pm9c5t2o9

January 23, 2025

[ ]:
```
import pandas as pd
```

[ ]:
```
data={
    "Name":["Raman","Suresh","Bikash"],
    "Age":[22,23,24],
    "city":["Ktm",'ltp',"pok"]
}
df=pd.DataFrame(data)
df.info() #checks if dataframe is null or not
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3 entries, 0 to 2
Data columns (total 3 columns):
 #   Column  Non-Null Count  Dtype 
---  ------  --------------  ----- 
 0   Name    3 non-null      object
 1   Age     3 non-null      int64 
 2   city    3 non-null      object
dtypes: int64(1), object(2)
memory usage: 200.0+ bytes
```

[ ]:
```
#deleting specific row
df.drop(1, axis=0, inplace=True)
df.head()
```

[ ]:
```
     Name  Age city
0   Raman   22  Ktm
2  Bikash   24  pok
```

[ ]:
```
df.drop(['Name','Age'],axis=1)
```

[ ]:
```
  city
0  Ktm
1  ltp
2  pok
```

[ ]:
```
#accessing individual colums
df[['city','Age']]
```

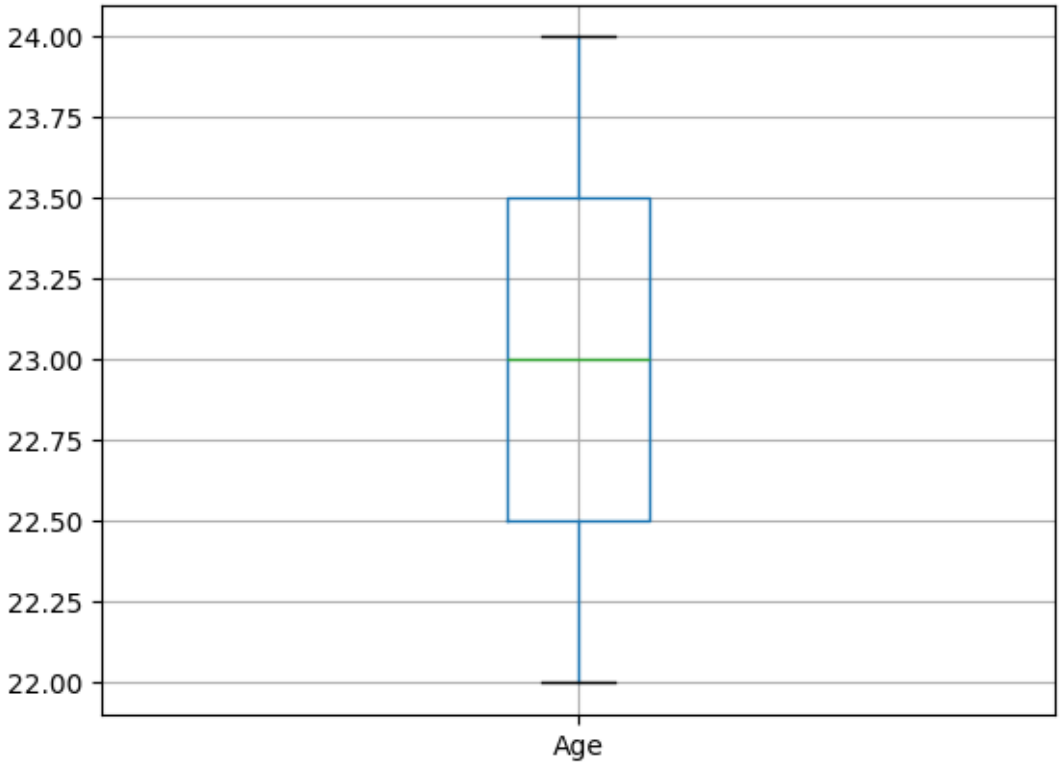
[ ]:
```
  city  Age
0  Ktm   22
1  ltp   23
2  pok   24
```

