```
 1 import matplotlib.pyplot as plt
 2 import pandas as pd
 3 import xlrd
 4 import seaborn as sns
 5 import numpy as np
 6
 7 # ----------------------------------------------------------------------
 8 # Step 1: Import the the Data using Pandas (Data_Set_1.xlsx)
 9
10 df = pd.read_excel("Data_Set_1.xlsx", "Sheet1")
11 # ----------------------------------------------------------------------
12 # Comment: Plots in matplotlib reside within a figure object, use plt.figure to create new figure.
13 # Step2: Extract the Ages from the data set and plot the distribution.
14
15 fig = plt.figure(1)
16 plt.hist(df['Age'], bins=7)
17 plt.title('Age distribution')
18 plt.xlabel('Age')
19 plt.ylabel('#Employee')
20 # ----------------------------------------------------------------------
21 # Step3: Extract the Ages from the data set use boxplot to plot the Ages measures.
22
23 fig = plt.figure(2)
24 plt.boxplot(df['Age'])
25 # ----------------------------------------------------------------------
26 # Step4: Extract the Ages from the data set and find the relationship between Age and Gender.
27 # Comment: Use seaborn libraray and use the violinplot.
28
29 fig=plt.figure(3)
30 sns.violinplot(df['Age'], df['Gender'])
31
32 # ----------------------------------------------------------------------
33 # Step5:  Find the sum of sales by each group at gender level. Store them in the variable.
34 # Comment: Use pandas and use groupby method.
35 var = df.groupby('Gender').Sales.sum()
```

```
36 # ----------------------------------------------------------------------
37 # Step6:  Plot sum of sales by each group at gender level using bar plot.
38
39 fig = plt.figure(4)
40 plt.xlabel('Gender')
41 plt.ylabel('Sum of Sales')
42 plt.title("Gender wise Sum of Sales")
43 var.plot(kind='bar')
44 # ----------------------------------------------------------------------
45 # Step7:  Plot sum of sales by each group at BMI level using line plot.
46
47 var = df.groupby('BMI').Sales.sum()
48 fig = plt.figure(5)
49 plt.xlabel('BMI')
50 plt.ylabel('Sum of Sales')
51 plt.title("BMI wise Sum of Sales")
52 var.plot(kind='line')
53 # ----------------------------------------------------------------------
54 # Step8:  Plot sum of sales by each group and BMI level using bar plot (Stacked Bar plot).
55
56 var = df.groupby(['BMI','Gender']).Sales.sum()
57 var.unstack().plot(kind='bar',stacked=True,  color=['red','blue'], grid=False)
58 # ----------------------------------------------------------------------
59 # Step9:  Plot sum of sales by each group and BMI level using bar plot (UnStacked Bar plot).
60
61 var = df.groupby(['BMI','Gender']).Sales.sum()
62 var.unstack().plot(kind='bar',stacked=False,  color=['red','blue'], grid=False)
63 # ----------------------------------------------------------------------
64 # Step10:  Find the relationship between age and Sales.
65
66 fig = plt.figure(7)
67 plt.scatter(df['Age'],df['Sales'])
68 # ----------------------------------------------------------------------
69 # Step11:  Find the relationship between age and Sales and find a way to include income into this Rela
70 # Comment: Added third variable income as size of the bubble
```

```
71
72 fig = plt.figure(8)
73 plt.scatter(df['Age'], df['Sales'], s=df['Income'])
74
75 # ----------------------------------------------------------------------
76 # Step12:  Use pie chart to shows the Sales for each gender.
77
78 var = df.groupby(['Gender']).sum().stack()
79 temp = var.unstack()
80 type(temp)
81 Sales = temp['Sales']
82 Group_Gender = temp.index
83 plt.axis("equal")
84 plt.pie(Sales, labels=Group_Gender, autopct="%1.1f%%")
85 plt.title("Expenses")
86 plt.show()
87
88
```