•

---

1) Slip 1 & 11

```
# Q 2 A
d=pd.read_csv('C:\\User\\Iris.csv')
ax=plt.subplots(1,1,figsize=(10,8))
d['Species'].value_counts().plot.pie()
plt.title("Iris Species %")
plt.show()
# Q 2 B
import pandas as p
df = pd.) read_csv('C:\\winequality-red.csv')
df.shape
df.describe()
df.info()
df.dtypes
```

---

2) Slip 2 & 6

```
#Q2 A)
import pandas as p
import numpy as n
d=p.read_csv('D:Data.csv')
v=d['age'].mean()
v1=d['salary'].mean()
d['age'].fillna(v,inplace=True)
d['salary'].fillna(v1,inplace=True)
print(d)
#Q 2 B
import numpy as np
```

```
import matplotlib.pyplot as plt
import pandas as p
df=p.DataFrame({'name' :
['kunal','rekha','satish','ashish','radha'],
'age':[20,23,22,20,21],
'per':[98,80,95,92,85],
'salary':[1000,3000,2000,3000,8000] })
df.plot(x="name",y="salary")
plt.show()
# Q 2 C
import pandas as p
df=p.read_csv("ht&wt.csv")
print("first 10 rows \n",df.head(10))
print("\n random 20 rows\n",df.sample(20))
print("\n shape \n",df.shape)
```

---

3) Slip 3

```
# Q 2 A
import pandas as p
d=p.read_csv('C:\Iris.csv')
new_data=d[["SepalLengthCm","SepalWidthCm",
"PetalLengthCm","PetalWidthCm"]]
print(new_data)
plt.figure(figsize=(10,7))
new_data.boxplot()
# Q 2 B
import pandas as p
df=pd.read_csv('C:\\Users\\Iris.csv')
df.shape
df.describe()
df.info()
df.dtypes
```

---

4) Slip 4 & 5

```
# Q 2 A
import matplotlib.pyplot as plt
import numpy as np
x=np.random.randn(50)
y=np.random.randn(50)
plt.plot(x,y)
plt.show()
plt.scatter(x,y)
plt.show()
plt.hist(x)
plt.show()
plt.boxplot(y,vert=False)
plt.show()
# Q 2 B
import pandas as p
df=pd.read_csv('C:\\Users\\Iris.csv')
df.shape
df.describe()
df.info()
df.dtypes
```

---

5) Slip 7 & 29

```
# Q 2
import pandas as pd
from sklearn import preprocessing
data = pd.read_csv("C:\\Data1.csv")
one_hot_encoded_data=pd.get_dummies(data,
columns=['Country'])
label_encoder=preprocessing.LabelEncoder()
one_hot_encoded_data['Purchased']=
```

label_encoder.fit_transform(one_hot_encoded_data

['Purchased'])

print("After applying OneHot and Label Encoding:")

print(one_hot_encoded_data)

---

6) Slip 9 & 15

```
# Q 2 A
import matplotlib.pyplot as plt
import numpy as np
from matplotlib import colors
from matplotlib.ticker import PercentFormatter
no_of_balls=50
x=np.random.randn(50)
y=np.random.randn(50)
colors=[np.random.randint(1, 4) for i in range(no_of_balls)]
plt.plot(x,y)
plt.show()
plt.scatter(x,y,c=colors)
plt.show()
# Q 2 B
from matplotlib import pyplot as plt
import numpy as np
subjects=['TCS','Data Science','OS','JAVA','PHP','Python']
marks=[23,17,35,29,12,33]
fig=plt.figure(figsize=(10,7))
plt.pie(marks,labels=subjects)
plt.show()
# Q 2 C
import pandas as p
df=pd.read_csv('C:\\winequality-red.csv')
print("\n",df.shape)
print("\n",df.describe())
```

```
df.head(3)
```

---

7) Slip 10

```
# Q 2 A
import pandas as p
df=p.read_csv("\\wt.csv")
print("mean is \n",df.mean)
print("median is \n",df.median)
# Q 2 B
def distancesum(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=[-1,1,3,2]
y=[5,6,5,3]
n=len(x)
print(distancesum(x, y, n))
```

---

8) Slip 12

```
# Q 2 A
import matplotlib.pyplot as plt
import numpy as np
x = np.random.randn(50)
y = np.random.randn(50)
plt.plot(x,y)
plt.show()
plt.scatter(x,y)
plt.show()
plt.hist(x)
plt.show()
```

```
plt.boxplot(y, vert=False)
plt.show()
# Q 2 B
import pandas as p
df=p.DataFrame({'name':
['kunal','rekha','satish','ashish','radha'],
'dept':['production','computer','manufacturing',
None,'manufacturing'],
'salary':[100000,300000,20000,300000,80000] })
print(df)
d=df.dropna()
print(d)
```

