.31155779,  0.  ,  0.  ,  0.26229508,  0.52040816],
[0.31658291,  0.  ,  0.  ,  0.26229508,  0.59183673],
[0.32160804,  1.  ,  0.  ,  0.2704918 ,  0.51020408],
[0.32663317,  1.  ,  0.  ,  0.2704918 ,  0.59183673],
[0.33165829,  0.  ,  0.  ,  0.2704918 ,  0.5       ],
[0.33668342,  0.  ,  0.  ,  0.2704918 ,  0.47959184],
[0.34170854,  1.  ,  0.  ,  0.2704918 ,  0.59183673],
[0.34673367,  0.  ,  0.  ,  0.2704918 ,  0.46938776],
[0.35175879,  1.  ,  0.  ,  0.27868852,  0.55102041],
[0.35678392,  0.  ,  0.  ,  0.27868852,  0.41836735],
[0.36180905,  0.  ,  0.  ,  0.28688525,  0.48979592],
[0.36683417,  0.  ,  0.  ,  0.28688525,  0.56122449],
[0.3718593 ,  1.  ,  0.  ,  0.31967213,  0.46938776],
[0.37688442,  1.  ,  0.  ,  0.31967213,  0.54081633],
[0.38190955,  0.  ,  0.  ,  0.31967213,  0.53061224],
[0.38693467,  1.  ,  0.  ,  0.31967213,  0.47959184],
[0.3919598 ,  0.  ,  0.  ,  0.31967213,  0.52040816],
[0.39698492,  0.  ,  0.  ,  0.31967213,  0.41836735],
[0.40201005,  1.  ,  0.  ,  0.31967213,  0.51020408],
[0.40703518,  1.  ,  0.  ,  0.31967213,  0.55102041],
[0.4120603 ,  1.  ,  0.  ,  0.31967213,  0.40816327],
[0.41708543,  0.  ,  0.  ,  0.31967213,  0.43877551],
[0.42211055,  0.  ,  0.  ,  0.31967213,  0.57142857],
[0.42713568,  1.  ,  0.  ,  0.31967213,  0.45918367],
[0.4321608 ,  0.  ,  0.  ,  0.3442623 ,  0.58163265],
[0.43718593,  0.  ,  0.  ,  0.3442623 ,  0.55102041],
[0.44221106,  0.  ,  0.  ,  0.35245902,  0.60204082],
[0.44723618,  0.  ,  0.  ,  0.35245902,  0.45918367],
[0.45226131,  0.  ,  0.  ,  0.36065574,  0.55102041],
[0.45728643,  1.  ,  0.  ,  0.36065574,  0.40816327],
[0.46231156,  1.  ,  0.  ,  0.36885246,  0.48979592],
[0.46733668,  0.  ,  0.  ,  0.36885246,  0.39795918],
[0.47236181,  0.  ,  0.  ,  0.36885246,  0.41836735],
[0.47738693,  1.  ,  0.  ,  0.36885246,  0.52040816],
[0.48241206,  0.  ,  0.  ,  0.36885246,  0.46938776],
[0.48743719,  0.  ,  0.  ,  0.36885246,  0.5       ],
[0.49246231,  1.  ,  0.  ,  0.37704918,  0.41836735],
```

```
[0.49748744,  1.        ,  0.        ,  0.37704918,  0.48979592],
[0.50251256,  0.        ,  0.        ,  0.3852459 ,  0.40816327],
[0.50753769,  0.        ,  0.        ,  0.3852459 ,  0.47959184],
[0.51256281,  1.        ,  0.        ,  0.3852459 ,  0.59183673],
[0.51758794,  1.        ,  0.        ,  0.3852459 ,  0.55102041],
[0.52261307,  1.        ,  0.        ,  0.3852459 ,  0.56122449],
[0.52763819,  0.        ,  0.        ,  0.3852459 ,  0.41836735],
[0.53266332,  0.        ,  0.        ,  0.39344262,  0.5       ],
```

```
[0.53768844,  1.        ,  0.        ,  0.39344262,  0.45918367],
[0.54271357,  1.        ,  0.        ,  0.39344262,  0.42857143],
[0.54773869,  1.        ,  0.        ,  0.39344262,  0.47959184],
[0.55276382,  1.        ,  0.        ,  0.39344262,  0.52040816],
[0.55778894,  0.        ,  0.        ,  0.39344262,  0.54081633],
[0.56281407,  0.        ,  0.        ,  0.40163934,  0.41836735],
[0.5678392 ,  1.        ,  0.        ,  0.40163934,  0.45918367],
[0.57286432,  0.        ,  0.        ,  0.40983607,  0.47959184],
[0.57788945,  0.        ,  0.        ,  0.40983607,  0.5       ],
[0.58291457,  0.        ,  0.        ,  0.40983607,  0.42857143],
[0.5879397 ,  0.        ,  0.        ,  0.40983607,  0.59183673],
[0.59296482,  0.        ,  0.        ,  0.42622951,  0.42857143],
[0.59798995,  0.        ,  0.        ,  0.42622951,  0.57142857],
[0.60301508,  1.        ,  0.        ,  0.42622951,  0.56122449],
[0.6080402 ,  0.        ,  0.        ,  0.42622951,  0.39795918],
[0.61306533,  0.        ,  0.        ,  0.44262295,  0.58163265],
[0.61809045,  1.        ,  0.        ,  0.44262295,  0.91836735],
[0.62311558,  0.        ,  0.        ,  0.45081967,  0.28571429],
[0.6281407 ,  0.        ,  0.        ,  0.45081967,  0.7755102 ],
[0.63316583,  1.        ,  0.        ,  0.45901639,  0.34693878],
[0.63819095,  1.        ,  0.        ,  0.45901639,  0.95918367],
[0.64321608,  1.        ,  0.        ,  0.45901639,  0.10204082],
[0.64824121,  1.        ,  0.        ,  0.45901639,  0.75510204],
[0.65326633,  1.        ,  0.        ,  0.45901639,  0.08163265],
[0.65829146,  1.        ,  0.        ,  0.45901639,  0.75510204],
[0.66331658,  0.        ,  0.        ,  0.46721311,  0.33673469],
[0.66834171,  0.        ,  0.        ,  0.46721311,  0.71428571],
[0.67336683,  1.        ,  0.        ,  0.47540984,  0.04081633],
[0.67839196,  0.        ,  0.        ,  0.47540984,  0.8877551 ],
[0.68341709,  0.        ,  0.        ,  0.47540984,  0.06122449],
[0.68844221,  1.        ,  0.        ,  0.47540984,  0.73469388],
[0.69346734,  1.        ,  0.        ,  0.48360656,  0.09183673],
[0.69849246,  0.        ,  0.        ,  0.48360656,  0.7244898 ],
[0.70351759,  0.        ,  0.        ,  0.49180328,  0.04081633],
[0.70854271,  1.        ,  0.        ,  0.49180328,  0.93877551],
[0.71356784,  0.        ,  0.        ,  0.5       ,  0.39795918],
[0.71859296,  0.        ,  0.        ,  0.5       ,  0.87755102],
[0.72361809,  1.        ,  0.        ,  0.50819672,  0.1122449 ],
[0.72864322,  1.        ,  0.        ,  0.50819672,  0.97959184],
[0.73366834,  1.        ,  0.        ,  0.50819672,  0.35714286],
[0.73869347,  0.        ,  0.        ,  0.50819672,  0.74489796],
[0.74371859,  0.        ,  0.        ,  0.51639344,  0.21428571],
```

