**Lab Pandas Data Analysis (Day-28)**

**Lab1: Write a Pandas program to split the following data frame into groups based**

**on Class and count the number of students in that particular class.Also generate**

**a bar chart based on the result and explain the conclusion.**

**Input:**

**student\_data = pd.DataFrame({**

**'school\_code': ['s001','s002','s003','s001','s002','s004'],**

**'class': ['V', 'VI', 'VI', 'VI', 'V', 'VI'],**

**'name': ['Alberto Franco','Gino Mcneill','Ryan Parkes', 'Eesha Hinton', 'Gino Mcneill',**

**'David Parkes'],**

**'age': [12, 12, 13, 13, 14, 12],**

**'height': [173, 192, 186, 167, 151, 159],**

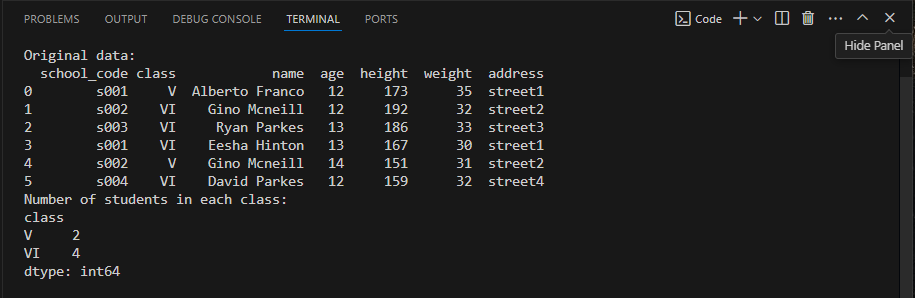
**'weight': [35, 32, 33, 30, 31, 32],**

**'address': ['street1', 'street2', 'street3', 'street1', 'street2', 'street4']},**

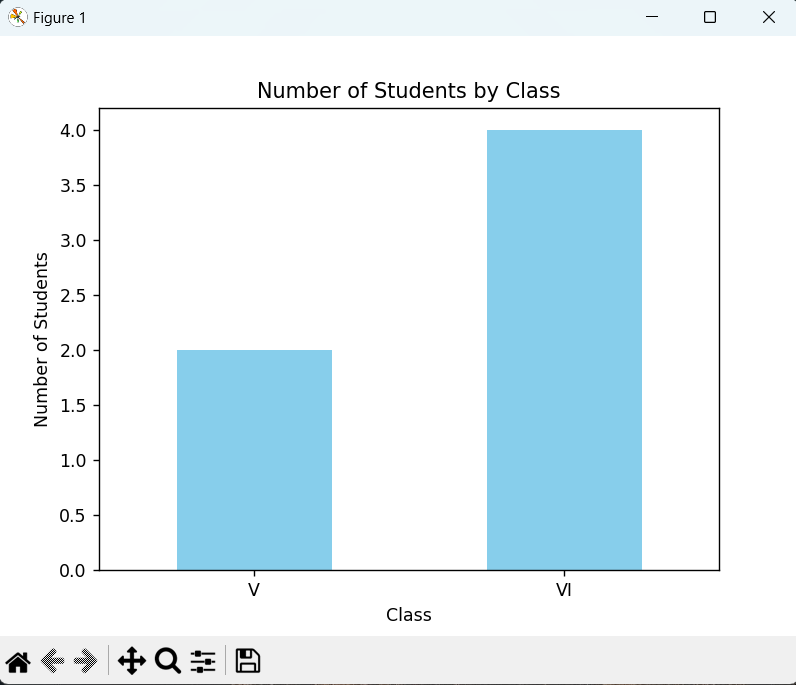
