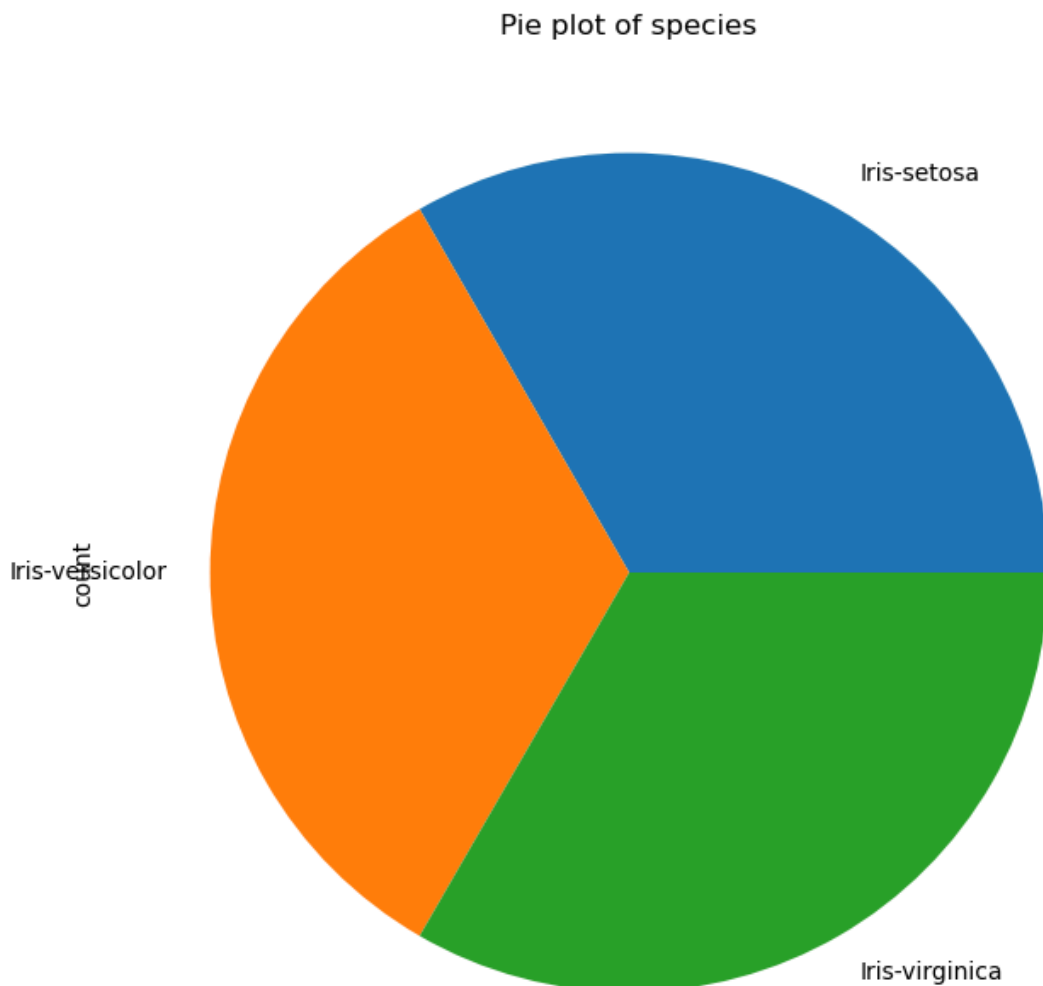


```
In [70]: import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sns
from sklearn import preprocessing
```

slip 1,11

que 2 a

```
In [12]: df=pd.read_csv('C:/Users/PRANJAL/Desktop/html slips/IRIS.csv')
ax=plt.subplots(1,1,figsize=(10,8))
df['species'].value_counts().plot.pie()
plt.title("Pie plot of species")
plt.show()
```



que 2 b

```
In [ ]: df=pd.read_csv('C:/Users/PRANJAL/Desktop/html slips/IRIS.csv')
```

```
In [14]: df.shape
```

```
Out[14]: (150, 5)
```

```
In [16]: df.size
```

```
Out[16]: 750
```

```
In [18]: df.dtypes
```

```
Out[18]: sepal_length    float64
sepal_width    float64
petal_length    float64
petal_width    float64
species        object
dtype: object
```

```
In [20]: df.describe()
```

```
Out[20]:
```

	sepal_length	sepal_width	petal_length	petal_width
count	150.000000	150.000000	150.000000	150.000000
mean	5.843333	3.054000	3.758667	1.198667
std	0.828066	0.433594	1.764420	0.763161
min	4.300000	2.000000	1.000000	0.100000
25%	5.100000	2.800000	1.600000	0.300000
50%	5.800000	3.000000	4.350000	1.300000
75%	6.400000	3.300000	5.100000	1.800000
max	7.900000	4.400000	6.900000	2.500000

```
In [22]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 5 columns):
#   Column          Non-Null Count  Dtype
---  -
0   sepal_length    150 non-null   float64
1   sepal_width     150 non-null   float64
2   petal_length    150 non-null   float64
3   petal_width     150 non-null   float64
4   species         150 non-null   object
dtypes: float64(4), object(1)
memory usage: 6.0+ KB
```

slip 2 , 6

que 2 a

```
In [24]: df=pd.read_csv('C:/Users/PRANJAL/Desktop/html slips/ds.csv')
print(df)
print("-----")
v1=df['Age'].mean()
v2=df['Salary'].mean()
df['Age'].fillna(v1,inplace=True)
df['Salary'].fillna(v2,inplace=True)
print(df)
```

	Country	Age	Salary	Purchased
0	France	44.0	72000.0	No
1	Spain	27.0	48000.0	YES
2	Germany	30.0	54000.0	No
3	Spain	38.0	61000.0	No
4	Germany	40.0	NaN	YES
5	France	35.0	58000.0	YES
6	Spain	NaN	52000.0	No
7	France	48.0	79000.0	YES
8	Germany	50.0	83000.0	No
9	France	37.0	67000.0	Yes

	Country	Age	Salary	Purchased
0	France	44.000000	72000.000000	No
1	Spain	27.000000	48000.000000	YES
2	Germany	30.000000	54000.000000	No
3	Spain	38.000000	61000.000000	No
4	Germany	40.000000	63777.777778	YES
5	France	35.000000	58000.000000	YES
6	Spain	38.777778	52000.000000	No
7	France	48.000000	79000.000000	YES
8	Germany	50.000000	83000.000000	No
9	France	37.000000	67000.000000	Yes

C:\Users\PRANJAL\AppData\Local\Temp\ipykernel_20348\4383282.py:6: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method. The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform the operation inplace on the original object.

```
df['Age'].fillna(v1,inplace=True)
```

C:\Users\PRANJAL\AppData\Local\Temp\ipykernel_20348\4383282.py:7: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method. The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform the operation inplace on the original object.

```
df['Salary'].fillna(v2,inplace=True)
```

que 2 b

```
In [29]: df=pd.read_csv('C:/Users/PRANJAL/Desktop/html slips/ds.csv')
plt.figure(figsize=(10,7))
df.plot(x="Country",y="Salary")
plt.xlabel("species")
plt.ylabel("values")
plt.title("line plot")
plt.show()
```

<Figure size 1000x700 with 0 Axes>



que 2 c

```
In [33]: df=pd.read_csv('C:/Users/PRANJAL/Desktop/html slips/ds.csv')
print("1st 10 rows is \n ",df.head(10))
print("-----")
print("last 10 rows is \n",df.tail(10))
print("-----")
print("random 20 rows is \n ",df.sample(10))
print("-----")
print("shape is \n",df.shape)
```

```
1st 10 rows is
```

	Country	Age	Salary	Purchased
0	France	44.0	72000.0	No
1	Spain	27.0	48000.0	YES
2	Germany	30.0	54000.0	No
3	Spain	38.0	61000.0	No
4	Germany	40.0	NaN	YES
5	France	35.0	58000.0	YES
6	Spain	NaN	52000.0	No
7	France	48.0	79000.0	YES
8	Germany	50.0	83000.0	No
9	France	37.0	67000.0	Yes

```
last 10 rows is
```

	Country	Age	Salary	Purchased
0	France	44.0	72000.0	No
1	Spain	27.0	48000.0	YES
2	Germany	30.0	54000.0	No
3	Spain	38.0	61000.0	No
4	Germany	40.0	NaN	YES
5	France	35.0	58000.0	YES
6	Spain	NaN	52000.0	No
7	France	48.0	79000.0	YES
8	Germany	50.0	83000.0	No
9	France	37.0	67000.0	Yes

```
random 20 rows is
```

	Country	Age	Salary	Purchased
0	France	44.0	72000.0	No
5	France	35.0	58000.0	YES
4	Germany	40.0	NaN	YES
2	Germany	30.0	54000.0	No
3	Spain	38.0	61000.0	No
6	Spain	NaN	52000.0	No
9	France	37.0	67000.0	Yes
1	Spain	27.0	48000.0	YES
8	Germany	50.0	83000.0	No
7	France	48.0	79000.0	YES

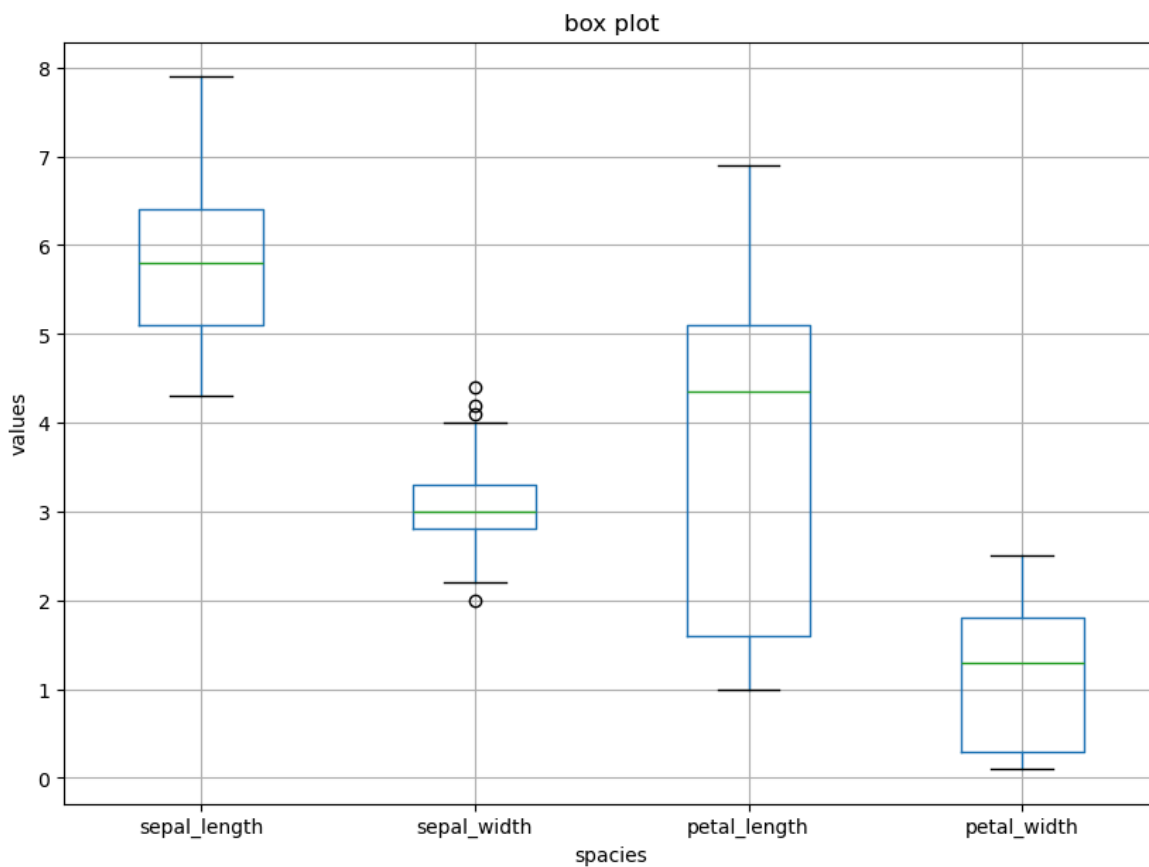
```
shape is
(10, 4)
```

slip 3

que 2 a

```
In [35]: df=pd.read_csv('C:/Users/PRANJAL/Desktop/html slips/IRIS.csv')
data=df[["sepal_length","sepal_width","petal_length","petal_width"]]
plt.figure(figsize=(10,7))
plt.xlabel("spacies")
plt.ylabel("values")
plt.title("box plot")
data.boxplot()
```

```
Out[35]: <Axes: title={'center': 'box plot'}, xlabel='spacies', ylabel='values'>
```



que 2 b

```
In [41]: df=pd.read_csv('C:/Users/PRANJAL/Desktop/html slips/weight-height.csv')  
df.shape
```

```
Out[41]: (10000, 3)
```

```
In [42]: df.size
```

```
Out[42]: 30000
```

```
In [43]: df.dtypes
```

```
Out[43]: Gender      object  
Height    float64  
Weight    float64  
dtype: object
```

```
In [45]: df.describe()
```

Out[45]:

	Height	Weight
count	10000.000000	10000.000000
mean	66.367560	161.440357
std	3.847528	32.108439
min	54.263133	64.700127
25%	63.505620	135.818051
50%	66.318070	161.212928
75%	69.174262	187.169525
max	78.998742	269.989699