```
[0.74874372,  1.        ,  0.        ,  0.51639344,  0.90816327],
[0.75376884,  1.        ,  0.        ,  0.51639344,  0.16326531],
[0.75879397,  1.        ,  0.        ,  0.51639344,  0.8877551 ],
[0.7638191 ,  0.        ,  0.        ,  0.51639344,  0.19387755],
[0.76884422,  0.        ,  0.        ,  0.51639344,  0.76530612],
[0.77386935,  0.        ,  0.        ,  0.51639344,  0.15306122],
[0.77889447,  0.        ,  0.        ,  0.51639344,  0.89795918],
[0.7839196 ,  1.        ,  0.        ,  0.51639344,  0.        ],
```

[0.78894472, 0. , 0. , 0.51639344, 0.78571429],
[0.79396985, 1. , 0. , 0.51639344, 0.],
[0.79899497, 0. , 0. , 0.51639344, 0.73469388],
[0.8040201 , 0. , 0. , 0.52459016, 0.34693878],
[0.80904523, 0. , 0. , 0.52459016, 0.83673469],
[0.81407035, 1. , 0. , 0.54098361, 0.04081633],
[0.81909548, 0. , 0. , 0.54098361, 0.93877551],
[0.8241206 , 1. , 0. , 0.57377049, 0.25510204],
[0.82914573, 0. , 0. , 0.57377049, 0.75510204],
[0.83417085, 1. , 0. , 0.58196721, 0.19387755],
[0.83919598, 0. , 0. , 0.58196721, 0.95918367],
[0.84422111, 0. , 0. , 0.59016393, 0.26530612],
[0.84924623, 1. , 0. , 0.59016393, 0.63265306],
[0.85427136, 1. , 0. , 0.59016393, 0.12244898],
[0.85929648, 1. , 0. , 0.59016393, 0.75510204],
[0.86432161, 1. , 0. , 0.59016393, 0.09183673],
[0.86934673, 1. , 0. , 0.59016393, 0.92857143],
[0.87437186, 0. , 0. , 0.59836066, 0.12244898],
[0.87939698, 0. , 0. , 0.59836066, 0.86734694],
[0.88442211, 1. , 0. , 0.59836066, 0.14285714],
[0.88944724, 1. , 0. , 0.59836066, 0.69387755],
[0.89447236, 1. , 0. , 0.63934426, 0.13265306],
[0.89949749, 1. , 0. , 0.63934426, 0.90816327],
[0.90452261, 0. , 0. , 0.67213115, 0.31632653],
[0.90954774, 0. , 0. , 0.67213115, 0.86734694],
[0.91457286, 1. , 0. , 0.68032787, 0.14285714],
[0.91959799, 0. , 0. , 0.68032787, 0.8877551],
[0.92462312, 0. , 0. , 0.68852459, 0.3877551],
[0.92964824, 1. , 0. , 0.68852459, 0.97959184],
[0.93467337, 0. , 0. , 0.70491803, 0.23469388],
[0.93969849, 1. , 0. , 0.70491803, 0.68367347],
[0.94472362, 0. , 0. , 0.72131148, 0.16326531],
[0.94974874, 0. , 0. , 0.72131148, 0.85714286],
[0.95477387, 0. , 0. , 0.72131148, 0.2244898],
[0.95979899, 0. , 0. , 0.72131148, 0.69387755],
[0.96482412, 1. , 0. , 0.80327869, 0.07142857],
[0.96984925, 0. , 0. , 0.80327869, 0.91836735],
[0.97487437, 0. , 0. , 0.86065574, 0.15306122],
[0.9798995 , 0. , 0. , 0.86065574, 0.79591837],
[0.98492462, 0. , 0. , 0.90983607, 0.2755102],
[0.98994975, 1. , 0. , 0.90983607, 0.74489796],
[0.99497487, 1. , 0. , 1. , 0.17346939],

[1. , 1. , 0. , 1. , 0.83673469]])

Perform any of the clustering algorithms

```
from sklearn.cluster import KMeanskm =
KMeans()
clus = km.fit_predict(x_scaled)clus
```

```
array([4, 4, 1, 1, 1, 1, 1, 1, 4, 1, 4, 1, 1, 1, 4, 4, 1, 4, 4, 1, 4,
4,
       1, 4, 1, 4, 1, 4, 1, 1, 4, 1, 4, 4, 1, 1, 1, 1, 1, 1, 1, 4, 4,
1,
       1, 1, 1, 1, 1, 1, 1, 4, 1, 4, 1, 4, 1, 3, 1, 3, 3, 3, 2, 2, 3,
3,
       2, 2, 3, 2, 3, 2, 2, 2, 3, 3, 2, 3, 2, 2, 3, 3, 3, 2, 2, 3, 2,
2,
       2, 2, 2, 3, 3, 2, 2, 3, 2, 2, 3, 3, 2, 2, 3, 3, 3, 2, 2, 3, 3,
3,
       3, 2, 2, 3, 2, 2, 2, 2, 2, 2, 3, 2, 2, 7, 2, 5, 3, 7, 0, 7, 0,
7,
       2, 5, 0, 5, 6, 7, 0, 5, 6, 7, 6, 5, 0, 7, 0, 5, 6, 7, 0, 7, 6,
5,
       6, 5, 0, 5, 0, 5, 6, 5, 0, 5, 0, 5, 0, 5, 6, 7, 0, 7, 0, 7, 6,
5,
       0, 7, 0, 7, 6, 5, 0, 5, 6, 7, 6, 7, 6, 5, 6, 5, 0, 5, 6, 5, 6,
7,
       0, 7])
```