**)**

**Output:**

**1.**

****

**2.**

****

**Lab2: Write a Pandas program to split the following dataframe by school**

**code and get mean, min, and max value of age for each school.**

**Also generate a horizontal bar chart based on the result and explain the**

**conclusion.**

**Input:**

**student\_data = pd.DataFrame({**

**'school\_code': ['s001','s002','s003','s001','s002','s004'],**

**'class': ['V', 'V', 'VI', 'VI', 'V', 'VI'],**

**'name': ['Alberto Franco','Gino Mcneill','Ryan Parkes', 'Eesha Hinton', 'Gino Mcneill',**

**'David Parkes'],**

**'age': [12, 12, 13, 13, 14, 12],**

**'height': [173, 192, 186, 167, 151, 159],**

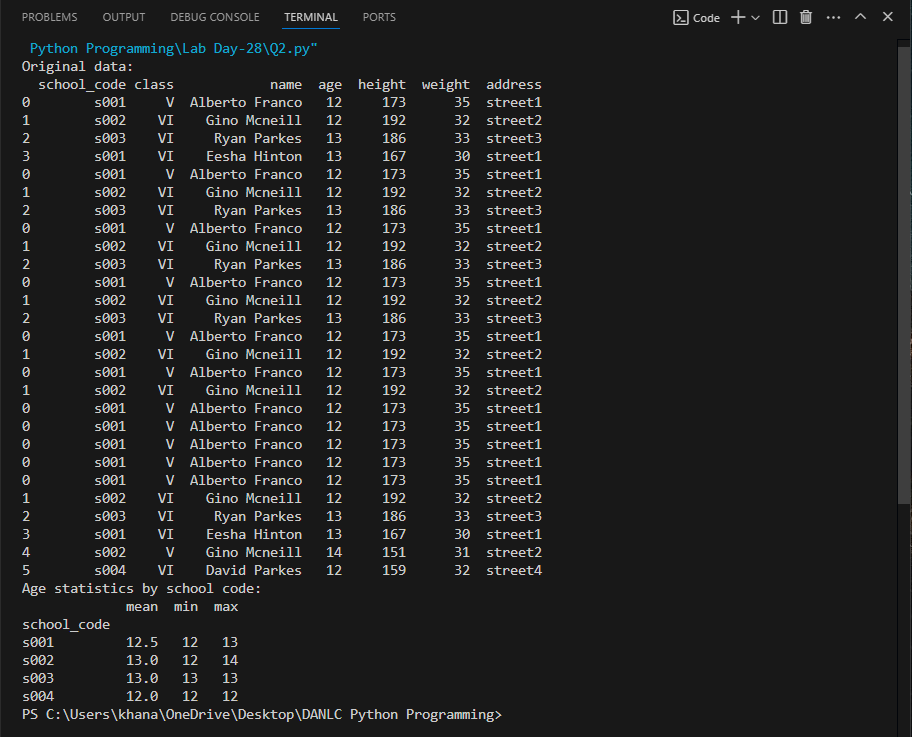
**'weight': [35, 32, 33, 30, 31, 32],**

**'address': ['street1', 'street2', 'street3', 'street1', 'street2', 'street4']},**

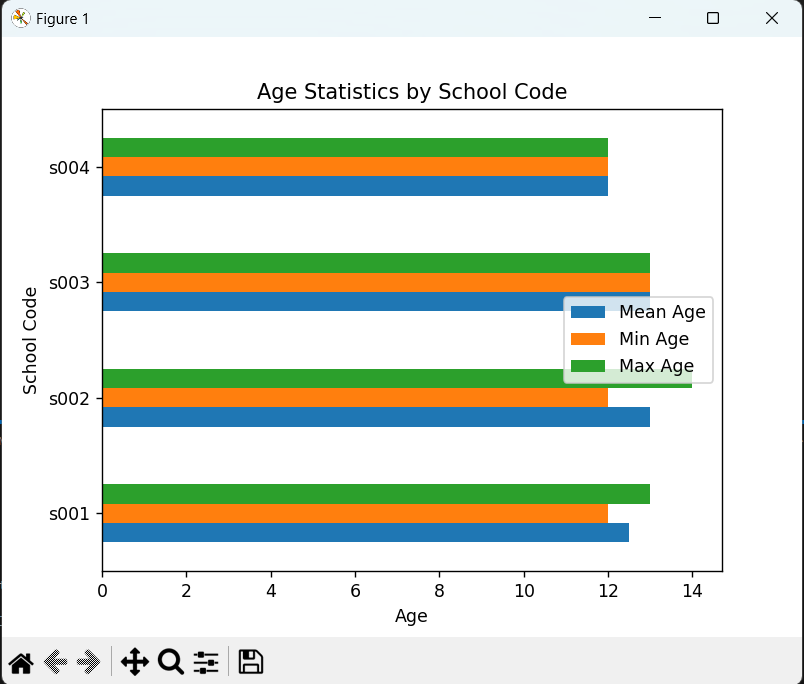
**)**

**Output:**

**1.**

****

**2.