In [47]: `df.info()`

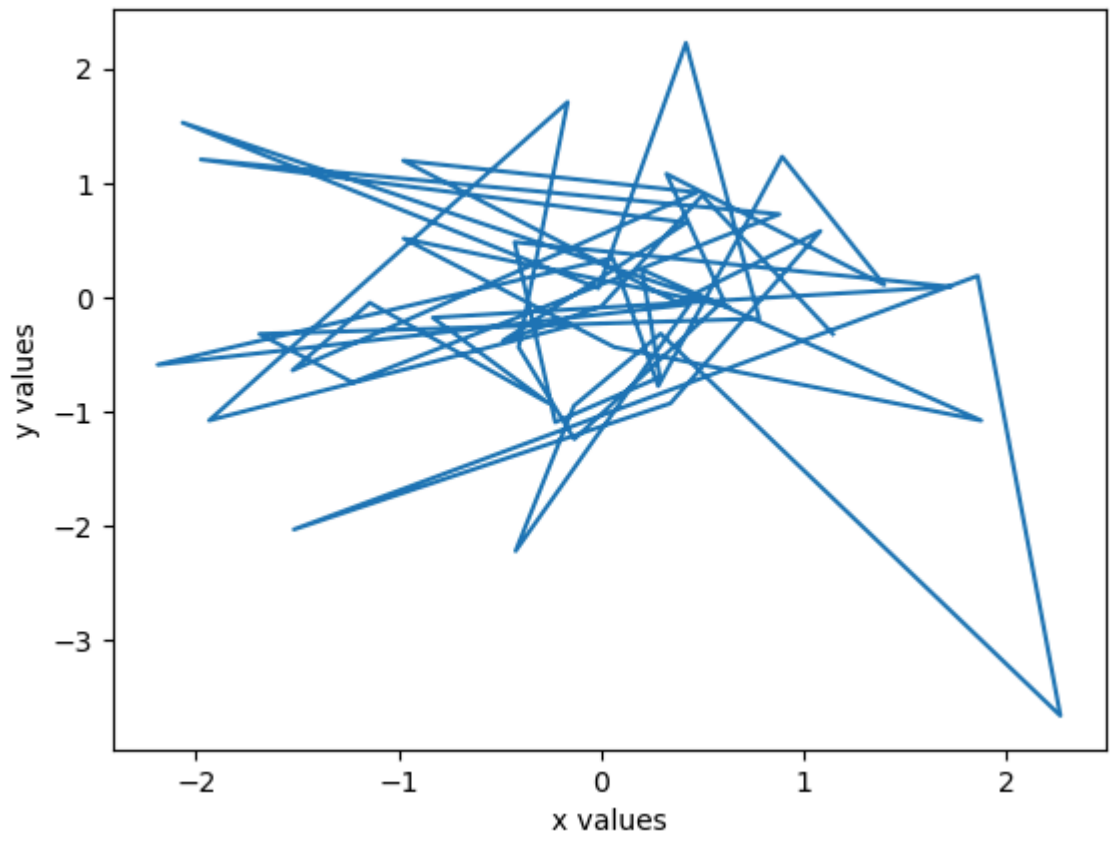
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10000 entries, 0 to 9999
Data columns (total 3 columns):
#   Column  Non-Null Count  Dtype
---  -
0   Gender   10000 non-null    object
1   Height   10000 non-null    float64
2   Weight   10000 non-null    float64
dtypes: float64(2), object(1)
memory usage: 234.5+ KB
```

slip 4 ,5

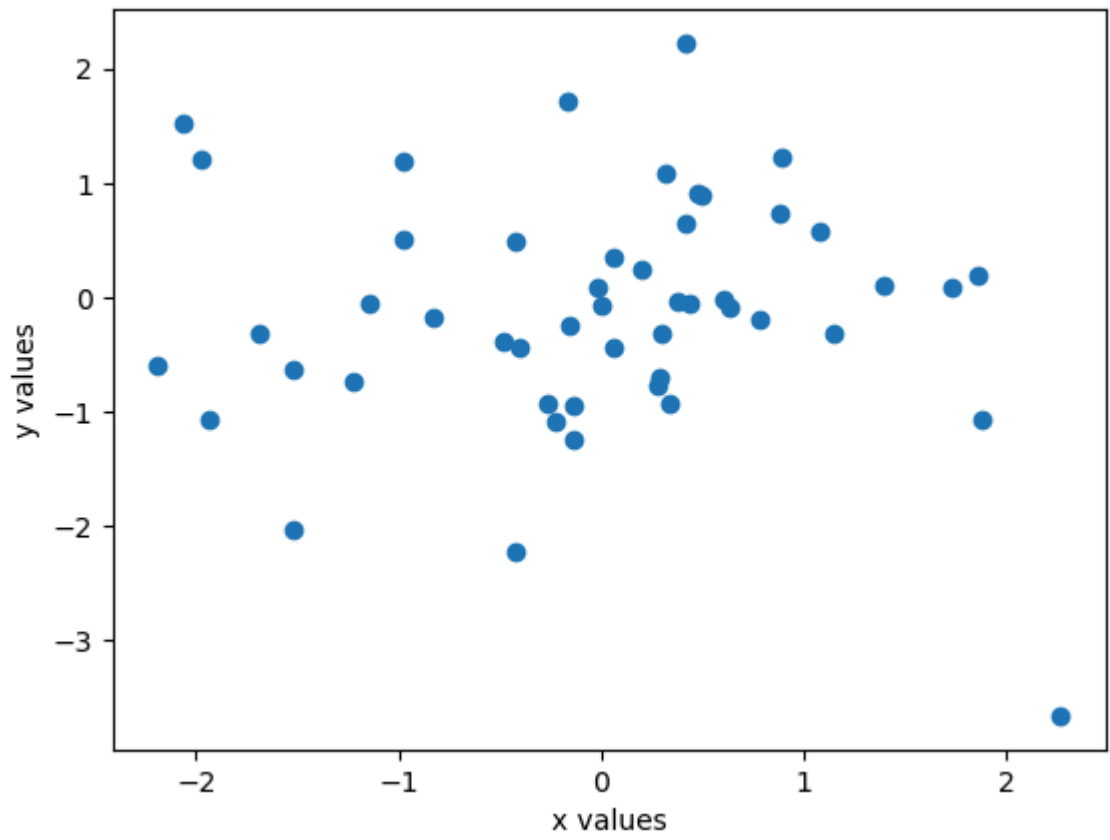
que 2 a

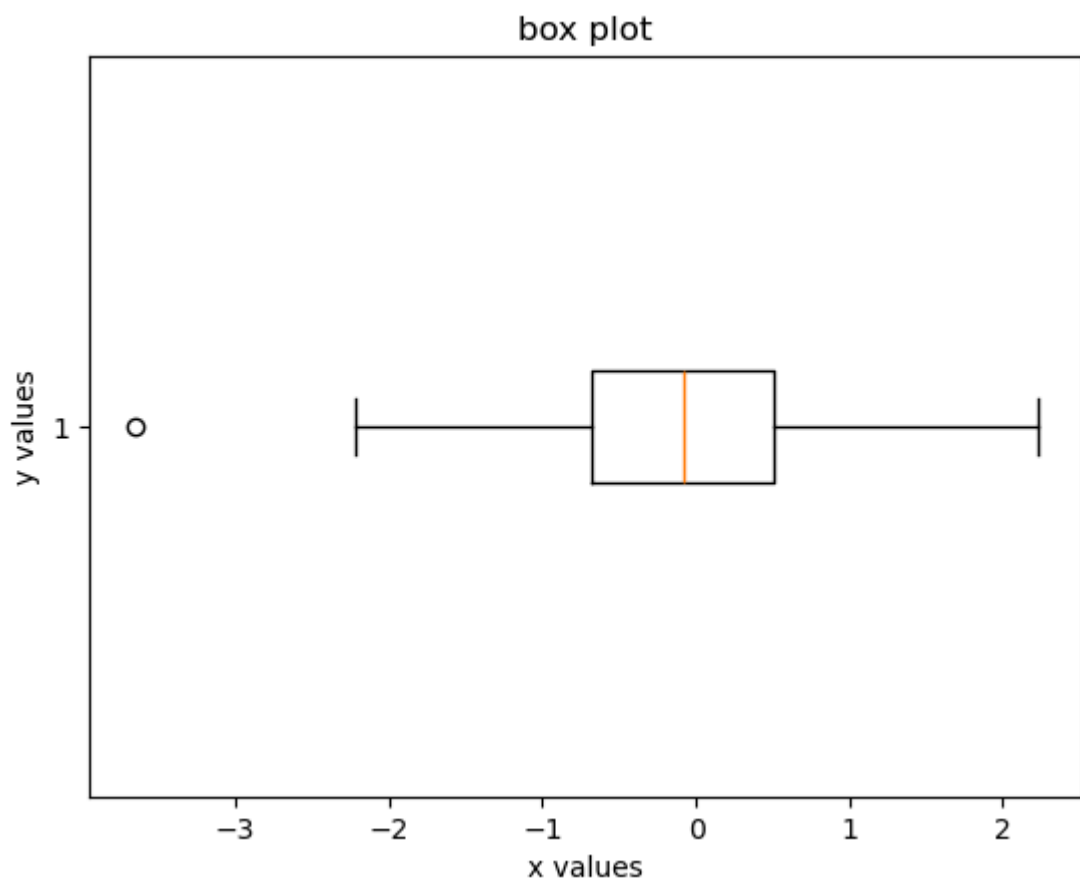
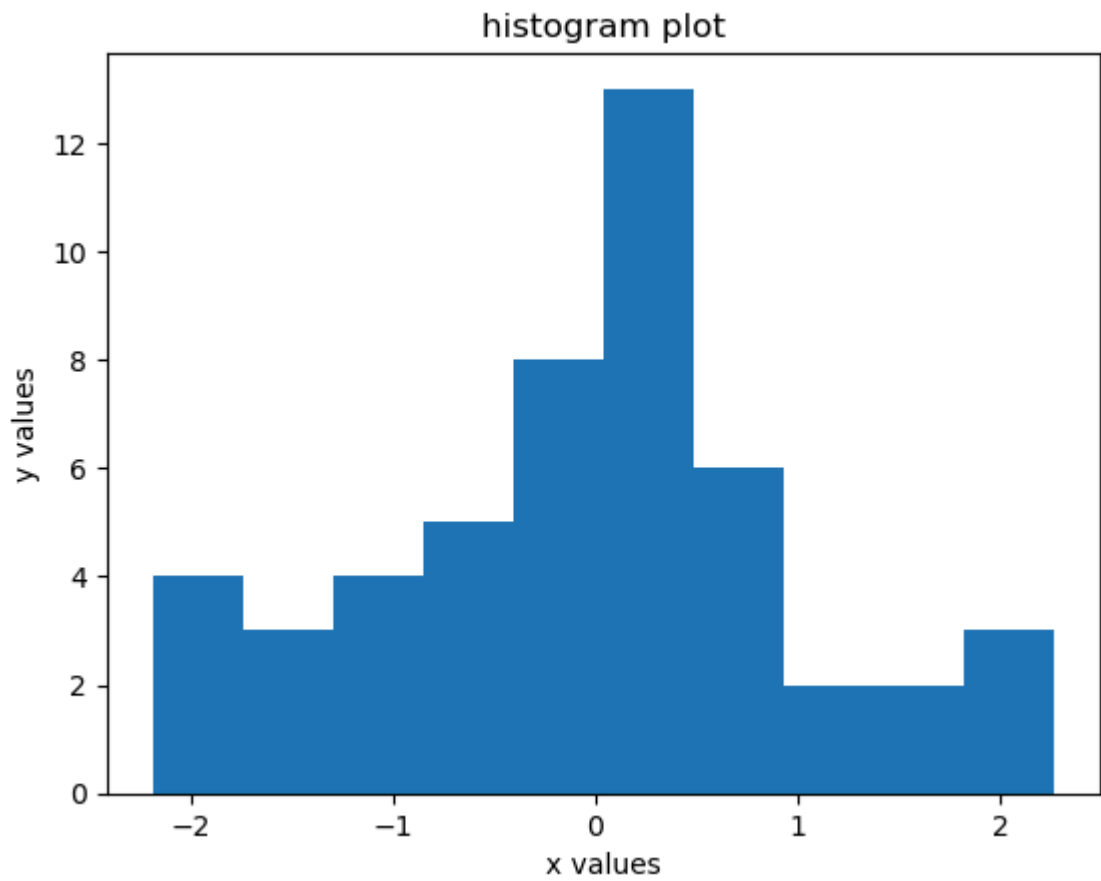
```
In [58]: x=np.random.randn(50)
y=np.random.randn(50)
plt.plot(x,y)
plt.xlabel("x values")
plt.ylabel("y values")
plt.title("line plot")
plt.show()
plt.scatter(x,y)
plt.xlabel("x values")
plt.ylabel("y values")
plt.title("scatter plot")
plt.show()
plt.hist(x)
plt.xlabel("x values")
plt.ylabel("y values")
plt.title("histogram plot")
plt.show()
plt.boxplot(y,vert=False)
plt.xlabel("x values")
plt.ylabel("y values")
plt.title("box plot")
plt.show()
```


line plot



scatter plot





que 2 b

```
In [60]: df=pd.read_csv('C:/Users/PRANJAL/Desktop/html slips/ds.csv')
df.shape
```

Out[60]: (10, 4)

In [62]: `df.size`

Out[62]: 40

In [64]: `df.dtypes`

Out[64]: Country object
Age float64
Salary float64
Purchased object
dtype: object

In [66]: `df.info()`

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 10 entries, 0 to 9  
Data columns (total 4 columns):  
#   Column      Non-Null Count  Dtype  
---  -  
0   Country     10 non-null    object  
1   Age         9 non-null     float64  
2   Salary      9 non-null     float64  
3   Purchased   10 non-null    object  
dtypes: float64(2), object(2)  
memory usage: 452.0+ bytes
```

In [68]: `df.describe()`

Out[68]:

	Age	Salary
count	9.000000	9.000000
mean	38.777778	63777.777778
std	7.693793	12265.579662
min	27.000000	48000.000000
25%	35.000000	54000.000000
50%	38.000000	61000.000000
75%	44.000000	72000.000000
max	50.000000	83000.000000

slip 7,slip 29,slip 27

In [79]: `df=pd.read_csv('C:/Users/PRANJAL/Desktop/html slips/ds.csv')
print(df)
print("-----")
ohe=pd.get_dummies(df,columns=['Country'])
print(ohe)
print("-----")
le=preprocessing.LabelEncoder()`

```
df['Purchased']=le.fit_transform(df['Purchased'])
print(df)
```