```
names = data.columnsnames
Index(['CustomerID', 'Gender', 'Age', 'Annual Income (k$)','Spending Score (1-100)'],
          dtype='object')
```

```
data1 = pd.DataFrame(x_scaled, columns=names)data1.head()
    CustomerID   Gender   Age   Annual Income (k$)   Spending Score (1-100)0                0.000000
       1.0   0.0            0.000000                    0.387755
1      0.005025           1.0   0.0                    0.000000                    0.816327
2      0.010050           0.0   0.0                    0.008197                    0.051020
3      0.015075           0.0   0.0                    0.008197                    0.775510
4      0.020101           0.0   0.0                    0.016393                    0.397959
```

Add the cluster data with the primary dataset

```
data1['New clus'] = pd.Series(clus)data1.head()
```

```
   CustomerID  Gender  Age  Annual Income (k$)  Spending Score (1-100)
\
0    0.000000     1.0  0.0            0.000000                0.387755

1    0.005025     1.0  0.0            0.000000                0.816327

2    0.010050     0.0  0.0            0.008197                0.051020
```

3	0.015075	0.0	0.0	0.008197	0.775510
4	0.020101	0.0	0.0	0.016393	0.397959

	New clus
0	4
1	4
2	1
3	1
4	1

Split the data into dependent and independent variables

```
data.shape
(200, 5)
x = data.iloc[:, 0:5]x
```

	CustomerID	Gender	Age	Annual Income (k$)	Spending Score (1-100)
0	1	1	4	15	39
1	2	1	4	15	81
2	3	0	4	16	6
3	4	0	4	16	77
4	5	0	4	17	40
..	...	...	...	...	. ..
195	196	0	4	120	79
196	197	0	4	126	28
197	198	1	4	126	74
198	199	1	4	137	18
199	200	1	4	137	83

[200 rows x 5 columns]

```
y = data['Spending Score (1-100)']y
```

```
0      39
1      81
2       6
3      77
4      40
       ..
195    79
196    28
197    74
198    18
199    83
Name:  Spending Score (1-100), Length: 200, dtype: int64
```

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_scaled, y,test_size = 0.5, random_state = 0)
```

```
x_train
```

```
array([[0.79396985, 1.        , 0.        , 0.51639344, 0.        ],
       [0.95979899, 0.        , 0.        , 0.72131148, 0.69387755],
       [0.25125628, 0.        , 0.        , 0.22131148, 0.52040816],
       [0.        , 1.        , 0.        , 0.        , 0.3877551 ],
       [0.47236181, 0.        , 0.        , 0.36885246, 0.41836735],
       [0.55276382, 1.        , 0.        , 0.39344262, 0.52040816],
       [0.47738693, 1.        , 0.        , 0.36885246, 0.52040816],
       [0.32160804, 1.        , 0.        , 0.2704918 , 0.51020408],
       [0.83919598, 0.        , 0.        , 0.58196721, 0.95918367],
       [0.20603015, 1.        , 0.        , 0.18852459, 0.92857143],
       [0.34673367, 0.        , 0.        , 0.2704918 , 0.46938776],
       [0.24623116, 0.        , 0.        , 0.20491803, 0.41836735],
       [0.24120603, 0.        , 0.        , 0.20491803, 0.41836735],
       [0.42713568, 1.        , 0.        , 0.31967213, 0.45918367],
       [0.06532663, 0.        , 0.        , 0.04098361, 0.7755102 ],
       [0.80904523, 0.        , 0.        , 0.52459016, 0.83673469],
       [0.11557789, 1.        , 0.        , 0.08196721, 0.73469388],
       [0.93467337, 0.        , 0.        , 0.70491803, 0.23469388],
       [0.67839196, 0.        , 0.        , 0.47540984, 0.8877551 ],
       [0.10050251, 1.        , 0.        , 0.07377049, 0.34693878],
       [0.07537688, 1.        , 0.        , 0.04098361, 0.79591837],
       [0.3919598 , 0.        , 0.        , 0.31967213, 0.52040816],
       [0.52261307, 1.        , 0.        , 0.3852459 , 0.56122449],
       [0.26130653, 0.        , 0.        , 0.2295082 , 0.54081633],
       [0.50251256, 0.        , 0.        , 0.3852459 , 0.40816327],
       [0.38190955, 0.        , 0.        , 0.31967213, 0.53061224],
```