**

****

**Lab3: Write a Pandas program to split a dataset, group by one column and**

**get mean, min, and max values by group. Using the following dataset find**

**the mean, min, and max values of purchase amount (purch\_amt) group by**

**customer id (customer\_id).Also generate a line chart based on the result and**

**explain the conclusion.**

**Input:**

**orders\_data = pd.DataFrame({**

**'ord\_no':[70001,70009,70002,70004,70007,70005,70008,70010,70003,70012,70011,**

**70013],**

**'purch\_amt':[150.5,270.65,65.26,110.5,948.5,2400.6,5760,1983.43,2480.4,250.45,**

**75.29,3045.6],**

**'ord\_date':**

**['2012-10-05','2012-09-10','2012-10-05','2012-08-17','2012-09-10','2012-07-27','2012-**

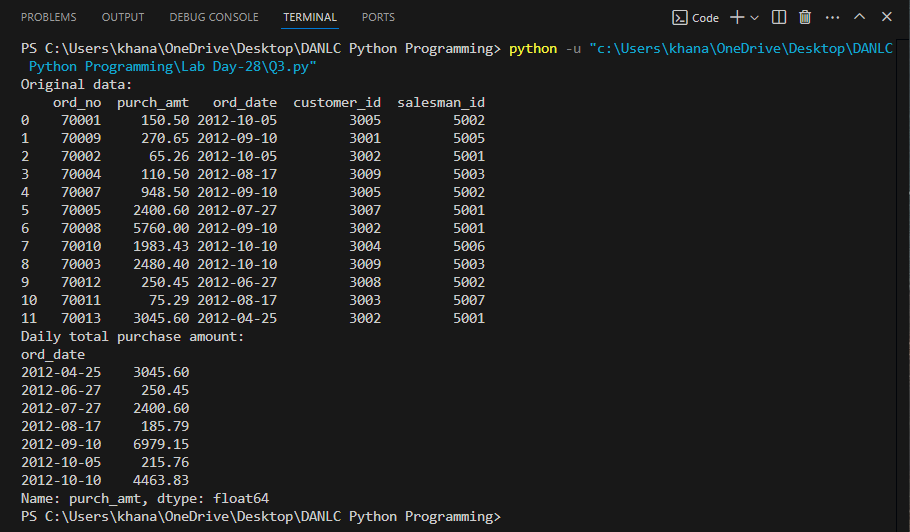
**09-10','2012-10-10','2012-10-10','2012-06-27','2012-08-17','2012-04-25'],**

**'customer\_id':[3005,3001,3002,3009,3005,3007,3002,3004,3009,3008,3003,3002],**

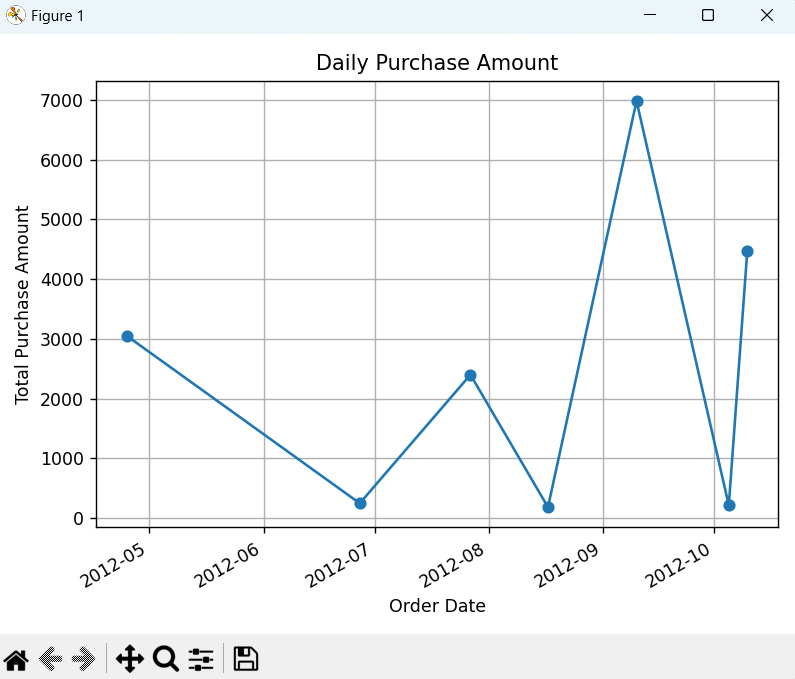
**'salesman\_id': [5002,5005,5001,5003,5002,5001,5001,5006,5003,5002,5007,5001]})**