	Country	Age	Salary	Purchased
0	France	44.0	72000.0	No
1	Spain	27.0	48000.0	YES
2	Germany	30.0	54000.0	No
3	Spain	38.0	61000.0	No
4	Germany	40.0	NaN	YES
5	France	35.0	58000.0	YES
6	Spain	NaN	52000.0	No
7	France	48.0	79000.0	YES
8	Germany	50.0	83000.0	No
9	France	37.0	67000.0	Yes

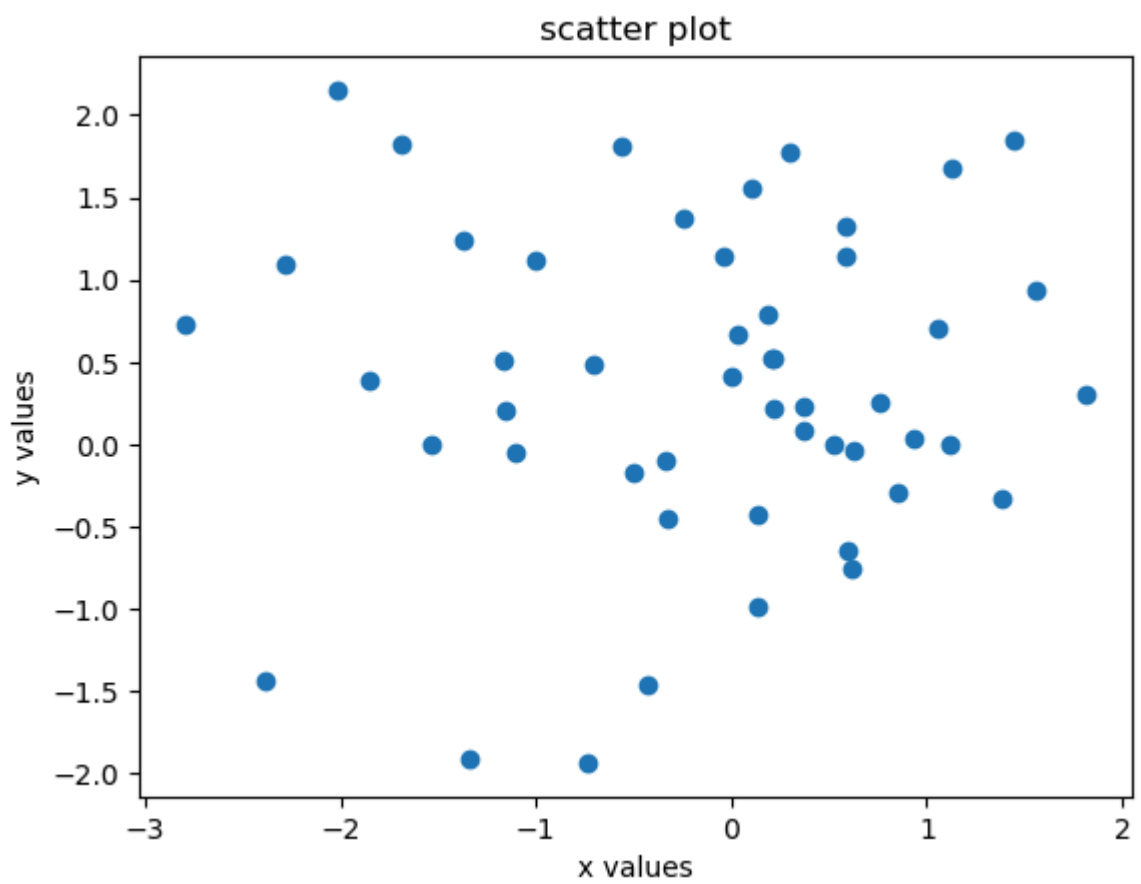
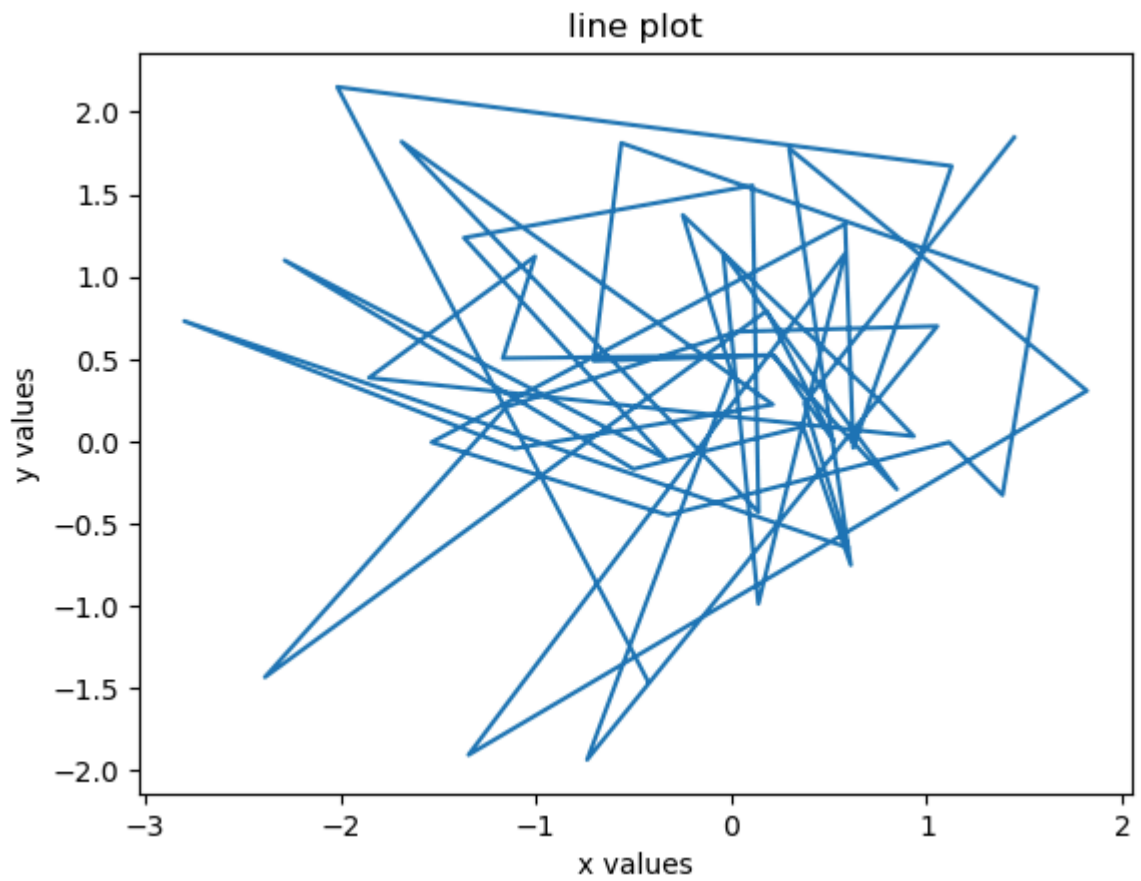
	Age	Salary	Purchased	Country_France	Country_Germany	Country_Spain
0	44.0	72000.0	No	True	False	False
1	27.0	48000.0	YES	False	False	True
2	30.0	54000.0	No	False	True	False
3	38.0	61000.0	No	False	False	True
4	40.0	NaN	YES	False	True	False
5	35.0	58000.0	YES	True	False	False
6	NaN	52000.0	No	False	False	True
7	48.0	79000.0	YES	True	False	False
8	50.0	83000.0	No	False	True	False
9	37.0	67000.0	Yes	True	False	False

	Country	Age	Salary	Purchased
0	France	44.0	72000.0	0
1	Spain	27.0	48000.0	1
2	Germany	30.0	54000.0	0
3	Spain	38.0	61000.0	0
4	Germany	40.0	NaN	1
5	France	35.0	58000.0	1
6	Spain	NaN	52000.0	0
7	France	48.0	79000.0	1
8	Germany	50.0	83000.0	0
9	France	37.0	67000.0	2

slip 9

que 2 a

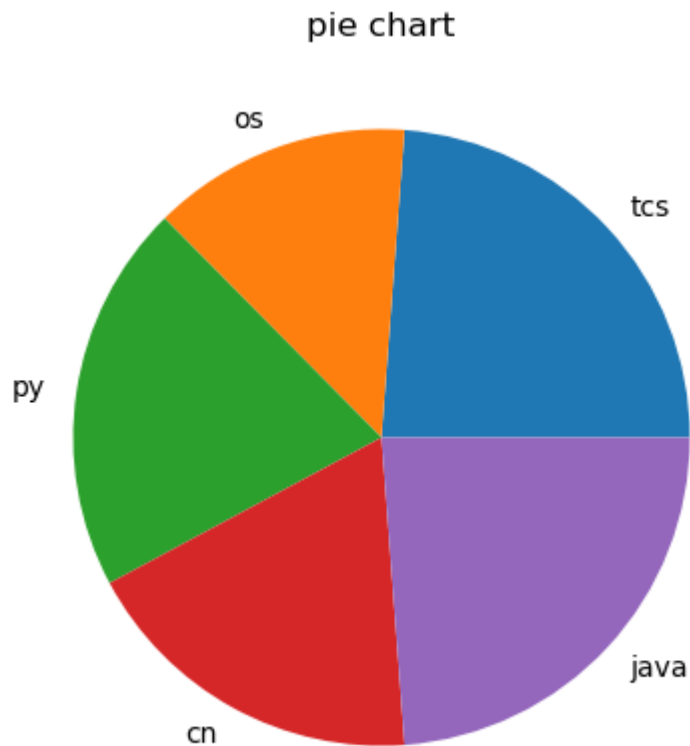
```
In [80]: x=np.random.randn(50)
y=np.random.randn(50)
plt.plot(x,y)
plt.xlabel("x values")
plt.ylabel("y values")
plt.title("line plot")
plt.show()
plt.scatter(x,y)
plt.xlabel("x values")
plt.ylabel("y values")
plt.title("scatter plot")
plt.show()
```



2 b

```
In [87]: subject=['tcs','os','py','cn','java']  
marks=[35,20,30,27,35]
```

```
plt.figure(figsize=(5,5))
plt.title("pie chart")
plt.pie(marks,labels=subject)
plt.show()
```



2 c

```
In [89]: df=pd.read_csv('C:/Users/PRANJAL/Desktop/html slips/winequality-red.csv')
df.shape
```

```
Out[89]: (1599, 1)
```

```
In [91]: df.size
```

```
Out[91]: 1599
```

```
In [93]: df.describe()
```

```
Out[93]:
```

	fixed acidity;"volatile acidity";"citric acid";"residual sugar";"chlorides";"free sulfur dioxide";"total sulfur dioxide";"density";"pH";"sulphates";"alcohol";"quality"
count	1599
unique	1359
top	7.5;0.51;0.02;1.7;0.084;13;31;0.99538;3.36;0.5...
freq	4

```
In [97]: df.head(3)
```

```
Out[97]:      fixed acidity;volatile acidity;citric acid;residual sugar;chlorides;free sulfur
          dioxide;total sulfur dioxide;density;pH;sulphates;alcohol;quality"

0      7.4;0.7;0;1.9;0.076;11;34;0.9978;3.51;0.56;9.4;5
1      7.8;0.88;0;2.6;0.098;25;67;0.9968;3.2;0.68;9.8;5
2      7.8;0.76;0.04;2.3;0.092;15;54;0.997;3.26;0.65;...
```

slip 10

que 2 a

```
In [100... df=pd.read_csv('C:/Users/PRANJAL/Desktop/html slips/SOCR-HeightWeight.csv')
print("mean is ",df.mean)
print("median is ",df.median)
```

mean is <bound method DataFrame.mean of				Index	Height(Inches)	Weight(Pou nds)
0	1	65.78331	112.9925			
1	2	71.51521	136.4873			
2	3	69.39874	153.0269			
3	4	68.21660	142.3354			
4	5	67.78781	144.2971			
...			
24995	24996	69.50215	118.0312			
24996	24997	64.54826	120.1932			
24997	24998	64.69855	118.2655			
24998	24999	67.52918	132.2682			
24999	25000	68.87761	124.8742			


```
[25000 rows x 3 columns]>
```

median is <bound method DataFrame.median of				Index	Height(Inches)	Weight
0	1	65.78331	112.9925			
1	2	71.51521	136.4873			
2	3	69.39874	153.0269			
3	4	68.21660	142.3354			
4	5	67.78781	144.2971			
...			
24995	24996	69.50215	118.0312			
24996	24997	64.54826	120.1932			
24997	24998	64.69855	118.2655			
24998	24999	67.52918	132.2682			
24999	25000	68.87761	124.8742			


```
[25000 rows x 3 columns]>
```

que 2 b

```
In [103... def sumdistance(x,y,n):
    sum=0
    for i in range(n):
        for j in range(i+1,n):
            sum+=(abs(x[i]-x[j])+abs(y[i]-y[j]))
    return sum
```