[0.01507538, 0. , 0. , 0.00819672, 0.7755102],
[0.58291457, 0. , 0. , 0.40983607, 0.42857143],
[0.8241206 , 1. , 0. , 0.57377049, 0.25510204],
[0.99497487, 1. , 0. , 1. , 0.17346939],
[0.03015075, 0. , 0. , 0.02459016, 0.05102041],

```
[0.34170854,  1.        ,  0.        ,  0.2704918 ,  0.59183673],
[0.42211055,  0.        ,  0.        ,  0.31967213,  0.57142857],
[0.6080402 ,  0.        ,  0.        ,  0.42622951,  0.39795918],
[0.77889447,  0.        ,  0.        ,  0.51639344,  0.89795918],
[0.85929648,  1.        ,  0.        ,  0.59016393,  0.75510204],
[0.7839196 ,  1.        ,  0.        ,  0.51639344,  0.        ],
[0.45728643,  1.        ,  0.        ,  0.36065574,  0.40816327],
[1.        ,  1.        ,  0.        ,  1.        ,  0.83673469],
[0.05527638,  0.        ,  0.        ,  0.03278689,  1.        ],
[0.59798995,  0.        ,  0.        ,  0.42622951,  0.57142857],
[0.51256281,  1.        ,  0.        ,  0.3852459 ,  0.59183673],
[0.1758794 ,  0.        ,  0.        ,  0.14754098,  0.81632653],
[0.28643216,  1.        ,  0.        ,  0.23770492,  0.45918367],
[0.32663317,  1.        ,  0.        ,  0.2704918 ,  0.59183673],
[0.00502513,  1.        ,  0.        ,  0.        ,  0.81632653],
[0.60301508,  1.        ,  0.        ,  0.42622951,  0.56122449],
[0.81407035,  1.        ,  0.        ,  0.54098361,  0.04081633],
[0.21105528,  1.        ,  0.        ,  0.19672131,  0.35714286],
[0.52763819,  0.        ,  0.        ,  0.3852459 ,  0.41836735],
[0.66331658,  0.        ,  0.        ,  0.46721311,  0.33673469],
[0.86934673,  1.        ,  0.        ,  0.59016393,  0.92857143],
[0.08542714,  1.        ,  0.        ,  0.04918033,  0.66326531],
[0.19095477,  0.        ,  0.        ,  0.18032787,  0.25510204],
[0.66834171,  0.        ,  0.        ,  0.46721311,  0.71428571],
[0.26633166,  1.        ,  0.        ,  0.2295082 ,  0.60204082],
[0.78894472,  0.        ,  0.        ,  0.51639344,  0.78571429],
[0.64321608,  1.        ,  0.        ,  0.45901639,  0.10204082],
[0.17085427,  0.        ,  0.        ,  0.14754098,  0.13265306],
[0.14070352,  0.        ,  0.        ,  0.1147541 ,  0.30612245],
[0.57286432,  0.        ,  0.        ,  0.40983607,  0.47959184],
[0.75879397,  1.        ,  0.        ,  0.51639344,  0.8877551 ],
[0.15577889,  0.        ,  0.        ,  0.12295082,  0.73469388],
[0.83417085,  1.        ,  0.        ,  0.58196721,  0.19387755],
[0.63819095,  1.        ,  0.        ,  0.45901639,  0.95918367],
[0.88442211,  1.        ,  0.        ,  0.59836066,  0.14285714],
[0.16080402,  1.        ,  0.        ,  0.14754098,  0.03061224],
[0.71356784,  0.        ,  0.        ,  0.5       ,  0.39795918],
[0.84924623,  1.        ,  0.        ,  0.59016393,  0.63265306],
[0.73869347,  0.        ,  0.        ,  0.50819672,  0.74489796],
[0.14572864,  0.        ,  0.        ,  0.1147541 ,  0.87755102],
[0.49748744,  1.        ,  0.        ,  0.37704918,  0.48979592],
[0.4120603 ,  1.        ,  0.        ,  0.31967213,  0.40816327],
```

```
[0.39698492, 0.        , 0.        , 0.31967213, 0.41836735],
[0.57788945, 0.        , 0.        , 0.40983607, 0.5       ],
[0.74371859, 0.        , 0.        , 0.51639344, 0.21428571],
[0.96984925, 0.        , 0.        , 0.80327869, 0.91836735],
[0.36180905, 0.        , 0.        , 0.28688525, 0.48979592],
[0.38693467, 1.        , 0.        , 0.31967213, 0.47959184],
[0.12562814, 1.        , 0.        , 0.10655738, 0.82653061],
[0.82914573, 0.        , 0.        , 0.57377049, 0.75510204],
```

```
       [0.40703518,  1.        ,  0.        ,  0.31967213,  0.55102041],
       [0.94472362,  0.        ,  0.        ,  0.72131148,  0.16326531],
       [0.87437186,  0.        ,  0.        ,  0.59836066,  0.12244898],
       [0.95477387,  0.        ,  0.        ,  0.72131148,  0.2244898 ],
       [0.1959799 ,  0.        ,  0.        ,  0.18032787,  0.75510204],
       [0.29145729,  0.        ,  0.        ,  0.25409836,  0.51020408],
       [0.70351759,  0.        ,  0.        ,  0.49180328,  0.04081633],
       [0.44221106,  0.        ,  0.        ,  0.35245902,  0.60204082],
       [0.35175879,  1.        ,  0.        ,  0.27868852,  0.55102041],
       [0.43718593,  0.        ,  0.        ,  0.3442623 ,  0.55102041],
       [0.18090452,  0.        ,  0.        ,  0.1557377 ,  0.16326531],
       [0.10552764,  1.        ,  0.        ,  0.07377049,  0.73469388],
       [0.04522613,  0.        ,  0.        ,  0.03278689,  0.7244898 ],
       [0.51758794,  1.        ,  0.        ,  0.3852459 ,  0.55102041],
       [0.33668342,  0.        ,  0.        ,  0.2704918 ,  0.47959184],
       [0.96482412,  1.        ,  0.        ,  0.80327869,  0.07142857],
       [0.5879397 ,  0.        ,  0.        ,  0.40983607,  0.59183673],
       [0.2361809 ,  0.        ,  0.        ,  0.20491803,  0.46938776],
       [0.86432161,  1.        ,  0.        ,  0.59016393,  0.09183673]])
```

x_train.shape

(100, 5)