**Output:**

**1.**

****

**2.**

****

**Lab4: Write a Pandas program to split the following data frame into groups**

**and calculate monthly purchase amount.Also generate a bar chart based on**

**the result and explain the conclusion.**

**Input:**

**df = pd.DataFrame({**

**'ord\_no':[70001,70009,70002,70004,70007,70005,70008,70010,70003,70012,70011,**

**70013],**

**'purch\_amt':[150.5,270.65,65.26,110.5,948.5,2400.6,5760,1983.43,2480.4,250.45,**

**75.29,3045.6],**

**'ord\_date':**

**['05-10-2012','09-10-2012','05-10-2012','08-17-2012','10-09-2012','07-27-2012','10-09-**

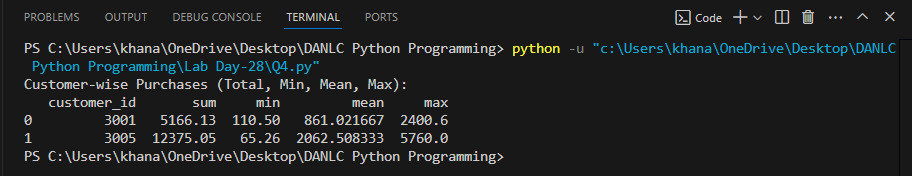
**2012','10-10-2012','10-10-2012','06-17-2012','07-08-2012','04-25-2012'],**

**'customer\_id':[3001,3001,3005,3001,3005,3001,3005,3001,3005,3001,3005,3005],**

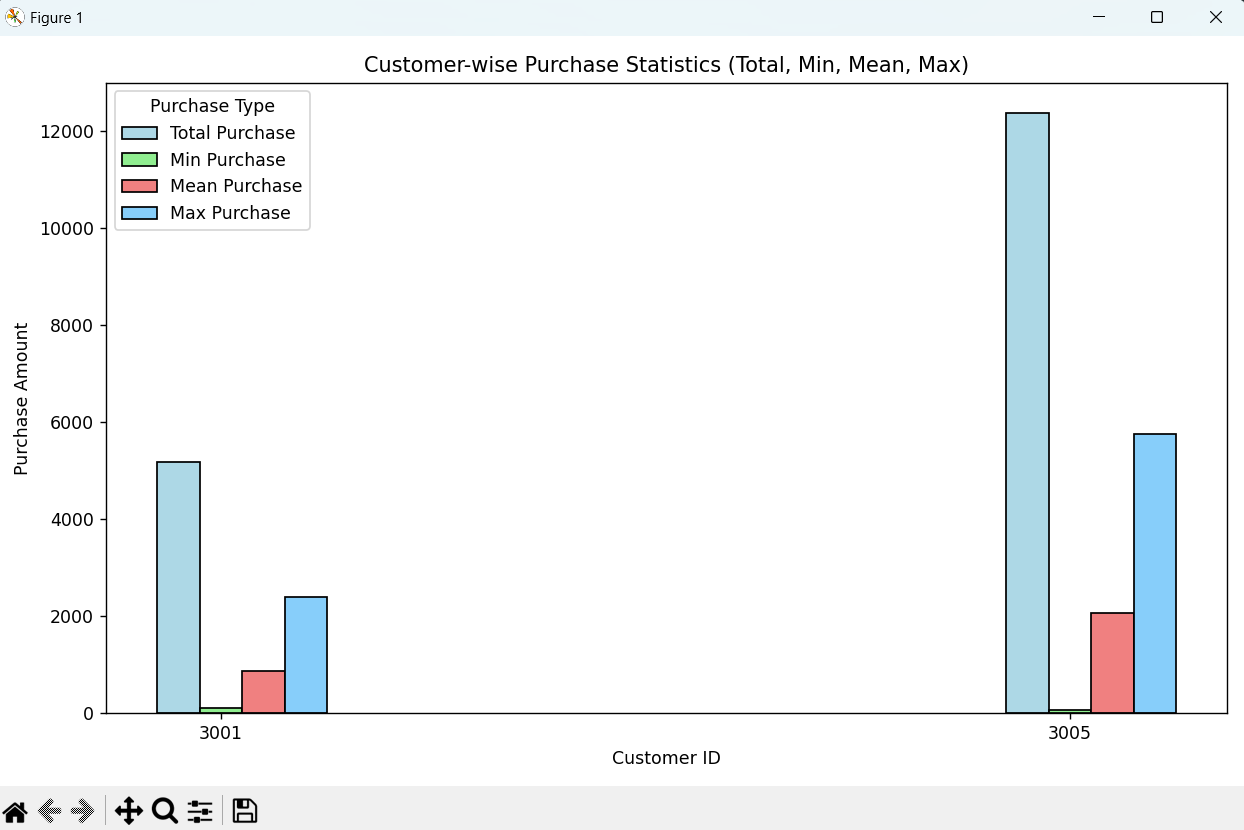
**'salesman\_id': [5002,5005,5001,5003,5002,5001,5001,5006,5003,5002,5007,5001]})**

**Output:**

**1.**

****

**2.**

****