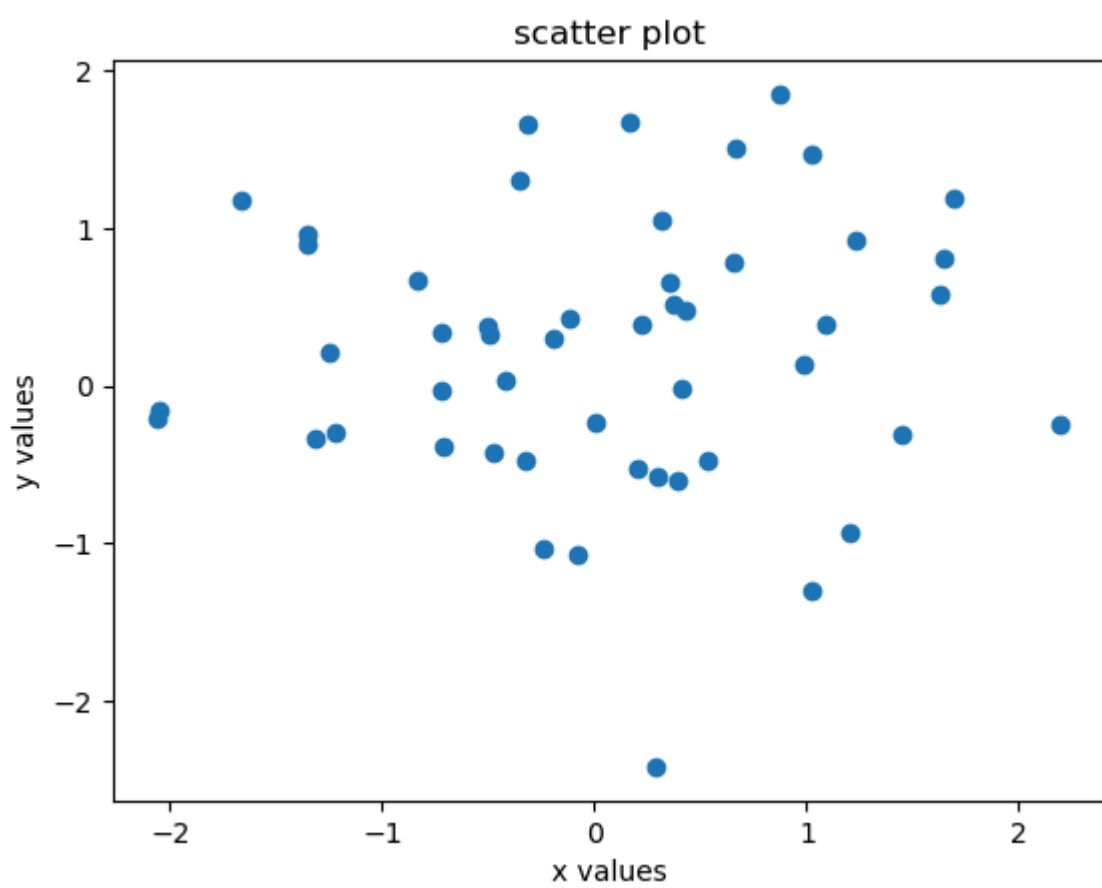
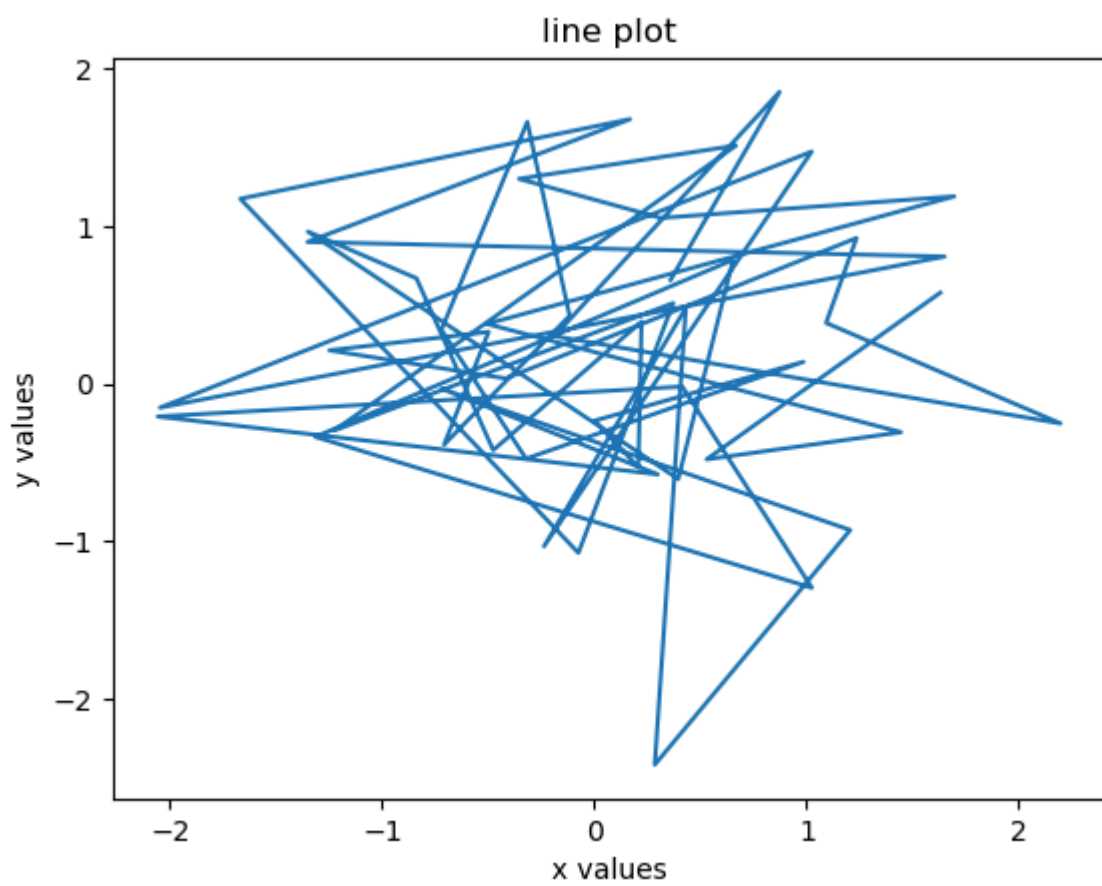
```
x=[1,-1,2,3]
y=[5,2,5,6]
n=len(x)
print(sumdistan(x,y,n))
```

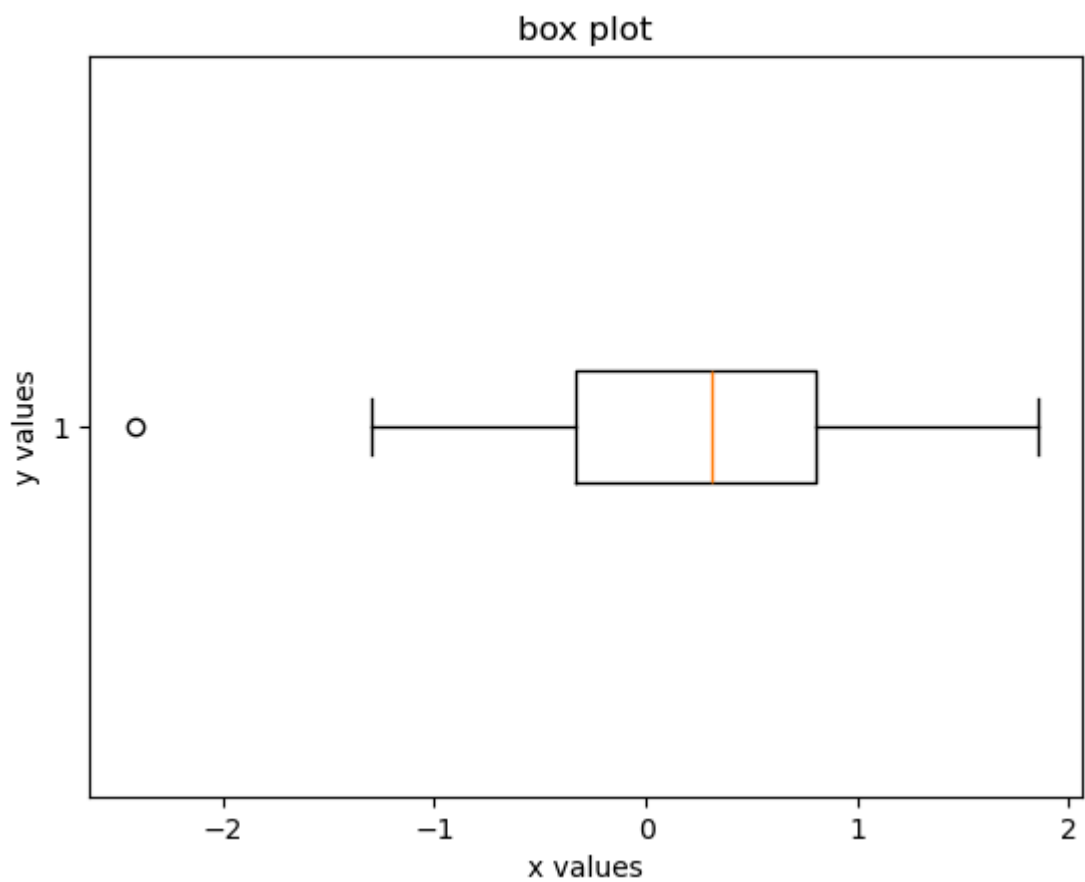
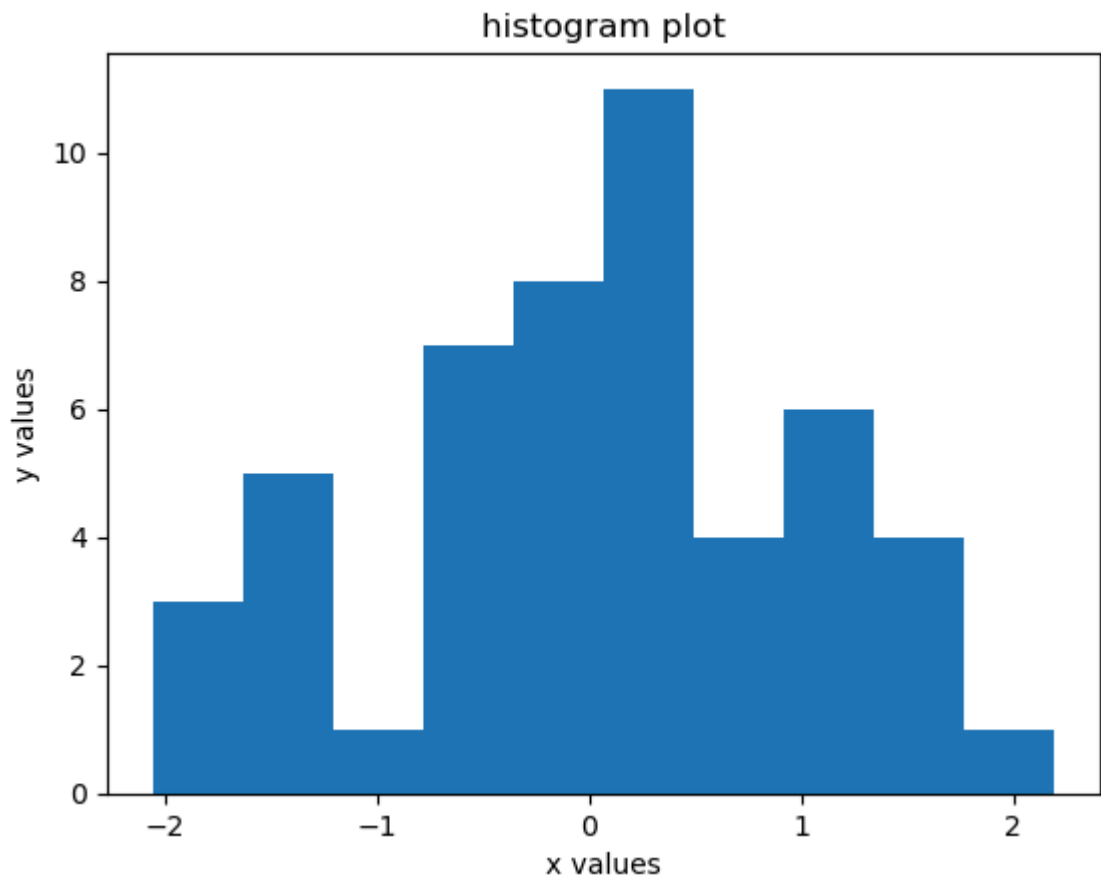
5

slip 12

que 2 a

```
In [105... x=np.random.randn(50)
y=np.random.randn(50)
plt.plot(x,y)
plt.xlabel("x values")
plt.ylabel("y values")
plt.title("line plot")
plt.show()
plt.scatter(x,y)
plt.xlabel("x values")
plt.ylabel("y values")
plt.title("scatter plot")
plt.show()
plt.hist(x)
plt.xlabel("x values")
plt.ylabel("y values")
plt.title("histogram plot")
plt.show()
plt.boxplot(y,vert=False)
plt.xlabel("x values")
plt.ylabel("y values")
plt.title("box plot")
plt.show()
```



que 2 b

In [108... `df=pd.DataFrame({'name':['sai',None,'ram','raj','om','sahil',None,'ram','rohit'],
'sal':[None,2000,4000,3000,1000,5000,None,4000,3000,1000],`

```

'dept':['cs','mechanical','cs','elec',None,'cs',None,'mechanical']
print(df)

```

	name	sal	dept
0	sai	NaN	cs
1	None	2000.0	mechanical
2	ram	4000.0	cs
3	raj	3000.0	elec
4	om	1000.0	None
5	sahil	5000.0	cs
6	None	NaN	None
7	ram	4000.0	mechanical
8	rohit	3000.0	cs
9	nikhil	1000.0	elec

In [113... df.dropna()

Out[113...

	name	sal	dept
2	ram	4000.0	cs
3	raj	3000.0	elec
5	sahil	5000.0	cs
7	ram	4000.0	mechanical
8	rohit	3000.0	cs
9	nikhil	1000.0	elec