x_test

```
array([[0.09045226,  1.        ,  0.        ,  0.06557377,  0.28571429],
       [0.85427136,  1.        ,  0.        ,  0.59016393,  0.12244898],
       [0.53768844,  1.        ,  0.        ,  0.39344262,  0.45918367],
       [0.49246231,  1.        ,  0.        ,  0.37704918,  0.41836735],
       [0.88944724,  1.        ,  0.        ,  0.59836066,  0.69387755],
       [0.91457286,  1.        ,  0.        ,  0.68032787,  0.14285714],
       [0.02512563,  0.        ,  0.        ,  0.01639344,  0.76530612],
       [0.73366834,  1.        ,  0.        ,  0.50819672,  0.35714286],
       [0.06030151,  0.        ,  0.        ,  0.04098361,  0.14285714],
       [0.7638191 ,  0.        ,  0.        ,  0.51639344,  0.19387755],
       [0.30653266,  1.        ,  0.        ,  0.25409836,  0.55102041],
       [0.6281407 ,  0.        ,  0.        ,  0.45081967,  0.7755102 ],
       [0.90452261,  0.        ,  0.        ,  0.67213115,  0.31632653],
       [0.77386935,  0.        ,  0.        ,  0.51639344,  0.15306122],
       [0.40201005,  1.        ,  0.        ,  0.31967213,  0.51020408],
       [0.03517588,  0.        ,  0.        ,  0.02459016,  0.94897959],
       [0.16582915,  1.        ,  0.        ,  0.14754098,  0.92857143],
       [0.65326633,  1.        ,  0.        ,  0.45901639,  0.08163265],
       [0.18592965,  0.        ,  0.        ,  0.1557377 ,  0.73469388],
```

```
 [0.3718593 ,  1.        ,  0.        ,  0.31967213,  0.46938776],
[0.91959799,  0.        ,  0.        ,  0.68032787,  0.8877551 ],
[0.72864322,  1.        ,  0.        ,  0.50819672,  0.97959184],
[0.22613065,  0.        ,  0.        ,  0.19672131,  0.65306122],
[0.79899497,  0.        ,  0.        ,  0.51639344,  0.73469388],
[0.30150754,  1.        ,  0.        ,  0.25409836,  0.56122449],
```

```
[0.61809045,  1.        ,  0.        ,  0.44262295,  0.91836735],
[0.89949749,  1.        ,  0.        ,  0.63934426,  0.90816327],
[0.92964824,  1.        ,  0.        ,  0.68852459,  0.97959184],
[0.61306533,  0.        ,  0.        ,  0.44262295,  0.58163265],
[0.22110553,  0.        ,  0.        ,  0.19672131,  0.2755102 ],
[0.08040201,  0.        ,  0.        ,  0.04918033,  0.34693878],
[0.27638191,  1.        ,  0.        ,  0.2295082 ,  0.40816327],
[0.75376884,  1.        ,  0.        ,  0.51639344,  0.16326531],
[0.55778894,  0.        ,  0.        ,  0.39344262,  0.54081633],
[0.11055276,  0.        ,  0.        ,  0.08196721,  0.04081633],
[0.94974874,  0.        ,  0.        ,  0.72131148,  0.85714286],
[0.64824121,  1.        ,  0.        ,  0.45901639,  0.75510204],
[0.0201005 ,  0.        ,  0.        ,  0.01639344,  0.39795918],
[0.41708543,  0.        ,  0.        ,  0.31967213,  0.43877551],
[0.53266332,  0.        ,  0.        ,  0.39344262,  0.5       ],
[0.67336683,  1.        ,  0.        ,  0.47540984,  0.04081633],
[0.33165829,  0.        ,  0.        ,  0.2704918 ,  0.5       ],
[0.13065327,  0.        ,  0.        ,  0.10655738,  0.31632653],
[0.5678392 ,  1.        ,  0.        ,  0.40163934,  0.45918367],
[0.84422111,  0.        ,  0.        ,  0.59016393,  0.26530612],
[0.31658291,  0.        ,  0.        ,  0.26229508,  0.59183673],
[0.04020101,  1.        ,  0.        ,  0.03278689,  0.02040816],
[0.37688442,  1.        ,  0.        ,  0.31967213,  0.54081633],
[0.59296482,  0.        ,  0.        ,  0.42622951,  0.42857143],
[0.71859296,  0.        ,  0.        ,  0.5       ,  0.87755102],
[0.35678392,  0.        ,  0.        ,  0.27868852,  0.41836735],
[0.62311558,  0.        ,  0.        ,  0.45081967,  0.28571429],
[0.92462312,  0.        ,  0.        ,  0.68852459,  0.3877551 ],
[0.48743719,  0.        ,  0.        ,  0.36885246,  0.5       ],
[0.74874372,  1.        ,  0.        ,  0.51639344,  0.90816327],
[0.12060302,  0.        ,  0.        ,  0.10655738,  0.13265306],
[0.15075377,  1.        ,  0.        ,  0.12295082,  0.03061224],
[0.8040201 ,  0.        ,  0.        ,  0.52459016,  0.34693878],
[0.20100503,  0.        ,  0.        ,  0.18852459,  0.34693878],
[0.28140704,  0.        ,  0.        ,  0.23770492,  0.5       ],
[0.65829146,  1.        ,  0.        ,  0.45901639,  0.75510204],
[0.48241206,  0.        ,  0.        ,  0.36885246,  0.46938776],
[0.90954774,  0.        ,  0.        ,  0.67213115,  0.86734694],
[0.09547739,  0.        ,  0.        ,  0.06557377,  0.98979592],
[0.76884422,  0.        ,  0.        ,  0.51639344,  0.76530612],
[0.46231156,  1.        ,  0.        ,  0.36885246,  0.48979592],
[0.27135678,  0.        ,  0.        ,  0.2295082 ,  0.44897959],
```

```
[0.81909548,  0.  ,  0.  ,  0.54098361,  0.93877551],
[0.25628141,  1.  ,  0.  ,  0.22131148,  0.60204082],
[0.4321608 ,  0.  ,  0.  ,  0.3442623 ,  0.58163265],
[0.69849246,  0.  ,  0.  ,  0.48360656,  0.7244898 ],
[0.45226131,  0.  ,  0.  ,  0.36065574,  0.55102041],
[0.68844221,  1.  ,  0.  ,  0.47540984,  0.73469388],
[0.50753769,  0.  ,  0.  ,  0.3852459 ,  0.47959184],
[0.72361809,  1.  ,  0.  ,  0.50819672,  0.1122449 ],
```

```
       [0.44723618,  0.        ,  0.        ,  0.35245902,  0.45918367],
       [0.54773869,  1.        ,  0.        ,  0.39344262,  0.47959184],
       [0.07035176,  1.        ,  0.        ,  0.04098361,  0.12244898],
       [0.13567839,  1.        ,  0.        ,  0.10655738,  0.6122449 ],
       [0.70854271,  1.        ,  0.        ,  0.49180328,  0.93877551],
       [0.93969849,  1.        ,  0.        ,  0.70491803,  0.68367347],
       [0.23115578,  0.        ,  0.        ,  0.20491803,  0.55102041],
       [0.69346734,  1.        ,  0.        ,  0.48360656,  0.09183673],
       [0.9798995 ,  0.        ,  0.        ,  0.86065574,  0.79591837],
       [0.54271357,  1.        ,  0.        ,  0.39344262,  0.42857143],
       [0.31155779,  0.        ,  0.        ,  0.26229508,  0.52040816],
       [0.01005025,  0.        ,  0.        ,  0.00819672,  0.05102041],
       [0.29648241,  1.        ,  0.        ,  0.25409836,  0.45918367],
       [0.68341709,  0.        ,  0.        ,  0.47540984,  0.06122449],
       [0.98994975,  1.        ,  0.        ,  0.90983607,  0.74489796],
       [0.2160804 ,  0.        ,  0.        ,  0.19672131,  0.6122449 ],
       [0.05025126,  1.        ,  0.        ,  0.03278689,  0.13265306],
       [0.97487437,  0.        ,  0.        ,  0.86065574,  0.15306122],
       [0.36683417,  0.        ,  0.        ,  0.28688525,  0.56122449],
       [0.98492462,  0.        ,  0.        ,  0.90983607,  0.2755102 ],
       [0.89447236,  1.        ,  0.        ,  0.63934426,  0.13265306],
       [0.87939698,  0.        ,  0.        ,  0.59836066,  0.86734694],
       [0.63316583,  1.        ,  0.        ,  0.45901639,  0.34693878],
       [0.46733668,  0.        ,  0.        ,  0.36885246,  0.39795918],
       [0.56281407,  0.        ,  0.        ,  0.40163934,  0.41836735]])
```

```
x_test.shape(100,
5)
y_train
```

```
158      1
191     69
50      52
0       39
94      42
       ..
67      48
192      8
117     59
47      47
172     10
Name: Spending Score (1-100), Length: 100, dtype: int64
```

```
y_train.shape
```