---

8) Slip 13

```
# Q 2 A
import pandas as p
import matplotlib.pyplot as plt
d=p.read_csv('C:\\Users\\Iris.csv')
fig = d[d.Species=='Irissetosa']
.plot.bar(x='PetalLengthCm',
y='PetalWidthCm',color='orange',
label='Setosa')
d[d.Species=='Irisversicolor']
.plot.bar(x='PetalLengthCm',y=
'PetalWidthCm',color='blue',
label='versicolor',ax=fig)
d[d.Species=='Iris-virginica']
.plot.bar(x='PetalLengthCm',y=
'PetalWidthCm',color='green',
label='virginica', ax=fig)
fig.set_xlabel("Petal Length")
fig.set_ylabel("Petal Width")
```

```
fig.set_title(" Petal Length VS Petal Width")
plt.show()
# Q 2 B
import numpy as n
d=n.array([[0,1],[2,3]])
print(d.max())
print(d.min())
```

---

9) Slip 14

```
#Q 2 A
import numpy as np
array=np.arange(5)
print(array)
weights = np.arange(10,15)
print(weights)
res1=np.average(array,weights=weights)
print(res1)
#Q 2 B
import pandas as p
df = pd.read_csv('C:\\Advertising.csv')
df.shape
df.describe()
df.info()
df.dtypes
```

---

10) Slip 16

```
# Q 2 A
from matplotlib import pyplot as plt
import numpy as np
subjects=['TCS','Data Science','OS','JAVA','PHP','Python']
marks=[23,17,35,29,12,33]
fig = plt.figure(figsize =(10, 7))
```

```
plt.pie(marks, labels = subjects)
plt.show()
# Q 2 B
import pandas as p
import numpy as n
df=p.DataFrame({'name':
['kunal','rekha','satish','ashish','radha'],
'age':[20,23,22,20,21],
'per':[98,80,95,92,85]})
print(n.average(df['age']))
print(n.average(df['per']))
```

---

11) Slip 17

```
#Q 2 A
import pandas as p
import matplotlib.pyplot as plt
d=p.read_csv('C:\\Users\\Iris.csv')
fig = d[d.Species=='Iris-setosa'].
plot.scatter(x='PetalLengthCm',y=
'PetalWidthCm',color='orange',
label='Setosa')
d[d.Species=='Iris-versicolor'].
plot.scatter(x='PetalLengthCm',y=
'PetalWidthCm',color='blue',label='versicolor',ax=fig)
fig.set_xlabel("Petal Length")
fig.set_ylabel("Petal Width")
fig.set_title(" Petal Length VS Petal Width")
plt.show()
# Q 2 B
import pandas as p
df=p.DataFrame({'name':
['kunal','rekha','satish','ashish','radha'],
```

'age':[20,23,22,20,21],

'salary':[100000,300000,20000,300000,80000] })

df

---

12) Slip 18

```
# Q 2 A
import pandas as p
d=p.read_csv('C:\\Users\\Iris.csv')
new_data=d[["SepalLengthCm","SepalWidthCm",
"PetalLengthCm","PetalWidthCm"]]
print(new_data)
plt.figure(figsize=(10,7))
new_data.boxplot()
# Q 2 B
import pandas as p
df = pd.read_csv('C:\\Iris.csv')
print(df.head(5))
print(df.tail(5))
print(df.sample(10))
```

---

13) Slip 19 & 28

```
# Q2 A
import pandas as p
df=p.DataFrame(columns =['name','age','per'])
df.loc[0]=['rajesh',20,95]
df.loc[1]=['suresh',21,85]
df.loc[2]=['avinash',20,90]
df.loc[3]=['kunal',21,75]
df.loc[4]=['sakshi',20,80]
df.loc[6]=['xxx',np.nan,95]
df.loc[7]=['suresh',21,85]
df.loc[8]=['archana',22,91]
```

```
df.loc[9]=['kunal',20,np.nan]
print(df)
print(df.shape)
print(df.describe)
print(df.info())
print(df.dtypes)
df["remark"]=None
df
```

---

14) Slip 20

```
# Q 2 A
import matplotlib.pyplot as plt
import numpy as np
x = np.random.randn(50)
y = np.random.randn(50)
plt.plot(x,y)
plt.show()
plt.scatter(x,y)
plt.show()
plt.hist(x)
plt.show()
# Q 2 B
plt.boxplot(y, vert=False)
plt.show()
```