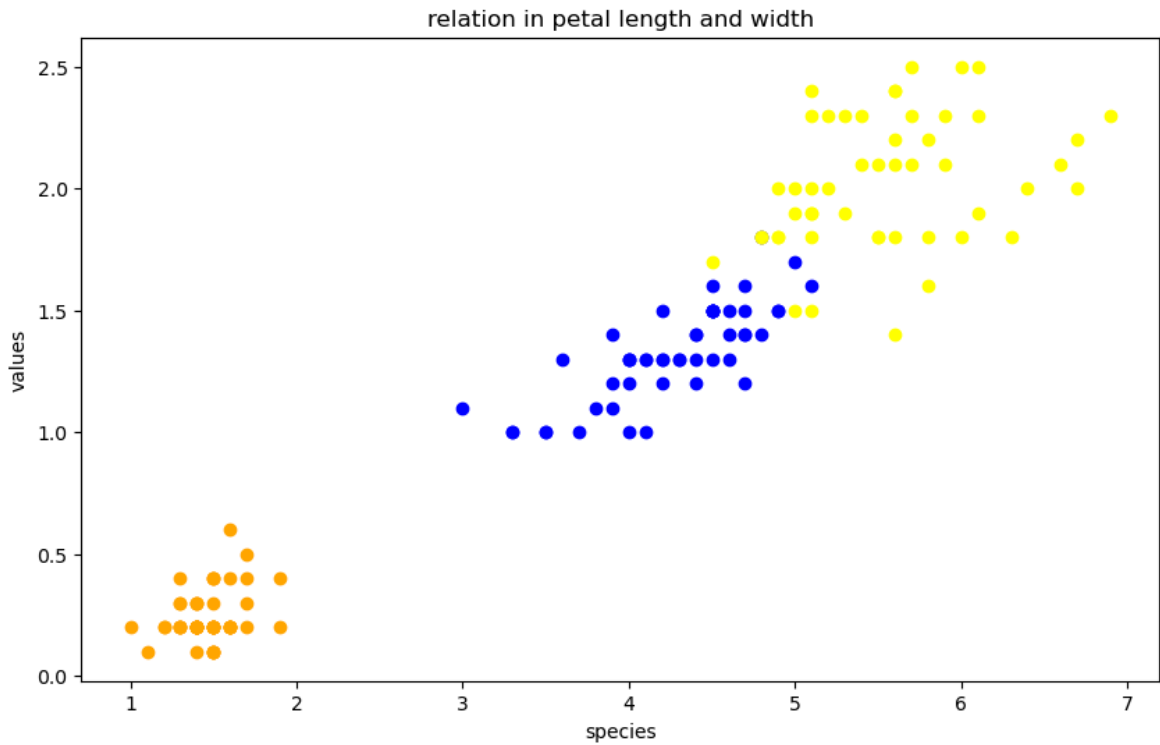
slip 13

que 2 a

```

df=pd.read_csv('C:/Users/PRANJAL/Desktop/html slips/IRIS.csv')
colors={'Iris-setosa':'orange','Iris-versicolor':'blue','Iris-virginica':'yellow'}
plt.figure(figsize=(10,6))
for species,color in colors.items():
    subset=df[df['species']==species]
    plt.scatter(subset['petal_length'],subset['petal_width'],c=color,label=species)
plt.xlabel('species')
plt.ylabel('values')
plt.title('relation in petal length and width')
plt.show()

```



que 2 b

In [119...

```
df=np.array([[2,3],[4,8]])
print(array)
print("maximum is ",df.max())
print("minimum is ",df.min())
```

```
[[2, 3], [4, 8]]
maximum is 8
minimum is 2
```

slip 14

que 2 a

In [120...

```
array=np.arange(5)
print(array)
weights=np.arange(10,15)
print(weights)
res=np.average(array,weights=weights)
print(res)
```

```
[0 1 2 3 4]
[10 11 12 13 14]
2.1666666666666665
```

que 2 b

In [122...

```
df=pd.read_csv('C:/Users/PRANJAL/Desktop/html slips/IRIS.csv')
print(df)
```

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa
3	4.6	3.1	1.5	0.2	Iris-setosa
4	5.0	3.6	1.4	0.2	Iris-setosa
..
145	6.7	3.0	5.2	2.3	Iris-virginica
146	6.3	2.5	5.0	1.9	Iris-virginica
147	6.5	3.0	5.2	2.0	Iris-virginica
148	6.2	3.4	5.4	2.3	Iris-virginica
149	5.9	3.0	5.1	1.8	Iris-virginica

[150 rows x 5 columns]

In [124... `df.shape`

Out[124... (150, 5)

In [126... `df.size`

Out[126... 750

In [128... `df.dtypes`

Out[128...
sepal_length float64
sepal_width float64
petal_length float64
petal_width float64
species object
dtype: object

In [130... `df.describe()`

Out[130...

	sepal_length	sepal_width	petal_length	petal_width
count	150.000000	150.000000	150.000000	150.000000
mean	5.843333	3.054000	3.758667	1.198667
std	0.828066	0.433594	1.764420	0.763161
min	4.300000	2.000000	1.000000	0.100000
25%	5.100000	2.800000	1.600000	0.300000
50%	5.800000	3.000000	4.350000	1.300000
75%	6.400000	3.300000	5.100000	1.800000
max	7.900000	4.400000	6.900000	2.500000

In [132... `df.info()`

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 5 columns):
#   Column          Non-Null Count  Dtype
---  -
0   sepal_length    150 non-null   float64
1   sepal_width     150 non-null   float64
2   petal_length    150 non-null   float64
3   petal_width     150 non-null   float64
4   species         150 non-null   object
dtypes: float64(4), object(1)
memory usage: 6.0+ KB

```

slip 15

que 2 a,slip 30(que 2a),slip 26(2a),slip 25(2a),slip 20(2a)

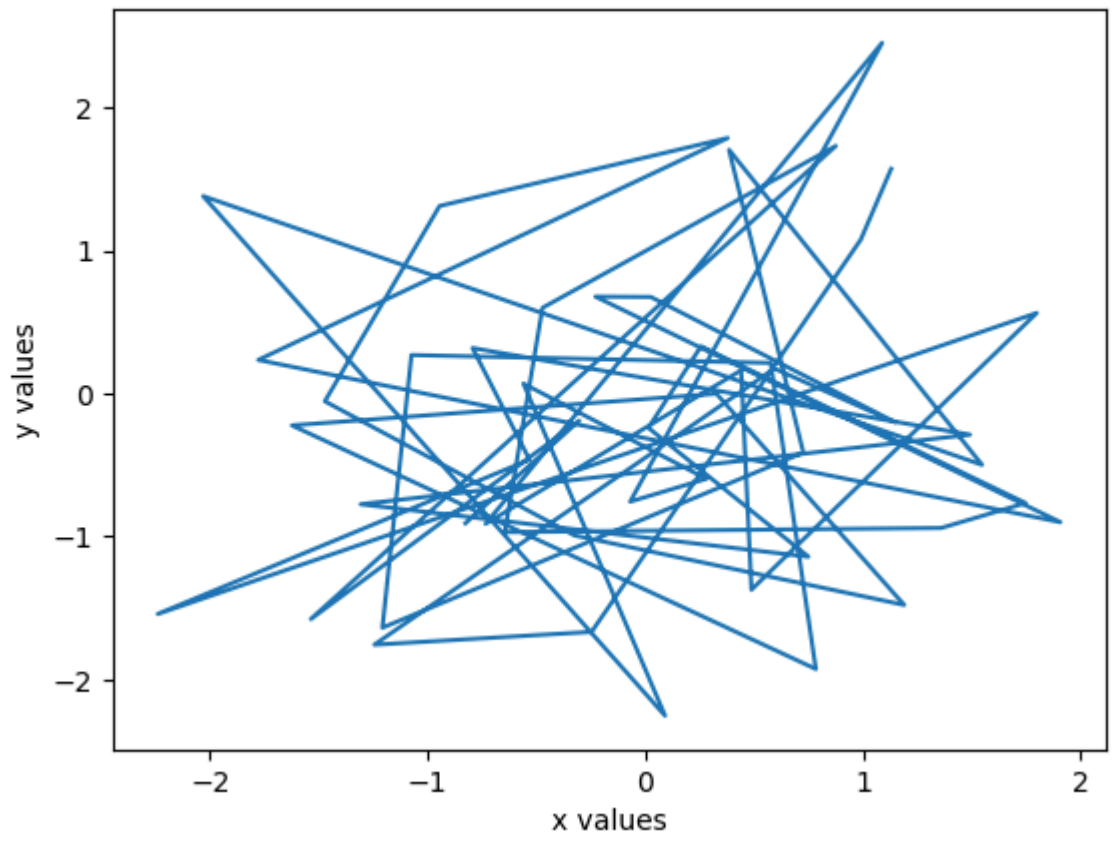
In [134...

```

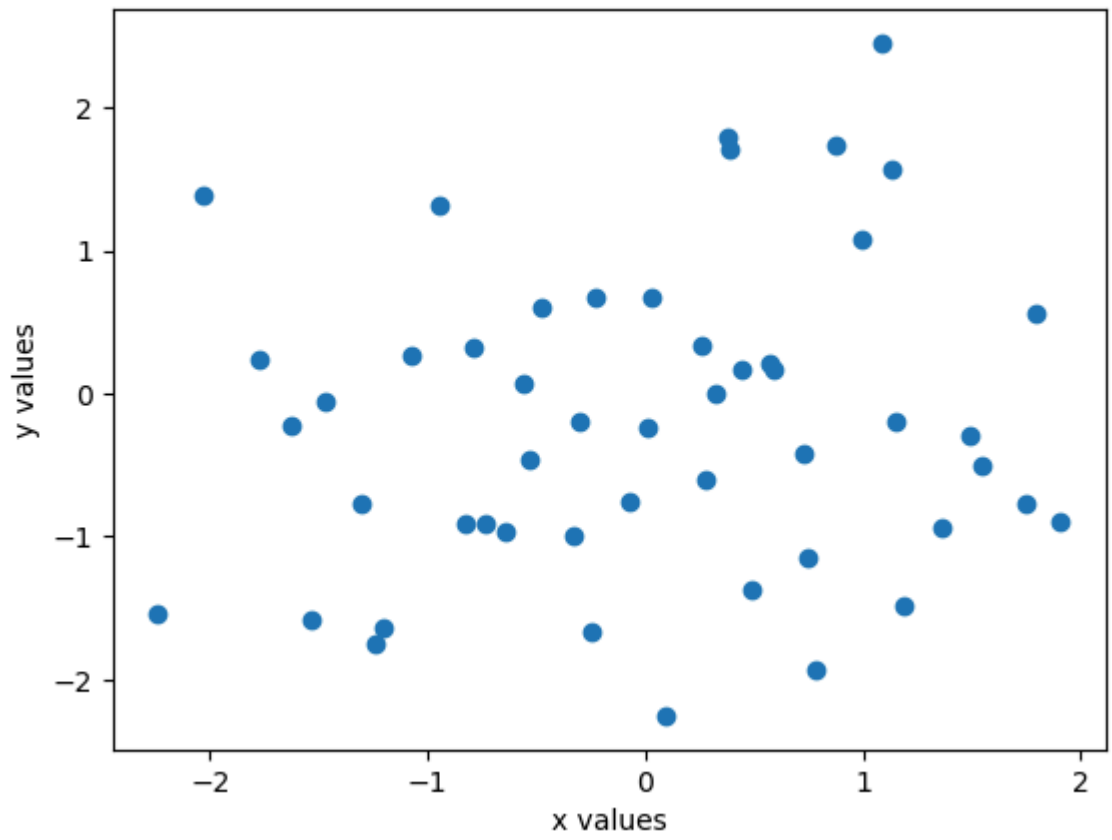
x=np.random.randn(50)
y=np.random.randn(50)
plt.plot(x,y)
plt.xlabel("x values")
plt.ylabel("y values")
plt.title("line plot")
plt.show()
plt.scatter(x,y)
plt.xlabel("x values")
plt.ylabel("y values")
plt.title("scatter plot")
plt.show()
plt.hist(x)
plt.xlabel("x values")
plt.ylabel("y values")
plt.title("histogram plot")
plt.show()
plt.boxplot(y,vert=False)
plt.xlabel("x values")
plt.ylabel("y values")
plt.title("box plot")
plt.show()

