GROUP B ASSIGNMENTS

ASSIGNMENTS BASED ON DATA ANALYTICS USING PYTHON

Assignment 1

Problem statement-

Perform the following operations using Python on the Facebook metrics data sets
a) Create data subsets.
b) Merge data
c) sort data
d) transposing data
e) Shape and reshape data

Importing python libraries

In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

Loading a CSV file into a dataframe

In [2]:
```
A = pd.read_csv(r"C:\Users\HP\Downloads\Facebook.csv")
```

In [3]:
```
A.head(5)
```

Out[3]:

	Page total likes	Type	Category	Post Month	Post Weekday	Post Hour	Paid	Lifetime Post Total Reach	Lifetime Post Total Impressions	Lifetime Engaged Users	L Con
0	139441	Photo	2	12	4	3	0.0	2752	5091	178	
1	139441	Status	2	12	3	10	0.0	10460	19057	1457	
2	139441	Photo	3	12	3	3	0.0	2413	4373	177	
3	139441	Photo	2	12	2	10	1.0	50128	87991	2211	
4	139441	Photo	2	12	2	3	0.0	7244	13594	671	

Checking the shape of dataset

In [4]:
```
A.shape
```

Out[4]:
```
(500, 19)
```

A) Creating data subsets

a. Column-wise subset:

In [5]:
```
#subset1
subset1=A.iloc[:,[1,2,18]]
subset1
```

Out[5]:

	Type	Category	Total Interactions
0	Photo	2	100
1	Status	2	164
2	Photo	3	80
3	Photo	2	1777
4	Photo	2	393
...	...	...	...
495	Photo	3	84
496	Photo	2	75
497	Photo	1	115
498	Photo	3	136
499	Photo	2	119

500 rows × 3 columns

In [6]:
```
#subset2
subset2=A.iloc[:, [1,15,16,17]]
subset2
```

Out[6]:

	Type	comment	like	share
0	Photo	4	79.0	17.0
1	Status	5	130.0	29.0
2	Photo	0	66.0	14.0
3	Photo	58	1572.0	147.0
4	Photo	19	325.0	49.0
...	...	...	...	...
495	Photo	5	53.0	26.0
496	Photo	0	53.0	22.0
497	Photo	4	93.0	18.0
498	Photo	7	91.0	38.0
499	Photo	0	91.0	28.0

500 rows × 4 columns

In [7]:
```
subset2.shape
```

Out[7]:
```
(500, 4)
```

b. Row-wise subset:

In [8]:
```
#subset3
subset3=A.iloc[[0,1,2,3,4,5,6,7],[0,1,2]]
```

```
subset3
```

Out[8]:

	Page total likes	Type	Category
0	139441	Photo	2
1	139441	Status	2
2	139441	Photo	3
3	139441	Photo	2
4	139441	Photo	2
5	139441	Status	2
6	139441	Photo	3
7	139441	Photo	3

In [9]:
```
subset3.shape
```

Out[9]:
```
(8, 3)
```

In [10]:
```
#subset4
subset4=A.iloc[[15,18,25,36,45,58,67],[0,1,2]]
subset4
```

Out[10]:

	Page total likes	Type	Category
15	138414	Status	2
18	138414	Status	3
25	138458	Status	2
36	138895	Photo	3
45	138353	Link	1
58	138329	Photo	1
67	138185	Photo	1

In [11]:
```
subset4.shape
```

Out[11]:
```
(7, 3)
```

B) Merging the data

In [12]:
```
m1=pd.concat([subset3,subset4])           #merging row wise subsets
m1
```

Out[12]:

	Page total likes	Type	Category
0	139441	Photo	2
1	139441	Status	2
2	139441	Photo	3
3	139441	Photo	2
4	139441	Photo	2
5	139441	Status	2
6	139441	Photo	3
7	139441	Photo	3
15	138414	Status	2
18	138414	Status	3
25	138458	Status	2
36	138895	Photo	3
45	138353	Link	1
58	138329	Photo	1
67	138185	Photo	1

In [13]:
```
m1.shape
```

Out[13]:
```
(15, 3)
```

In [14]:
```
m2=pd.concat([subset1,subset2])                    #merging column-wise subsets
m2
```

Out[14]:

	Type	Category	Total Interactions	comment	like	share
0	Photo	2.0	100.0	NaN	NaN	NaN
1	Status	2.0	164.0	NaN	NaN	NaN
2	Photo	3.0	80.0	NaN	NaN	NaN
3	Photo	2.0	1777.0	NaN	NaN	NaN
4	Photo	2.0	393.0	NaN	NaN	NaN
...	...	...	...	...	...	...
495	Photo	NaN	NaN	5.0	53.0	26.0
496	Photo	NaN	NaN	0.0	53.0	22.0
497	Photo	NaN	NaN	4.0	93.0	18.0
498	Photo	NaN	NaN	7.0	91.0	38.0
499	Photo	NaN	NaN	0.0	91.0	28.0

1000 rows × 6 columns

C) Sort the data

In [15]:
```
sort_data=m1.sort_values(by=["Category"])
sort_data
```

Out[15]:

	Page total likes	Type	Category
45	138353	Link	1
58	138329	Photo	1
67	138185	Photo	1
0	139441	Photo	2
1	139441	Status	2
3	139441	Photo	2
4	139441	Photo	2
5	139441	Status	2
15	138414	Status	2
25	138458	Status	2
2	139441	Photo	3
6	139441	Photo	3
7	139441	Photo	3
18	138414	Status	3
36	138895	Photo	3

In [16]:
```
sort_data.shape
```

Out[16]:
```
(15, 3)
```

D) Transposing the data

In [17]:
```
trans=sort_data.transpose()
trans
```

Out[17]:

	45	58	67	0	1	3	4	5	15	25	
Page total likes	138353	138329	138185	139441	139441	139441	139441	139441	138414	138458	139
Type	Link	Photo	Photo	Photo	Status	Photo	Photo	Status	Status	Status	Ph
Category	1	1	1	2	2	2	2	2	2	2	

In [18]:
```
trans.shape
```

Out[18]:
```
(3, 15)
```

E) Reshape the data

In [19]:
```
reshape=sort_data.melt(id_vars = ['Type'],value_vars =['Category'])
reshape
```

Out[19]:

	Type	variable	value
0	Link	Category	1
1	Photo	Category	1
2	Photo	Category	1
3	Photo	Category	2
4	Status	Category	2
5	Photo	Category	2
6	Photo	Category	2
7	Status	Category	2
8	Status	Category	2
9	Status	Category	2
10	Photo	Category	3
11	Photo	Category	3
12	Photo	Category	3
13	Status	Category	3
14	Photo	Category	3

In [20]:
```
reshape.shape
```

Out[20]:
```
(15, 3)
```