[ ]:
```
#deleting a specific column
df.drop(['city'],axis=1)
```

[ ]:
```
     Name  Age
0   Raman   22
2  Bikash   24
```

[ ]:
```
df.boxplot()
```

[ ]:
```
<Axes: >
```



[ ]:
```
new=df[df['Age']>23]
new[['Name']]
```

[ ]:
```
     Name
2  Bikash
```

[ ]:
```
data={
    "Name":["Raman","Suresh","Bikash","Raman"],
    "Age":[22,23,24,33],
    "city":["Ktm",'ltp',"pok","Bhak"]
}
df=pd.DataFrame(data)
#more useful in case of categorical data for eg male and female
df['Name'].duplicated() #checks for duplicate
df['Name'].value_counts() #checks for duplicate values
```

[ ]:
```
Name
Raman     2
Suresh    1
Bikash    1
Name: count, dtype: int64
```

[ ]:
```
data={
    "Name":["Raman","Suresh","Bikash"],
    "Age":[22,23,24],
    "city":["Ktm",'ltp',"pok"]
}
df=pd.DataFrame(data)
df.to_csv("output.csv",sep=",")
```

[ ]:
```
#create an excel data of your table mates with columns Name, Age, Gender(M/F).
#In one instance, filter out data of students whose age>19
#in another instance show all male and female student data deiifernetly
data1={
    "Name":["Raman","Salam","Suraj","Santosh","Sajin","Sajani"],
    "Age":[22,23,23,20,21,22],
    "Gender":["M",'m',"M","f","M","F"]
}
df=pd.DataFrame(data1)
df.to_csv("assignment.csv")
```

[ ]:
```
pd.read_csv("assignment.csv")
```

[ ]:
```
   Unnamed: 0     Name  Age Gender
0           0    Raman   22      M
1           1    Salam   23      m
2           2    Suraj   23      M
3           3  Santosh   20      f
4           4    Sajin   21      M
5           5   Sajani   22      F
```

[ ]:
```
df[df['Age']>19]
```

[ ]:
```
      Name  Age Gender
0    Raman   22      M
1    Salam   23      m
2    Suraj   23      M
3  Santosh   20      f
4    Sajin   21      M
5   Sajani   22      F
```

[ ]:
```
df[df['Gender'].str.upper() == 'F']
```

[ ]:
```
      Name  Age Gender
3  Santosh   20      f
5   Sajani   22      F
```

[ ]:
```
df[(df['Gender']=='F') | (df['Gender']=='f')]
```

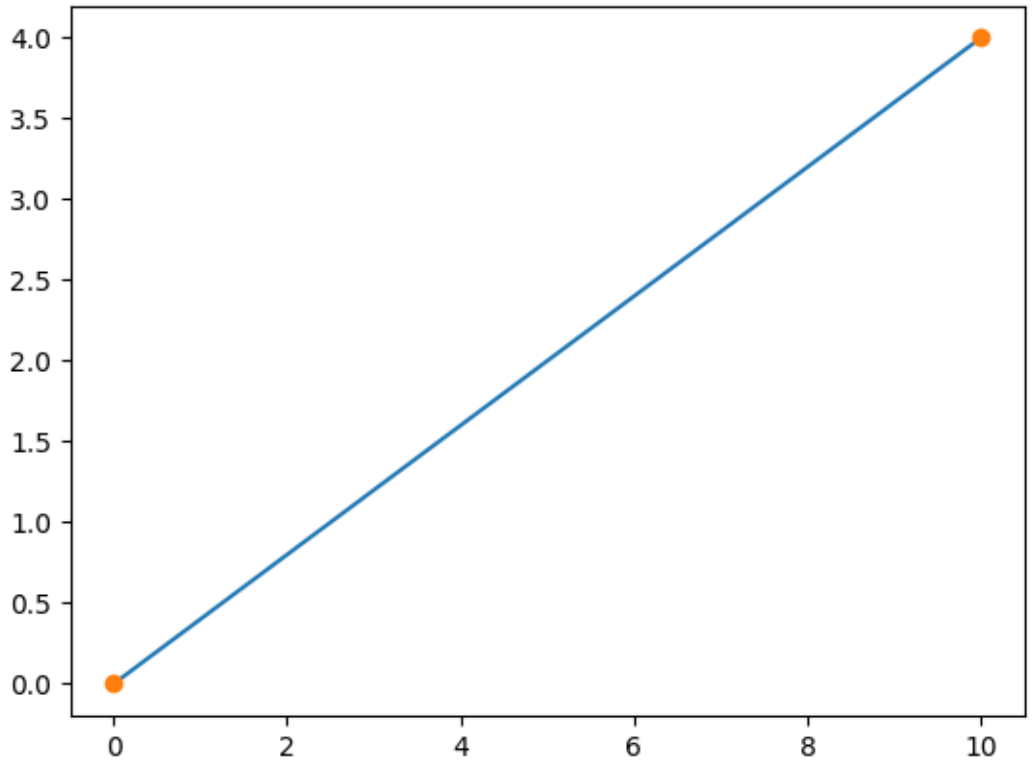
[ ]:
```
      Name  Age Gender
3  Santosh   20      f
5   Sajani   22      F
```