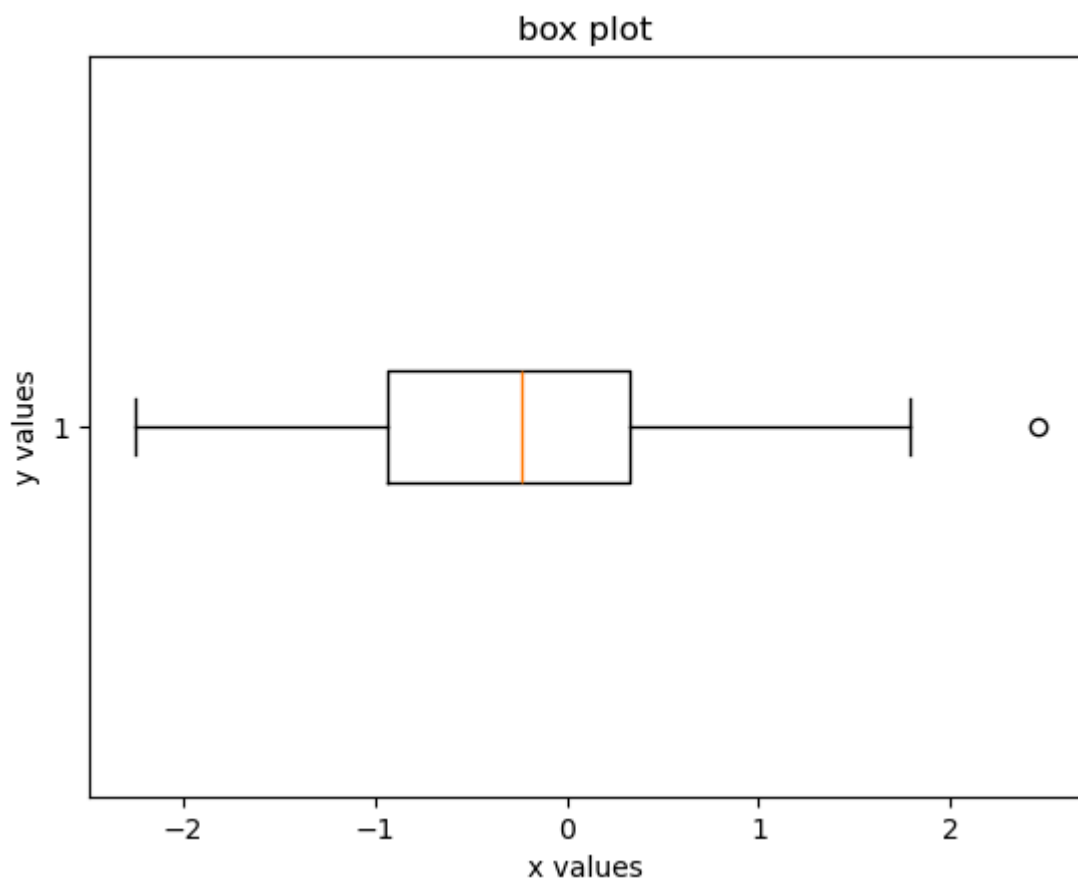
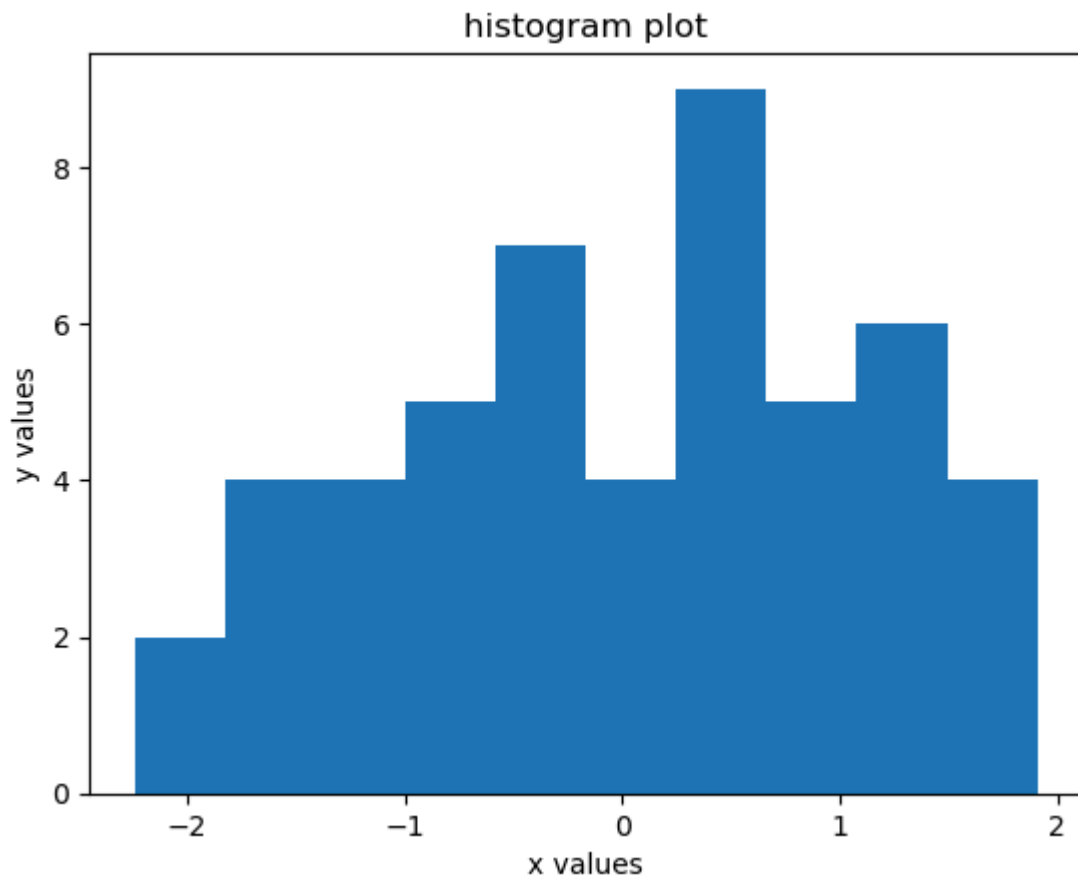
```

line plot



scatter plot

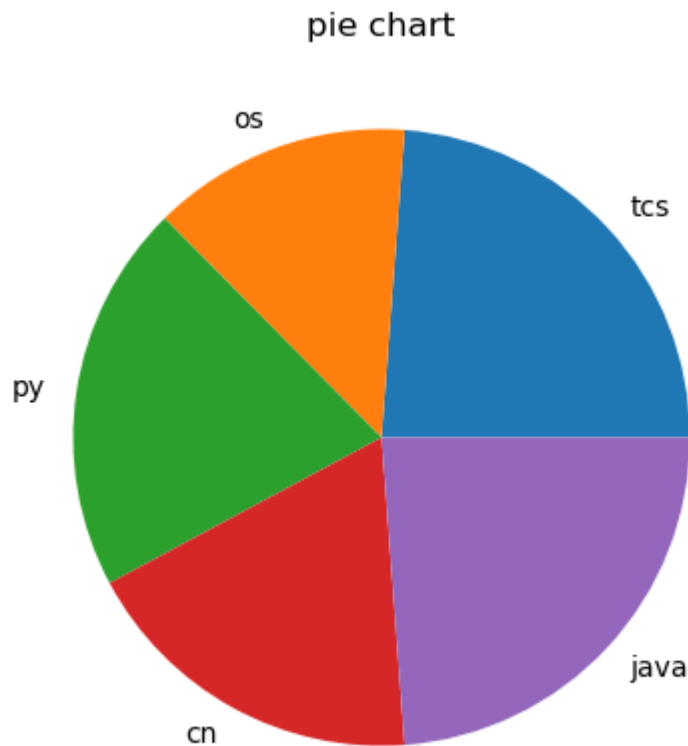




que 2 b

In [135... subject=['tcs','os','py','cn','java']
marks=[35,20,30,27,35]


```
plt.figure(figsize=(5,5))
plt.title("pie chart")
plt.pie(marks,labels=subject)
plt.show()
```

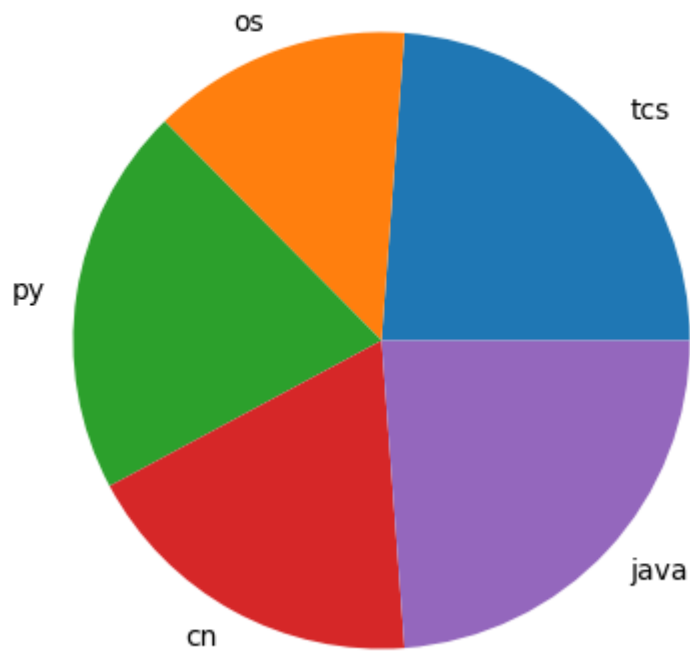


slip 16

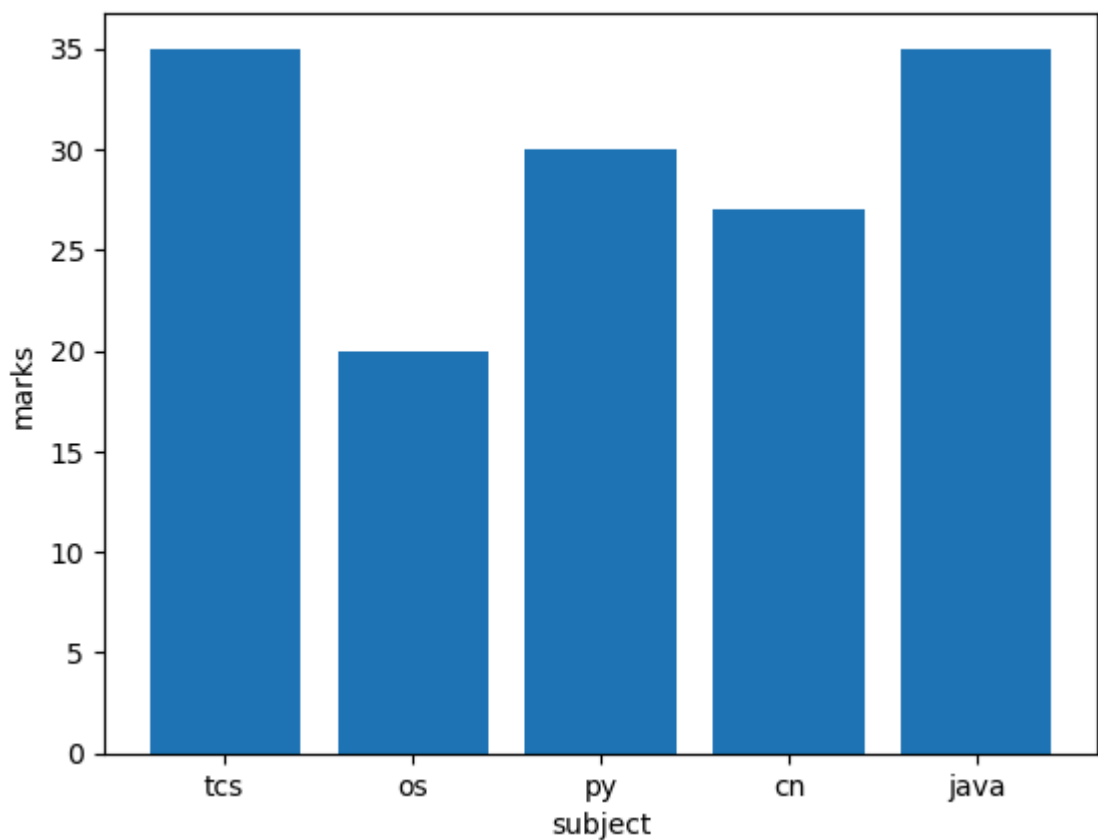
que 2a,slip 26(2b),slip 30(2b),slip 25(2b)

```
In [139... subject=['tcs','os','py','cn','java']
marks=[35,20,30,27,35]
plt.figure(figsize=(5,5))
plt.title("pie chart")
plt.pie(marks,labels=subject)
plt.show()
plt.title("bar chart")
plt.xlabel('subject')
plt.ylabel('marks')
plt.bar(subject,marks)
plt.show()
```

pie chart



bar chart



que 2 b

```
In [141]: df=pd.DataFrame({'name':['kunal','rekha','satish','ashish','radha'],  
                           'age':[20,23,22,20,21],  
                           'per':[98,80,95,92,85]})
```

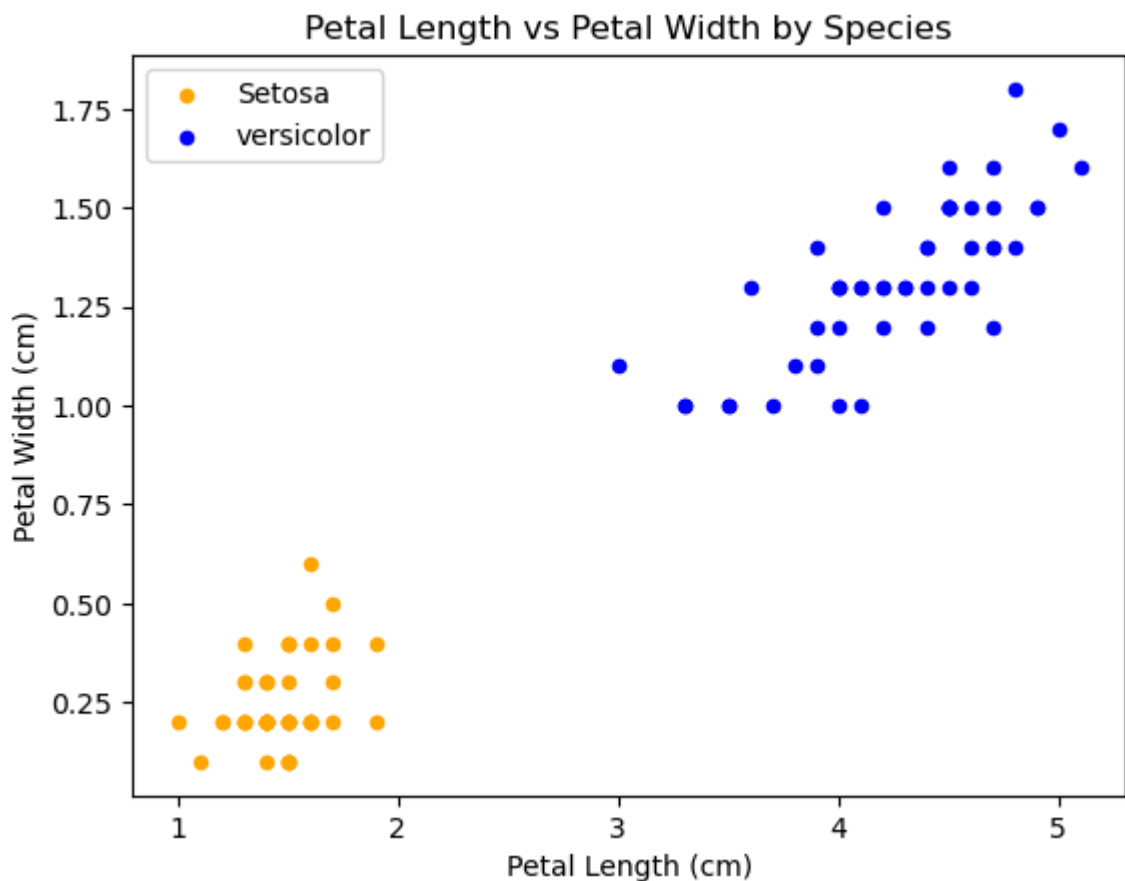
```
print(np.average(df['age']))
print(np.average(df['per']))
```

21.2
90.0

slip 17

que 2a