---

15) Slip 21 & 24

```
# Q 2 A
import pandas as p
import matplotlib.pyplot as plt
d=p.read_csv('C:\\Users\\Iris.csv')
d[d.Species=='Iris-setosa'].
plot.bar(x='PetalLengthCm',y=
```

```
'PetalWidthCm',color='orange', label='Setosa')
d[d.Species=='Iris-versicolor'].
plot.bar(x='PetalLengthCm',
y='PetalWidthCm',color='blue',label='versicolor')
d[d.Species=='Iris-virginica']
.plot.bar(x='PetalLengthCm',
y='PetalWidthCm',color='green', label='virginica')
fig.set_xlabel("Petal Length")
fig.set_ylabel("Petal Width")
fig.set_title(" Petal Length VS Petal Width")
plt.show()
# Q 2 B
import pandas as p
import matplotlib.pyplot as plt
d=p.read_csv('C:\\Users\\Iris.csv')
d[d.Species=='Iris-setosa']
.plot.hist(x='PetalLengthCm',
y='PetalWidthCm',color='orange', label='Setosa')
d[d.Species=='Iris-versicolor']
.plot.hist(x='PetalLengthCm',
y='PetalWidthCm',color='blue', label='versicolor')
d[d.Species=='Iris-virginica']
.plot.hist(x='PetalLengthCm',
y='PetalWidthCm',color='green', label='virginica')
fig.set_xlabel("Petal Length")
fig.set_ylabel("Petal Width")
fig.set_title(" Petal Length VS Petal Width")
plt.show()
```

---

16) Slip 25 & 26 & 30

```
# Q 2 A
import matplotlib.pyplot as plt
```

```
import numpy as np
x = np.random.randn(50)
y = np.random.randn(50)
plt.plot(x,y)
plt.show()
plt.scatter(x,y,color='green')
plt.show()
plt.hist(x,color='yellow')
plt.show()
plt.boxplot(y, vert=False)
plt.show()
# Q 2 B
from matplotlib import pyplot as plt
import numpy as np
subjects=['TCS','Data Science','OS','JAVA','PHP','Python']
marks = [23, 17, 35, 29, 12, 33]
fig = plt.figure(figsize =(10, 7))
plt.pie(marks, labels = subjects)
plt.show()
```

---

17) Slip 27

```
import pandas as pd
from sklearn import preprocessing
data = pd.read_csv("C:\\Users\\Data (2).csv")
one_hot_encoded_data = pd.get_dummies
(data, columns=['Country'])
label_encoder = preprocessing.LabelEncoder()
one_hot_encoded_data['Purchased'] =
label_encoder.fit_transform(one_hot_encoded_data
['Purchased'])
print("After applying OneHot and Label Encoding:")
print(one_hot_encoded_data)
```

---

18) Slip 8

```
import pandas as pd
from sklearn.preprocessing import StandardScaler
data = pd.read_csv("C:\\winequality-red (1).csv")
scaler = StandardScaler()
scaled_data = scaler.fit_transform(data)
standardized_df = pd.DataFrame(scaled_data,
columns=data.columns)
print("Standardized Data (mean = 0, std = 1):")
print(standardized_df)
```

---

19) Slip 22

```
import pandas as pd
from sklearn.preprocessing import
MinMaxScaler, StandardScaler, Normalizer
data = pd.read_csv("C:\\winequality-red.csv")
print("Original Data:\n")
print(data.head())
# a. Rescaling
minmax_scaler = MinMaxScaler()
rescaled_data = minmax_scaler.fit_transform(data)
rescaled_df = pd.DataFrame(rescaled_data,
columns=data.columns)
print("\nAfter Rescaling using MinMaxScaler:\n")
print(rescaled_df.head())
# b. Standardizing Data
standard_scaler = StandardScaler()
standardized_data =
standard_scaler.fit_transform(data)
standardized_df = pd.DataFrame(
standardized_data, columns=data.columns)
```

```
print("\nAfter Standardizing (Gaussian
Distribution):\n")
print(standardized_df.head())
# c. Normalizing Data
normalizer = Normalizer()
normalized_data = normalizer.fit_transform(data)
normalized_df = pd.DataFrame(normalized_data,
columns=data.columns)
print("\nAfter Normalizing using Normalizer class:\n")
```

---

20) Slip 23

```
import pandas as pd
from sklearn.preprocessing import
MinMaxScaler, StandardScaler, Binarizer
file_path = "C:\\winequality-red.csv"
df = pd.read_csv(file_path)
print("Original Data:\n", df.head(), "\n")
# Step 2: Rescaling
minmax_scaler = MinMaxScaler()
rescaled_data = minmax_scaler.fit_transform(df)
rescaled_df = pd.DataFrame(rescaled_data,
columns=df.columns)
print("Rescaled Data (MinMaxScaler):\n",
rescaled_df.head(), "\n")
# Step 3: Standardizing Data
standard_scaler = StandardScaler()
standardized_data = standard_scaler.fit_transform(df)
standardized_df = pd.DataFrame(standardized_data,
columns=df.columns)
print("Standardized Data (StandardScaler):\n",
standardized_df.head(), "\n")
# Step 4: Binarizing Data
```

```
binarizer=Binarizer(threshold=0.5)
binarized_data = binarizer.fit_transform(rescaled_df)
binarized_df = pd.DataFrame(binarized_data,
columns=df.columns)
print("Binarized Data (Binarizer):\n",
binarized_df.head())
print(normalized_df.head())
```