[ ]:
```
df[(df['Gender']=='M')|(df['Gender']=='m')]
```

[ ]:
```
    Name  Age Gender
0  Raman   22      M
1  Salam   23      m
2  Suraj   23      M
4  Sajin   21      M
```

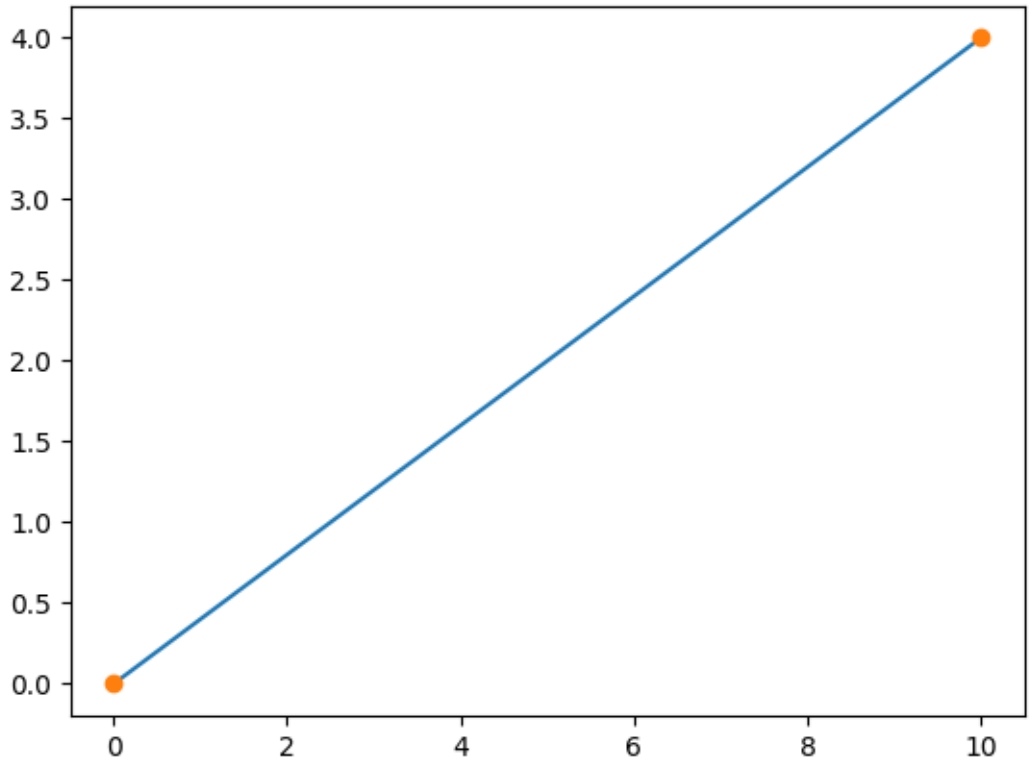
[ ]:
```
import matplotlib.pyplot as plt
import numpy as np

xpoints=np.array([0,10])
ypoints=np.array([0,4])
plt.plot(xpoints,ypoints)
plt.plot(xpoints,ypoints,"o")
plt.show()
```