```
In [153... df= pd.read_csv('C:/Users/PRANJAL/Desktop/html slips/IRIS.csv')
# Plot scatter for each species on the same figure
fig = df[df.species == 'Iris-setosa'].plot.scatter(x='petal_length', y='petal_wi
df[df.species == 'Iris-versicolor'].plot.scatter(x='petal_length', y='petal_wid
fig.set_xlabel("Petal Length (cm)")
fig.set_ylabel("Petal Width (cm)")
fig.set_title("Petal Length vs Petal Width by Species")
plt.legend()
plt.show()
```



que 2b

```
In [146... df=pd.DataFrame({'name':['kunal','rekha','satish','ashish','radha'],
                    'age':[20,23,22,20,21],
                    'salary':[100000,300000,20000,300000,80000] })
df
```

Out[146...

	name	age	salary
0	kunal	20	100000
1	rekha	23	300000
2	satish	22	20000
3	ashish	20	300000
4	radha	21	80000

slip 18

que 2 a

In [143...

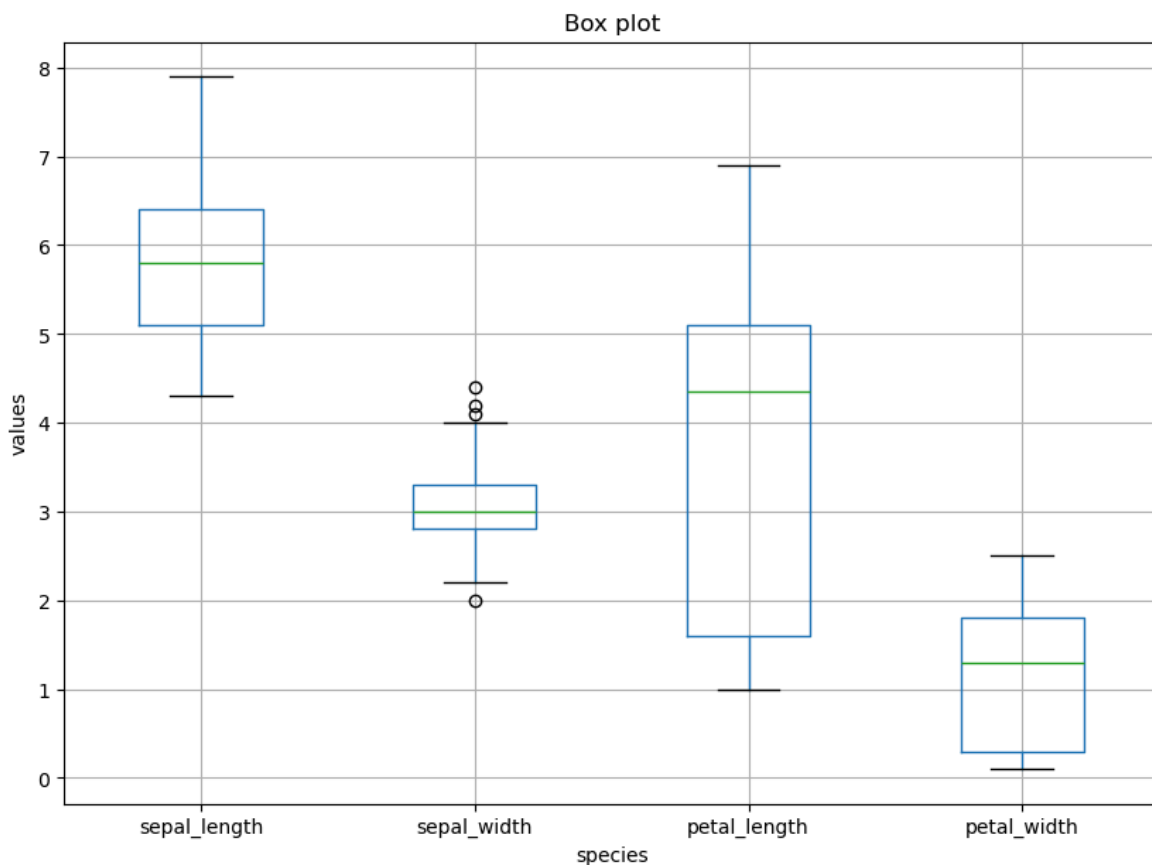
```
df=pd.read_csv('C:/Users/PRANJAL/Desktop/html slips/IRIS.csv')
data=df[["sepal_length","sepal_width","petal_length","petal_width"]]
print(data)
plt.figure(figsize=(10,7))
plt.title('Box plot')
plt.xlabel('species')
plt.ylabel('values')
data.boxplot()
```

	sepal_length	sepal_width	petal_length	petal_width
0	5.1	3.5	1.4	0.2
1	4.9	3.0	1.4	0.2
2	4.7	3.2	1.3	0.2
3	4.6	3.1	1.5	0.2
4	5.0	3.6	1.4	0.2
..
145	6.7	3.0	5.2	2.3
146	6.3	2.5	5.0	1.9
147	6.5	3.0	5.2	2.0
148	6.2	3.4	5.4	2.3
149	5.9	3.0	5.1	1.8

[150 rows x 4 columns]

Out[143...

<Axes: title={'center': 'Box plot'}, xlabel='species', ylabel='values'>



que 2b

In [144...]

```
df=pd.read_csv('C:/Users/PRANJAL/Desktop/html slips/ds.csv')
print("first 5 rows is ",df.head(5));
print("\nlast 5 rows is ",df.tail(5));
print("\n random rows is ",df.sample(10));
```

```
first 5 rows is      Country  Age  Salary Purchased
0   France  44.0  72000.0    No
1    Spain  27.0  48000.0   YES
2  Germany  30.0  54000.0    No
3    Spain  38.0  61000.0    No
4  Germany  40.0    NaN    YES
```

```
last 5 rows is      Country  Age  Salary Purchased
5   France  35.0  58000.0   YES
6    Spain  NaN  52000.0    No
7   France  48.0  79000.0   YES
8  Germany  50.0  83000.0    No
9   France  37.0  67000.0   Yes
```

```
random rows is      Country  Age  Salary Purchased
1    Spain  27.0  48000.0   YES
9   France  37.0  67000.0   Yes
6    Spain  NaN  52000.0    No
4  Germany  40.0    NaN    YES
5   France  35.0  58000.0   YES
2  Germany  30.0  54000.0    No
0   France  44.0  72000.0    No
7   France  48.0  79000.0   YES
3    Spain  38.0  61000.0    No
8  Germany  50.0  83000.0    No
```

slip 19

que 2a

```
In [155... df=pd.DataFrame(columns=['name','age','percentage'])
df
```

```
Out[155...    name  age  percentage
```

```
In [157... df.loc[1]=['sai',20,89.90]
df.loc[2]=['raj',19,78.45]
df.loc[3]=['om',20,76.12]
df.loc[4]=['ram',20,92.04]
df.loc[5]=['sejal',19,56.99]
df.loc[6]=['payal',20,69.34]
df.loc[7]=['nikita',19,74.67]
df.loc[8]=['sahil',19,96.11]
df.loc[9]=['soham',19,81.03]
df.loc[10]=['pallavi',20,59.50]
df
```

```
Out[157...    name  age  percentage
```

1	sai	20	89.90
2	raj	19	78.45
3	om	20	76.12
4	ram	20	92.04
5	sejal	19	56.99
6	payal	20	69.34
7	nikita	19	74.67
8	sahil	19	96.11
9	soham	19	81.03
10	pallavi	20	59.50

```
In [159... df.shape
```

```
Out[159... (10, 3)
```

```
In [161... df.size
```

```
Out[161... 30
```

```
In [163... df.dtypes
```

```
Out[163...  name          object
          age      int64
          percentage float64
          dtype: object
```

```
In [165... df.describe()
```

```
Out[165...      age  percentage
count  10.000000    10.000000
mean    19.500000    77.415000
std      0.527046    13.106207
min     19.000000    56.990000
25%     19.000000    70.672500
50%     19.500000    77.285000
75%     20.000000    87.682500
max     20.000000    96.110000
```

```
In [167... df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 10 entries, 1 to 10
Data columns (total 3 columns):
#   Column      Non-Null Count  Dtype
---  -
0   name         10 non-null     object
1   age          10 non-null     int64
2   percentage   10 non-null     float64
dtypes: float64(1), int64(1), object(1)
memory usage: 620.0+ bytes
```

```
In [169... df.loc[11]=['sai',20,89.90]
df.loc[12]=['sahil',None,96.11]
df.loc[13]=['nikita',19,74.67]
df.loc[14]=['sejal',19,None]
df.loc[15]=[None,None,None]
df
```

Out[169...

	name	age	percentage
1	sai	20	89.90
2	raj	19	78.45
3	om	20	76.12
4	ram	20	92.04
5	sejal	19	56.99
6	payal	20	69.34
7	nikita	19	74.67
8	sahil	19	96.11
9	soham	19	81.03
10	pallavi	20	59.50
11	sai	20	89.90
12	sahil	None	96.11
13	nikita	19	74.67
14	sejal	19	NaN
15	None	None	NaN

In [171...

```
df["remark"]=None  
df
```


Out[171...

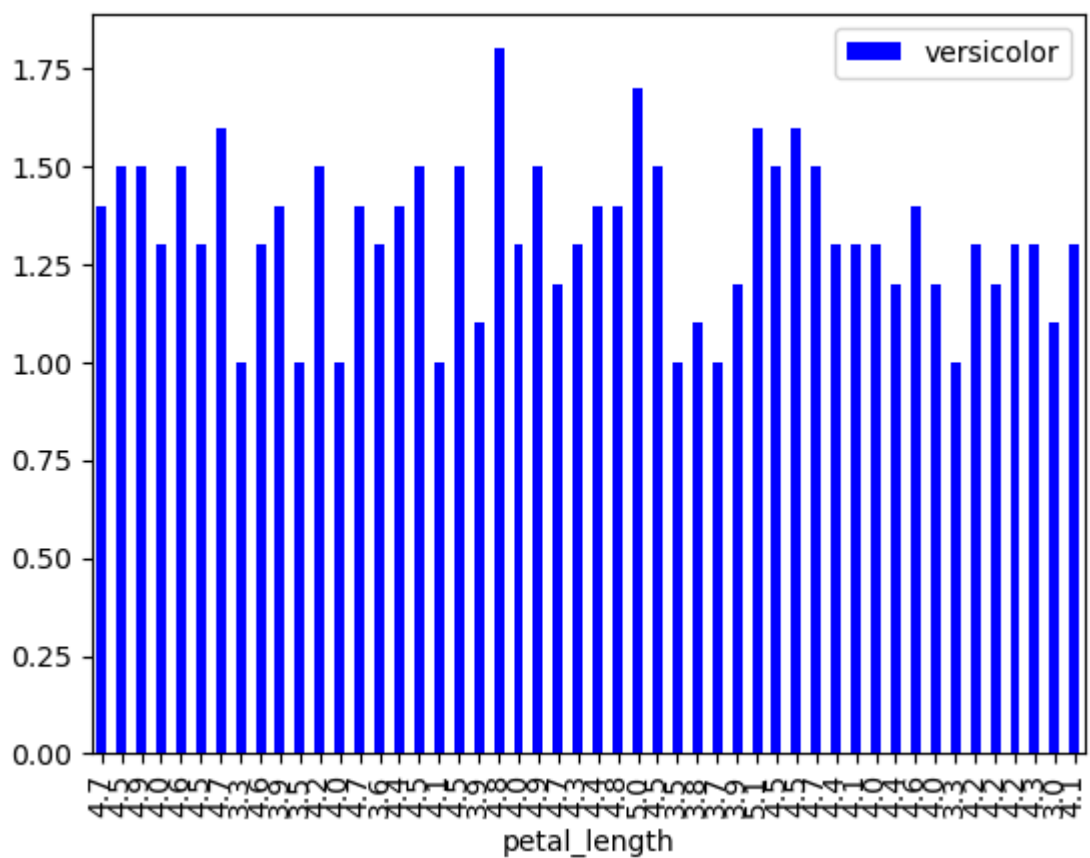
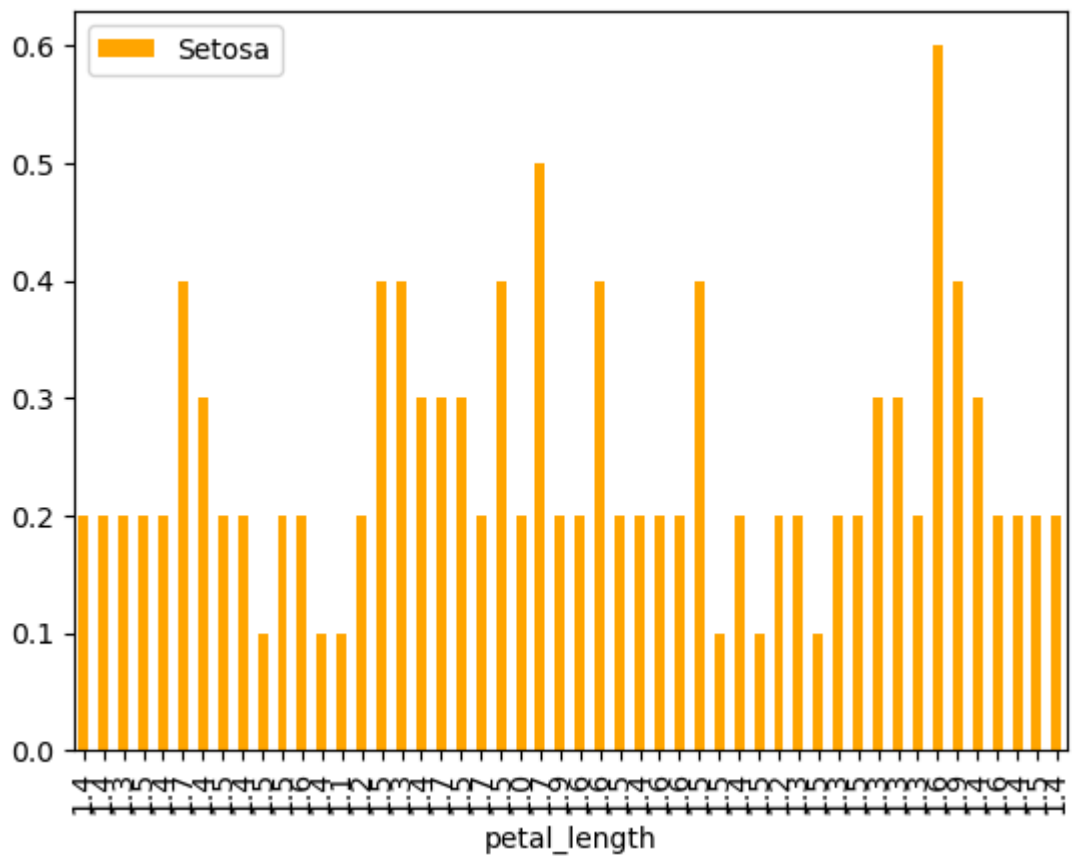
	name	age	percentage	remark
1	sai	20	89.90	None
2	raj	19	78.45	None
3	om	20	76.12	None
4	ram	20	92.04	None
5	sejal	19	56.99	None
6	payal	20	69.34	None
7	nikita	19	74.67	None
8	sahil	19	96.11	None
9	soham	19	81.03	None
10	pallavi	20	59.50	None
11	sai	20	89.90	None
12	sahil	None	96.11	None
13	nikita	19	74.67	None
14	sejal	19	NaN	None
15	None	None	NaN	None

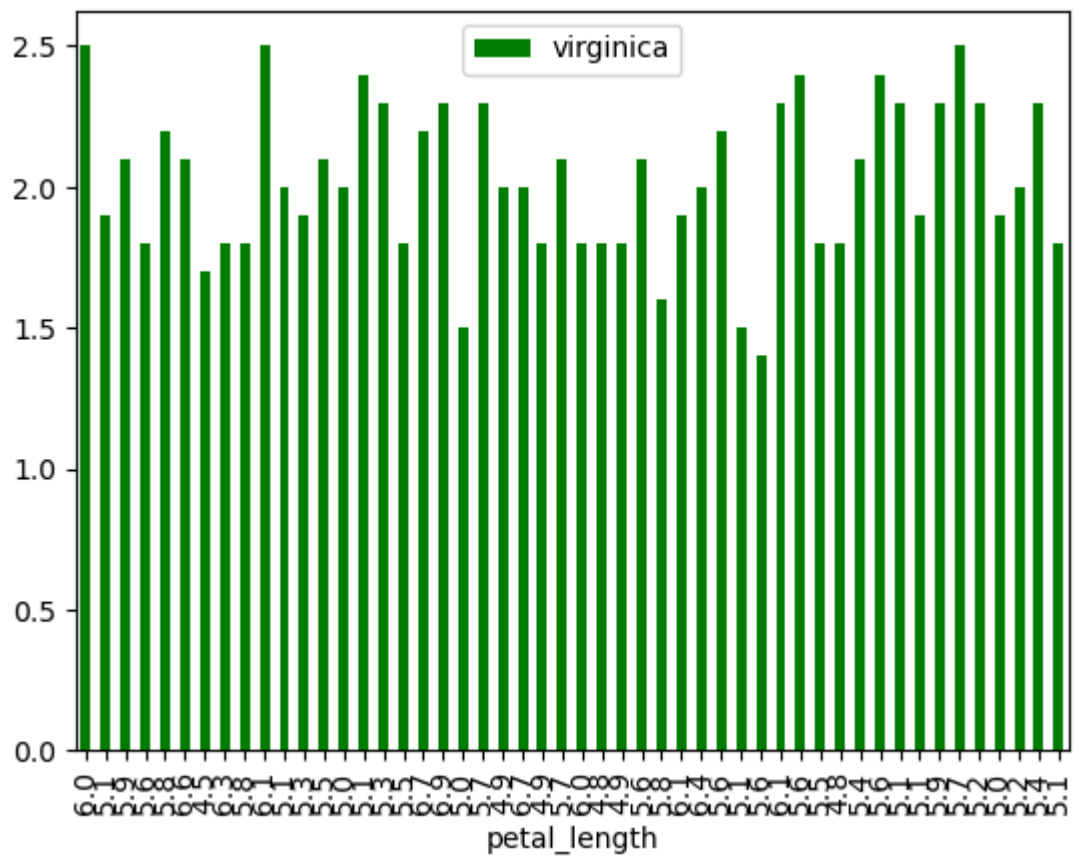
slip 21,24

que 2a

In [175...