(100,)
y_test

```
18      29
170     13
107     46
98      42
177     69
        ..
178     14
175     86
126     35
93      40
112     42
Name: Spending Score (1-100), Length: 100, dtype: int64y_test.shape
(100,)
```

Build the Model

```
from sklearn.ensemble import RandomForestRegressor
regr = RandomForestRegressor(max_depth=2, random_state=0,
                                                    n_estimators=100)
regr.fit(x_train, y_train)regr.fit(x_test,
y_test)
RandomForestRegressor(max_depth=2, random_state=0)y_train_pred =
regr.predict(x_train)
y_test_pred = regr.predict(x_test)

regr.score(x_train, y_train)
0.9752947343811537
regr.score(x_test, y_test)
0.9790787065797374
```

Train and Test

```
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split (x_scaled, y,test_size = 0.3, random_state = 0)

x_train

        array([[0.65829146, 1.        , 0.        , 0.45901639, 0.75510204],
               [0.48241206, 0.        , 0.        , 0.36885246, 0.46938776],
               [0.90954774, 0.        , 0.        , 0.67213115, 0.86734694],
               [0.09547739, 0.        , 0.        , 0.06557377, 0.98979592],
               [0.76884422, 0.        , 0.        , 0.51639344, 0.76530612],
               [0.46231156, 1.        , 0.        , 0.36885246, 0.48979592],
               [0.27135678, 0.        , 0.        , 0.2295082 , 0.44897959],
               [0.81909548, 0.        , 0.        , 0.54098361, 0.93877551],
```

```
[0.25628141, 1.        , 0.        , 0.22131148, 0.60204082],
[0.4321608 , 0.        , 0.        , 0.3442623 , 0.58163265],
[0.69849246, 0.        , 0.        , 0.48360656, 0.7244898 ],
[0.45226131, 0.        , 0.        , 0.36065574, 0.55102041],
[0.68844221, 1.        , 0.        , 0.47540984, 0.73469388],
[0.50753769, 0.        , 0.        , 0.3852459 , 0.47959184],
[0.72361809, 1.        , 0.        , 0.50819672, 0.1122449 ],
[0.44723618, 0.        , 0.        , 0.35245902, 0.45918367],
[0.54773869, 1.        , 0.        , 0.39344262, 0.47959184],
[0.07035176, 1.        , 0.        , 0.04098361, 0.12244898],
[0.13567839, 1.        , 0.        , 0.10655738, 0.6122449 ],
[0.70854271, 1.        , 0.        , 0.49180328, 0.93877551],
[0.93969849, 1.        , 0.        , 0.70491803, 0.68367347],
[0.23115578, 0.        , 0.        , 0.20491803, 0.55102041],
[0.69346734, 1.        , 0.        , 0.48360656, 0.09183673],
[0.9798995 , 0.        , 0.        , 0.86065574, 0.79591837],
[0.54271357, 1.        , 0.        , 0.39344262, 0.42857143],
[0.31155779, 0.        , 0.        , 0.26229508, 0.52040816],
[0.01005025, 0.        , 0.        , 0.00819672, 0.05102041],
[0.29648241, 1.        , 0.        , 0.25409836, 0.45918367],
[0.68341709, 0.        , 0.        , 0.47540984, 0.06122449],
[0.98994975, 1.        , 0.        , 0.90983607, 0.74489796],
[0.2160804 , 0.        , 0.        , 0.19672131, 0.6122449 ],
[0.05025126, 1.        , 0.        , 0.03278689, 0.13265306],
[0.97487437, 0.        , 0.        , 0.86065574, 0.15306122],
[0.36683417, 0.        , 0.        , 0.28688525, 0.56122449],
[0.98492462, 0.        , 0.        , 0.90983607, 0.2755102 ],
[0.89447236, 1.        , 0.        , 0.63934426, 0.13265306],
[0.87939698, 0.        , 0.        , 0.59836066, 0.86734694],
[0.63316583, 1.        , 0.        , 0.45901639, 0.34693878],
[0.46733668, 0.        , 0.        , 0.36885246, 0.39795918],
[0.56281407, 0.        , 0.        , 0.40163934, 0.41836735],
[0.79396985, 1.        , 0.        , 0.51639344, 0.        ],
[0.95979899, 0.        , 0.        , 0.72131148, 0.69387755],
[0.25125628, 0.        , 0.        , 0.22131148, 0.52040816],
[0.        , 1.        , 0.        , 0.        , 0.3877551 ],
[0.47236181, 0.        , 0.        , 0.36885246, 0.41836735],
[0.55276382, 1.        , 0.        , 0.39344262, 0.52040816],
[0.47738693, 1.        , 0.        , 0.36885246, 0.52040816],
[0.32160804, 1.        , 0.        , 0.2704918 , 0.51020408],
[0.83919598, 0.        , 0.        , 0.58196721, 0.95918367],
[0.20603015, 1.        , 0.        , 0.18852459, 0.92857143],
```

```
[0.34673367,  0.        ,  0.        ,  0.2704918 ,  0.46938776],
[0.24623116,  0.        ,  0.        ,  0.20491803,  0.41836735],
[0.24120603,  0.        ,  0.        ,  0.20491803,  0.41836735],
[0.42713568,  1.        ,  0.        ,  0.31967213,  0.45918367],
[0.06532663,  0.        ,  0.        ,  0.04098361,  0.7755102 ],
[0.80904523,  0.        ,  0.        ,  0.52459016,  0.83673469],
[0.11557789,  1.        ,  0.        ,  0.08196721,  0.73469388],
[0.93467337,  0.        ,  0.        ,  0.70491803,  0.23469388],
```

```
[0.67839196, 0.        , 0.        , 0.47540984, 0.8877551 ],
[0.10050251, 1.        , 0.        , 0.07377049, 0.34693878],
[0.07537688, 1.        , 0.        , 0.04098361, 0.79591837],
[0.3919598 , 0.        , 0.        , 0.31967213, 0.52040816],
[0.52261307, 1.        , 0.        , 0.3852459 , 0.56122449],
[0.26130653, 0.        , 0.        , 0.2295082 , 0.54081633],
[0.50251256, 0.        , 0.        , 0.3852459 , 0.40816327],