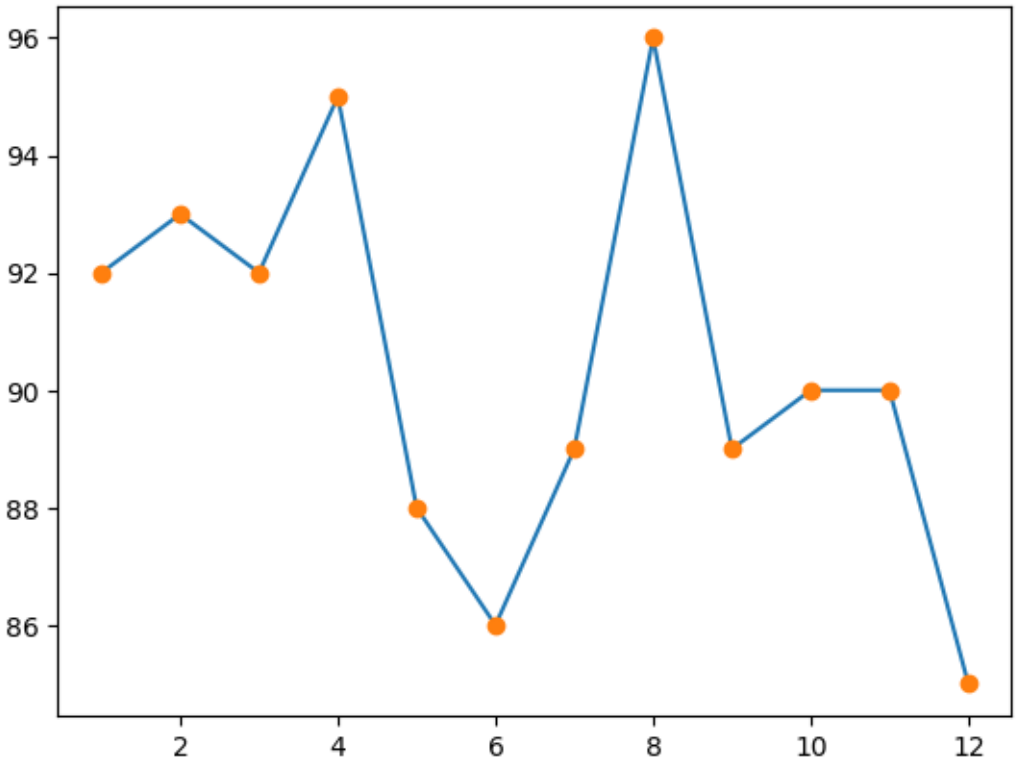
[ ]:
```
import matplotlib.pyplot as plt
import numpy as np

xpoints=np.array([0,10])
ypoints=np.array([0,4])
plt.plot(xpoints,ypoints, marker="*")
plt.plot(xpoints,ypoints,"o")
plt.show()
```



[ ]:
```
#create a numpy array of your class and percentage you scores (1,90),(2,85)...
#Dsiplay your academic progress in a graph using matplotlib

xpoints=np.array([1,2,3,4,5,6,7,8,9,10,11,12])
ypoints=np.array([92,93,92,95,88,86,89,96,89,90,90,85])
plt.plot(xpoints,ypoints, marker="*")
plt.plot(xpoints,ypoints,"o")
plt.show()
```


96
94
92
90
88
86
2
4
6
8
10
12