```
df=pd.read_csv('C:/Users/PRANJAL/Desktop/html slips/IRIS.csv')
df[df.species=='Iris-setosa'].plot.bar(x='petal_length',y='petal_width',color='c')
df[df.species=='Iris-versicolor'].plot.bar(x='petal_length',y='petal_width',color='m')
df[df.species=='Iris-virginica'].plot.bar(x='petal_length',y='petal_width',color='b')
plt.show()
```

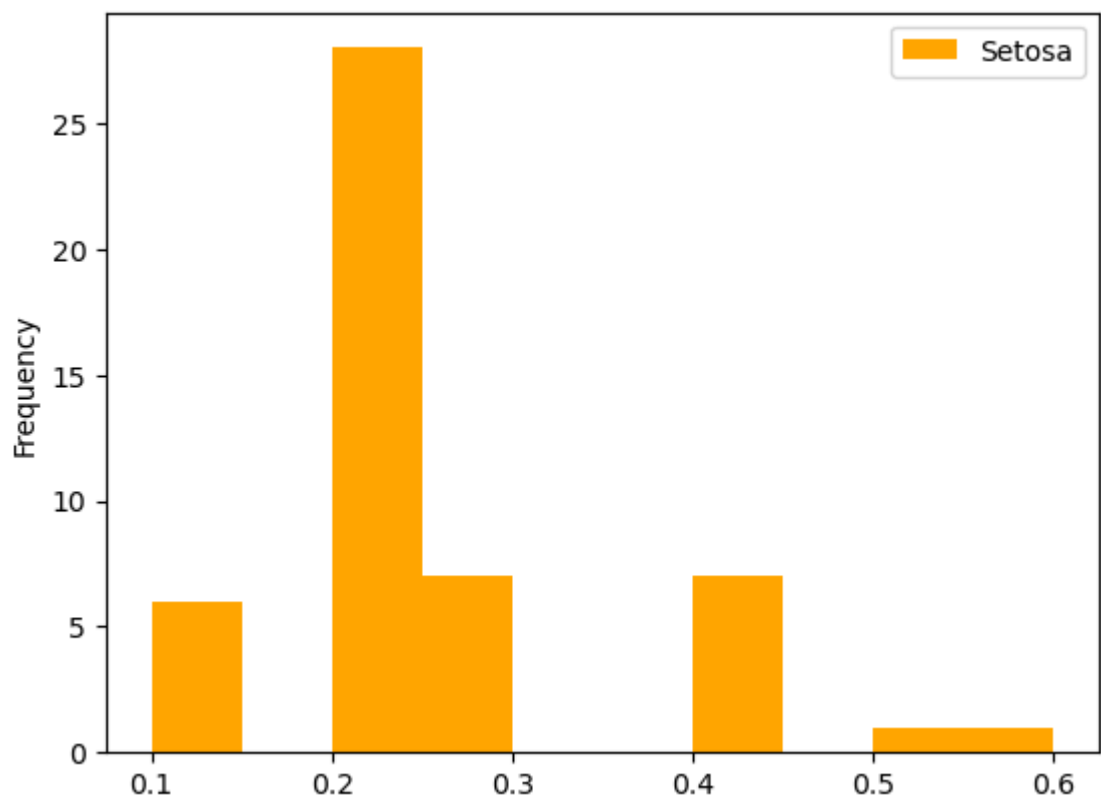


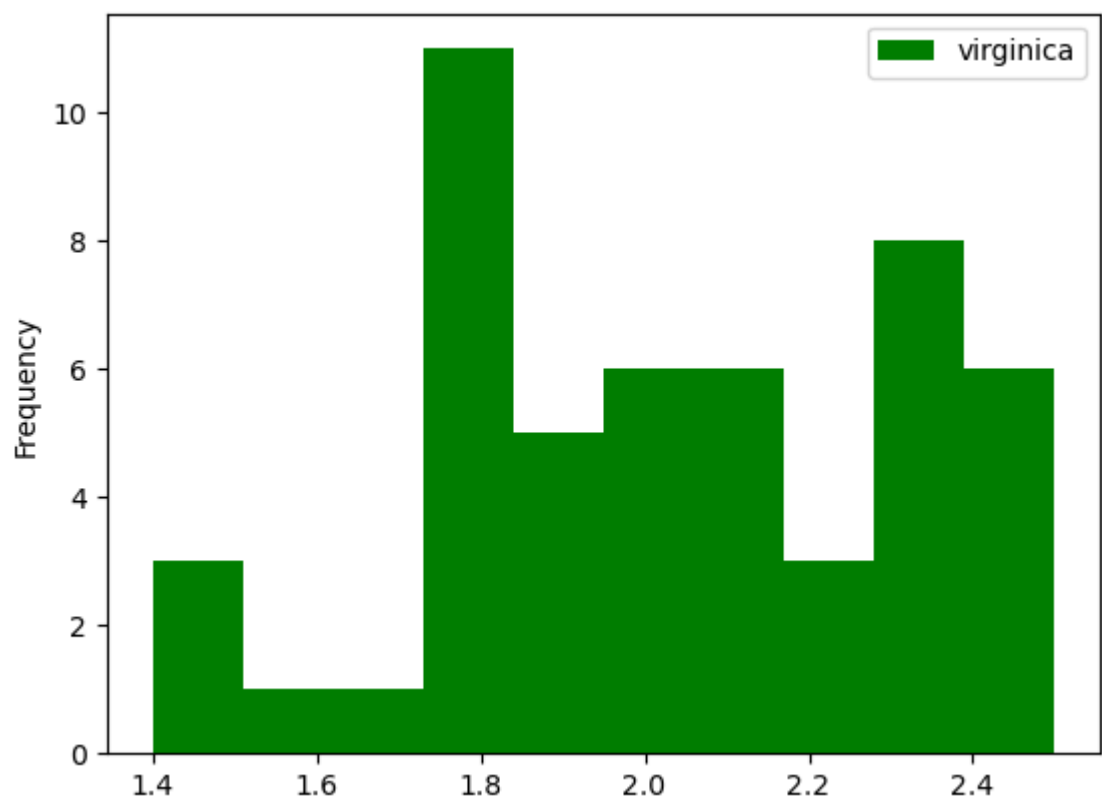
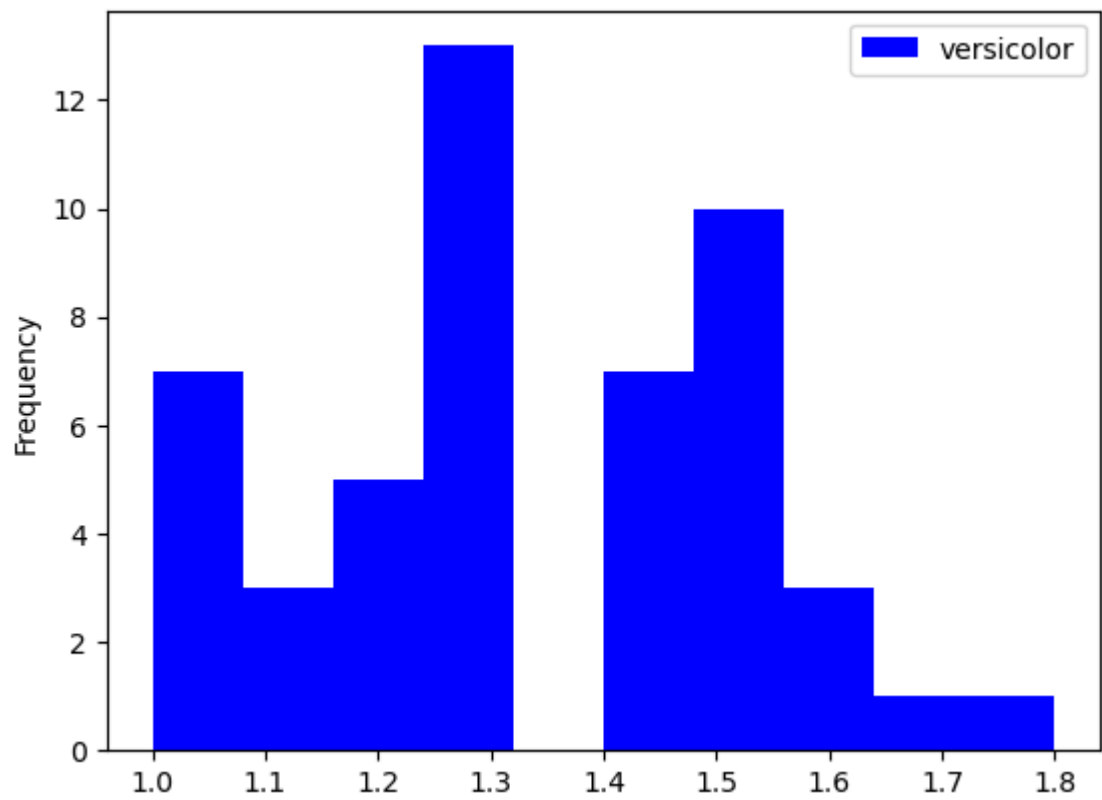


que 2b

In [176...

```
df=pd.read_csv('C:/Users/PRANJAL/Desktop/html slips/IRIS.csv')
df[df.species=='Iris-setosa'].plot.hist(x='petal_length',y='petal_width',color='
df[df.species=='Iris-versicolor'].plot.hist(x='petal_length',y='petal_width',col
df[df.species=='Iris-virginica'].plot.hist(x='petal_length',y='petal_width',colo
plt.show()
```





In []: