Name:Arpit Aditya

Roll No:21BCE8700

Section:Smart Internez Assign

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

In [ ]:
```
# Creating a random array of 4x4 and converting it into a dataframe
val=np.random.rand(4,4)
val[1,2]=np.nan
df=pd.DataFrame(val,columns=['A','B','C','D'])
display(df)
```

	A	B	C	D
0	0.751554	0.686288	0.956623	0.728394
1	0.003206	0.797169	NaN	0.153174
2	0.995318	0.005456	0.782088	0.882045
3	0.751148	0.968237	0.054956	0.414703

In [ ]:
```
# Renaming the columns of the dataframe
Val="Random_Value"
df.rename(columns={'A':f"{Val}_1",'B':f"{Val}_2",'C':f"{Val}_3",'D':f"{Val}_
display(df)
```

	Random_Value_1	Random_Value_2	Random_Value_3	Random_Value_4
0	0.751554	0.686288	0.956623	0.728394
1	0.003206	0.797169	NaN	0.153174
2	0.995318	0.005456	0.782088	0.882045
3	0.751148	0.968237	0.054956	0.414703

In [ ]:
```
#Descriptive Statistics
display(df.describe())
```

	Random_Value_1	Random_Value_2	Random_Value_3	Random_Value_4
count	4.000000	4.000000	3.000000	4.000000
mean	0.625306	0.614288	0.597889	0.544579
std	0.430384	0.422132	0.478224	0.325442
min	0.003206	0.005456	0.054956	0.153174
25%	0.564162	0.516080	0.418522	0.349321
50%	0.751351	0.741729	0.782088	0.571549
75%	0.812495	0.839936	0.869356	0.766807
max	0.995318	0.968237	0.956623	0.882045

In [ ]:
```
#Checking for null values
display(df.isnull())
display(df.isnull().sum())
#replace null values with 0
# df.fillna(0,inplace=True)
```

	Random_Value_1	Random_Value_2	Random_Value_3	Random_Value_4
0	False	False	False	False
1	False	False	True	False
2	False	False	False	False
3	False	False	False	False

```
Random_Value_1    0
Random_Value_2    0
Random_Value_3    1
Random_Value_4    0
dtype: int64
```

In [ ]:
```
# Finding datatypes of column
display(df.dtypes)
```

```
Random_Value_1    float64
Random_Value_2    float64
Random_Value_3    float64
Random_Value_4    float64
dtype: object
```

In [ ]:
```
# Index Location
display(df.iloc[:,1])
display(df.iloc[:,2])
# Location by column name
display(df.loc[:,'Random_Value_2'])
display(df.loc[:,'Random_Value_3'])
```

```
0    0.686288
1    0.797169
2    0.005456
3    0.968237
Name: Random_Value_2, dtype: float64
0    0.956623
1         NaN
2    0.782088
3    0.054956
Name: Random_Value_3, dtype: float64
0    0.686288
1    0.797169
2    0.005456
3    0.968237
Name: Random_Value_2, dtype: float64
0    0.956623
1         NaN
2    0.782088
3    0.054956
Name: Random_Value_3, dtype: float64
```

Assignment 2

In [ ]:
```
val=np.random.rand(10,5)
val[1,2]=np.nan
df=pd.DataFrame(val)
df=df.rename(columns={0:'A',1:'B',2:'C',3:'D',4:'E'})
print(df)
```

```
          A         B         C         D         E
0  0.360436  0.433614  0.312058  0.383004  0.297754
1  0.028404  0.298206       NaN  0.119028  0.581500
2  0.582183  0.156048  0.242483  0.815540  0.460389
3  0.254617  0.945861  0.354936  0.962532  0.283699
4  0.046822  0.326573  0.265029  0.984936  0.208637
5  0.863215  0.483557  0.731282  0.433527  0.756969
6  0.918555  0.332854  0.327234  0.467258  0.697382
7  0.108692  0.380389  0.896663  0.676804  0.030092
8  0.724264  0.890996  0.020625  0.840556  0.100078
9  0.376189  0.568317  0.476071  0.678712  0.232516
```

In [ ]:
```
#  Check the info of 'df'
display(df.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10 entries, 0 to 9
Data columns (total 5 columns):
 #   Column  Non-Null Count  Dtype  
---  ------  --------------  -----  
 0   A       10 non-null     float64
 1   B       10 non-null     float64
 2   C       9 non-null      float64
 3   D       10 non-null     float64
 4   E       10 non-null     float64
dtypes: float64(5)
memory usage: 532.0 bytes
None
```

In [ ]:
```
# Check the descriptive statistics of 'df'
display(df.describe())
```

	A	B	C	D	E
count	10.000000	10.000000	9.000000	10.000000	10.000000
mean	0.426338	0.481641	0.402931	0.636190	0.364902
std	0.330961	0.255819	0.265553	0.280215	0.248455
min	0.028404	0.156048	0.020625	0.119028	0.030092
25%	0.145173	0.328143	0.265029	0.441960	0.214607
50%	0.368312	0.407001	0.327234	0.677758	0.290726
75%	0.688744	0.547127	0.476071	0.834302	0.551222
max	0.918555	0.945861	0.896663	0.984936	0.756969

In [ ]:
```
# check the 4th index observation with 'loc' slicing operator.
print(df.loc[4])
```

```
A    0.046822
B    0.326573
C    0.265029
D    0.984936
E    0.208637
Name: 4, dtype: float64
```

In [ ]:
```
# Check the null values in your 'df'
print(df.isnull().sum())
```

```
A    0
B    0
C    1
D    0
E    0
dtype: int64
```