[0.38190955, 0.        , 0.        , 0.31967213, 0.53061224],
[0.01507538, 0.        , 0.        , 0.00819672, 0.7755102 ],
[0.58291457, 0.        , 0.        , 0.40983607, 0.42857143],
[0.8241206 , 1.        , 0.        , 0.57377049, 0.25510204],
[0.99497487, 1.        , 0.        , 1.        , 0.17346939],
[0.03015075, 0.        , 0.        , 0.02459016, 0.05102041],
[0.34170854, 1.        , 0.        , 0.2704918 , 0.59183673],
[0.42211055, 0.        , 0.        , 0.31967213, 0.57142857],
[0.6080402 , 0.        , 0.        , 0.42622951, 0.39795918],
[0.77889447, 0.        , 0.        , 0.51639344, 0.89795918],
[0.85929648, 1.        , 0.        , 0.59016393, 0.75510204],
[0.7839196 , 1.        , 0.        , 0.51639344, 0.        ],
[0.45728643, 1.        , 0.        , 0.36065574, 0.40816327],
[1.        , 1.        , 0.        , 1.        , 0.83673469],
[0.05527638, 0.        , 0.        , 0.03278689, 1.        ],
[0.59798995, 0.        , 0.        , 0.42622951, 0.57142857],
[0.51256281, 1.        , 0.        , 0.3852459 , 0.59183673],
[0.1758794 , 0.        , 0.        , 0.14754098, 0.81632653],
[0.28643216, 1.        , 0.        , 0.23770492, 0.45918367],
[0.32663317, 1.        , 0.        , 0.2704918 , 0.59183673],
[0.00502513, 1.        , 0.        , 0.        , 0.81632653],
[0.60301508, 1.        , 0.        , 0.42622951, 0.56122449],
[0.81407035, 1.        , 0.        , 0.54098361, 0.04081633],
[0.21105528, 1.        , 0.        , 0.19672131, 0.35714286],
[0.52763819, 0.        , 0.        , 0.3852459 , 0.41836735],
[0.66331658, 0.        , 0.        , 0.46721311, 0.33673469],
[0.86934673, 1.        , 0.        , 0.59016393, 0.92857143],
[0.08542714, 1.        , 0.        , 0.04918033, 0.66326531],
[0.19095477, 0.        , 0.        , 0.18032787, 0.25510204],
[0.66834171, 0.        , 0.        , 0.46721311, 0.71428571],
[0.26633166, 1.        , 0.        , 0.2295082 , 0.60204082],
[0.78894472, 0.        , 0.        , 0.51639344, 0.78571429],
[0.64321608, 1.        , 0.        , 0.45901639, 0.10204082],
[0.17085427, 0.        , 0.        , 0.14754098, 0.13265306],
[0.14070352, 0.        , 0.        , 0.1147541 , 0.30612245],
```

```
[0.57286432, 0.        , 0.        , 0.40983607, 0.47959184],
[0.75879397, 1.        , 0.        , 0.51639344, 0.8877551 ],
[0.15577889, 0.        , 0.        , 0.12295082, 0.73469388],
[0.83417085, 1.        , 0.        , 0.58196721, 0.19387755],
[0.63819095, 1.        , 0.        , 0.45901639, 0.95918367],
[0.88442211, 1.        , 0.        , 0.59836066, 0.14285714],
[0.16080402, 1.        , 0.        , 0.14754098, 0.03061224],
[0.71356784, 0.        , 0.        , 0.5       , 0.39795918],
```

```
       [0.84924623,  1.        ,  0.        ,  0.59016393,  0.63265306],
       [0.73869347,  0.        ,  0.        ,  0.50819672,  0.74489796],
       [0.14572864,  0.        ,  0.        ,  0.1147541 ,  0.87755102],
       [0.49748744,  1.        ,  0.        ,  0.37704918,  0.48979592],
       [0.4120603 ,  1.        ,  0.        ,  0.31967213,  0.40816327],
       [0.39698492,  0.        ,  0.        ,  0.31967213,  0.41836735],
       [0.57788945,  0.        ,  0.        ,  0.40983607,  0.5       ],
       [0.74371859,  0.        ,  0.        ,  0.51639344,  0.21428571],
       [0.96984925,  0.        ,  0.        ,  0.80327869,  0.91836735],
       [0.36180905,  0.        ,  0.        ,  0.28688525,  0.48979592],
       [0.38693467,  1.        ,  0.        ,  0.31967213,  0.47959184],
       [0.12562814,  1.        ,  0.        ,  0.10655738,  0.82653061],
       [0.82914573,  0.        ,  0.        ,  0.57377049,  0.75510204],
       [0.40703518,  1.        ,  0.        ,  0.31967213,  0.55102041],
       [0.94472362,  0.        ,  0.        ,  0.72131148,  0.16326531],
       [0.87437186,  0.        ,  0.        ,  0.59836066,  0.12244898],
       [0.95477387,  0.        ,  0.        ,  0.72131148,  0.2244898 ],
       [0.1959799 ,  0.        ,  0.        ,  0.18032787,  0.75510204],
       [0.29145729,  0.        ,  0.        ,  0.25409836,  0.51020408],
       [0.70351759,  0.        ,  0.        ,  0.49180328,  0.04081633],
       [0.44221106,  0.        ,  0.        ,  0.35245902,  0.60204082],
       [0.35175879,  1.        ,  0.        ,  0.27868852,  0.55102041],
       [0.43718593,  0.        ,  0.        ,  0.3442623 ,  0.55102041],
       [0.18090452,  0.        ,  0.        ,  0.1557377 ,  0.16326531],
       [0.10552764,  1.        ,  0.        ,  0.07377049,  0.73469388],
       [0.04522613,  0.        ,  0.        ,  0.03278689,  0.7244898 ],
       [0.51758794,  1.        ,  0.        ,  0.3852459 ,  0.55102041],
       [0.33668342,  0.        ,  0.        ,  0.2704918 ,  0.47959184],
       [0.96482412,  1.        ,  0.        ,  0.80327869,  0.07142857],
       [0.5879397 ,  0.        ,  0.        ,  0.40983607,  0.59183673],
       [0.2361809 ,  0.        ,  0.        ,  0.20491803,  0.46938776],
       [0.86432161,  1.        ,  0.        ,  0.59016393,  0.09183673]])
```

x_test

```
array([[0.09045226,  1.        ,  0.        ,  0.06557377,  0.28571429],
       [0.85427136,  1.        ,  0.        ,  0.59016393,  0.12244898],
       [0.53768844,  1.        ,  0.        ,  0.39344262,  0.45918367],
       [0.49246231,  1.        ,  0.        ,  0.37704918,  0.41836735],
       [0.88944724,  1.        ,  0.        ,  0.59836066,  0.69387755],
       [0.91457286,  1.        ,  0.        ,  0.68032787,  0.14285714],
       [0.02512563,  0.        ,  0.        ,  0.01639344,  0.76530612],
       [0.73366834,  1.        ,  0.        ,  0.50819672,  0.35714286],
       [0.06030151,  0.        ,  0.        ,  0.04098361,  0.14285714],
```

```
[0.7638191 ,  0.        ,  0.        ,  0.51639344,  0.19387755],
[0.30653266,  1.        ,  0.        ,  0.25409836,  0.55102041],
[0.6281407 ,  0.        ,  0.        ,  0.45081967,  0.7755102 ],
[0.90452261,  0.        ,  0.        ,  0.67213115,  0.31632653],
[0.77386935,  0.        ,  0.        ,  0.51639344,  0.15306122],
[0.40201005,  1.        ,  0.        ,  0.31967213,  0.51020408],
[0.03517588,  0.        ,  0.        ,  0.02459016,  0.94897959],
```

```
[0.16582915,  1.  ,  0.  ,  0.14754098,  0.92857143],
[0.65326633,  1.  ,  0.  ,  0.45901639,  0.08163265],
[0.18592965,  0.  ,  0.  ,  0.1557377 ,  0.73469388],
[0.3718593 ,  1.  ,  0.  ,  0.31967213,  0.46938776],
[0.91959799,  0.  ,  0.  ,  0.68032787,  0.8877551 ],
[0.72864322,  1.  ,  0.  ,  0.50819672,  0.97959184],
[0.22613065,  0.  ,  0.  ,  0.19672131,  0.65306122],
[0.79899497,  0.  ,  0.  ,  0.51639344,  0.73469388],
[0.30150754,  1.  ,  0.  ,  0.25409836,  0.56122449],
[0.61809045,  1.  ,  0.  ,  0.44262295,  0.91836735],
[0.89949749,  1.  ,  0.  ,  0.63934426,  0.90816327],
[0.92964824,  1.  ,  0.  ,  0.68852459,  0.97959184],
[0.61306533,  0.  ,  0.  ,  0.44262295,  0.58163265],
[0.22110553,  0.  ,  0.  ,  0.19672131,  0.2755102 ],
[0.08040201,  0.  ,  0.  ,  0.04918033,  0.34693878],
[0.27638191,  1.  ,  0.  ,  0.2295082 ,  0.40816327],
[0.75376884,  1.  ,  0.  ,  0.51639344,  0.16326531],
[0.55778894,  0.  ,  0.  ,  0.39344262,  0.54081633],
[0.11055276,  0.  ,  0.  ,  0.08196721,  0.04081633],
[0.94974874,  0.  ,  0.  ,  0.72131148,  0.85714286],
[0.64824121,  1.  ,  0.  ,  0.45901639,  0.75510204],
[0.0201005 ,  0.  ,  0.  ,  0.01639344,  0.39795918],
[0.41708543,  0.  ,  0.  ,  0.31967213,  0.43877551],
[0.53266332,  0.  ,  0.  ,  0.39344262,  0.5       ],
[0.67336683,  1.  ,  0.  ,  0.47540984,  0.04081633],
[0.33165829,  0.  ,  0.  ,  0.2704918 ,  0.5       ],
[0.13065327,  0.  ,  0.  ,  0.10655738,  0.31632653],
[0.5678392 ,  1.  ,  0.  ,  0.40163934,  0.45918367],
[0.84422111,  0.  ,  0.  ,  0.59016393,  0.26530612],
[0.31658291,  0.  ,  0.  ,  0.26229508,  0.59183673],
[0.04020101,  1.  ,  0.  ,  0.03278689,  0.02040816],
[0.37688442,  1.  ,  0.  ,  0.31967213,  0.54081633],
[0.59296482,  0.  ,  0.  ,  0.42622951,  0.42857143],
[0.71859296,  0.  ,  0.  ,  0.5       ,  0.87755102],
[0.35678392,  0.  ,  0.  ,  0.27868852,  0.41836735],
[0.62311558,  0.  ,  0.  ,  0.45081967,  0.28571429],
[0.92462312,  0.  ,  0.  ,  0.68852459,  0.3877551 ],
[0.48743719,  0.  ,  0.  ,  0.36885246,  0.5       ],
[0.74874372,  1.  ,  0.  ,  0.51639344,  0.90816327],
[0.12060302,  0.  ,  0.  ,  0.10655738,  0.13265306],
[0.15075377,  1.  ,  0.  ,  0.12295082,  0.03061224],
[0.8040201 ,  0.  ,  0.  ,  0.52459016,  0.34693878],
```

```
       [0.20100503,  0.        ,  0.        ,  0.18852459,  0.34693878],
       [0.28140704,  0.        ,  0.        ,  0.23770492,  0.5       ]])
```

y_train

131	75
96	47
181	86
19	98

153 76
.. ..
67 48
192 8
117 59
47 47
172 10
Name: Spending Score (1-100), Length: 140, dtype: int64y_test
18 29
170 13
107 46
98 42
177 69
182 15
5 76
146 36
12 15
152 20
61 55
125 77
180 32
154 16
80 51
7 94
33 92
130 9
37 73
74 47
183 88
145 97
45 65
159 73
60 56
123 91
179 90
185 97
122 58
44 28
16 35
55 41
150 17
111 54
22 5
189 85

129	75
4	40
83	44
106	50

```
134      5
66      50
26      32
113     46
168     27
63      59
8        3
75      54
118     43
143     87
71      42
124     29
184     39
97      50
149     90
24      14
30       4
160     35
40      35
56      50
Name:  Spending      Score (1-100), dtype: int64
```

Evaluation Metrics

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

0.9718954898467616

```
#MAE
from sklearn.metrics import mean_absolute_errorprint("MAE",mean_absolute_error(y_test,y_pred))
```

MAE 3.2810222489289385

```
#MSE
from sklearn.metrics import mean_squared_errorprint("MSE",mean_squared_error(y_test,y_pred))
```

MSE 17.504352105571893

```
#RMSE
print("RMSE",np.sqrt(mean_squared_error(y_test,y_pred)))
```

RMSE 4.183820276442559

```
#RMSLE
print("RMSE",np.log(np.sqrt(mean_squared_error(y_test,y_pred))))
```

RMSE 1.4312247708788837

```
# RSquared
from sklearn.metrics import r2_score
```

```
r2 = r2_score(y_test,y_pred)print(r2)
```

```
0.9755948089587156
```

```
#Adjusted RSquared
n=40
k=2
adj_r2_score = 1 - ((1-r2)*(n-1)/(n-k-1))print(adj_r2_score)
```

```
0.9742756094429705
```