In [ ]:

```
#Assignment 1 Group B
#Name:Samiksha Bandgar
#Roll No:3307
#Subject: DSBDAL
#Batch:A


#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
```

In [5]:

```
import pandas as pd
df=pd.read_csv(r"C:\Users\Samiksha Bandgar\OneDrive\Desktop\dataset_Facebook.csv",delim
iter=';')
```

In [6]:

```
df.head()
```

Out[6]:

	Page total likes	Type	Category	Post Month	Post Weekday	Post Hour	Paid	Lifetime Post Total Reach	Lifetime Post Total Impressions	Lifetime Engaged Users
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

In [7]:

```
df.shape
```

Out[7]:

```
(500, 19)
```

In [8]:

```
df.isnull()
```

Out[8]:

	Page total likes	Type	Category	Post Month	Post Weekday	Post Hour	Paid	Lifetime Post Total Reach	Lifetime Post Total Impressions	Lifetime Engaged Users
0	False	False	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False	False
...	...	...	...	...	...	...	...	...	...	...
495	False	False	False	False	False	False	False	False	False	False
496	False	False	False	False	False	False	False	False	False	False
497	False	False	False	False	False	False	False	False	False	False
498	False	False	False	False	False	False	False	False	False	False
499	False	False	False	False	False	False	True	False	False	False

500 rows × 19 columns

In [9]:

```
df.columns
```

Out[9]:

```
Index(['Page total likes', 'Type', 'Category', 'Post Month', 'Post Weekda
y',
       'Post Hour', 'Paid', 'Lifetime Post Total Reach',
       'Lifetime Post Total Impressions', 'Lifetime Engaged Users',
       'Lifetime Post Consumers', 'Lifetime Post Consumptions',
       'Lifetime Post Impressions by people who have liked your Page',
       'Lifetime Post reach by people who like your Page',
       'Lifetime People who have liked your Page and engaged with your pos
t',
       'comment', 'like', 'share', 'Total Interactions'],
      dtype='object')
```

In [10]:

```
#Creating Subsets
Subset1=df[['Type','Category','Post Month']]
```

In [11]:

```
Subset1
```

Out[11]:

	Type	Category	Post Month
0	Photo	2	12
1	Status	2	12
2	Photo	3	12
3	Photo	2	12
4	Photo	2	12
...	...	...	...
495	Photo	3	1
496	Photo	2	1
497	Photo	1	1
498	Photo	3	1
499	Photo	2	1

500 rows × 3 columns

In [12]:

```
Subset1.shape
```

Out[12]:

```
(500, 3)
```

In [15]:

```
Subset2=df.loc[0:200]
Subset2.shape
```

Out[15]:

```
(201, 19)
```

In [16]:

```
Subset3=df.loc[201:500]
Subset3.shape
```

Out[16]:

```
(299, 19)
```

In [17]:

```
Subset4=df.loc[0:2,['Type','Post Month']]
Subset4
```

Out[17]:

	Type	Post Month
0	Photo	12
1	Status	12
2	Photo	12

In [18]:

```
Subset4.shape
```

Out[18]:

```
(3, 2)
```

In [19]:

```
#Merge the subsets by row
S1=Subset2
S1.shape
```

Out[19]:

```
(201, 19)
```

In [20]:

```
S2=Subset3
S2.shape
```

Out[20]:

```
(299, 19)
```

In [21]:

```
mergedf=pd.concat([Subset2,Subset3])
mergedf.shape
```

Out[21]:

```
(500, 19)
```

In [22]:

```
#Merge the subsets by columns
m2=S1.merge(S2,on='Type')
```

In [23]:

```
m2.shape
```

Out[23]:

```
(43069, 37)
```

In [24]:

```
sb1=df[["Post Month","Type"]];
sb1.shape
```

Out[24]:

```
(500, 2)
```

In [25]:

```
sb2=df[["Category","Type"]];
sb2.shape
```

Out[25]:

```
(500, 2)
```

In [26]:

```
m3=sb1.merge(sb2,left_index=True,right_index=True)
m3.shape
```

Out[26]:

```
(500, 4)
```

In [27]:

```
#Sort the subset
st=df.sort_values(by='Type')
st
```

Out[27]:

	Page total likes	Type	Category	Post Month	Post Weekday	Post Hour	Paid	Lifetime Post Total Reach	Lifetime Post Total Impressions	Lifetime Engaged Users
438	98195	Link	2	3	6	6	0.0	5730	10083	103
470	91437	Link	1	2	3	13	0.0	9356	14986	448
41	138895	Link	1	12	6	3	1.0	18480	28438	517
43	138353	Link	1	12	5	3	1.0	2645	4270	134
45	138353	Link	1	12	4	3	1.0	7968	13023	206
...	...	...	...	...	...	...	...	...	...	...
71	137893	Video	1	11	5	3	1.0	100768	220447	2101
55	138329	Video	1	11	6	2	1.0	16416	31950	459
277	126424	Video	1	6	2	13	0.0	139008	277100	1779
243	130791	Video	1	7	3	11	1.0	21872	40413	3872
74	137893	Video	1	11	3	11	0.0	13544	30235	517

500 rows × 19 columns

In [28]:

```
#sort the dataframe based on multiple columns
st1=df.sort_values(by=['Type','Post Month'])
st1
```

Out[28]:

	Page total likes	Type	Category	Post Month	Post Weekday	Post Hour	Paid	Lifetime Post Total Reach	Lifetime Post Total Impressions	Lifetime Engaged Users
477	86909	Link	1	1	6	4	0.0	39600	7927	572
481	86491	Link	1	1	4	4	1.0	4938	7910	66
485	86491	Link	1	1	2	2	0.0	5168	8371	66
492	85979	Link	1	1	5	11	0.0	45920	5808	753
470	91437	Link	1	2	3	13	0.0	9356	14986	448
...	...	...	...	...	...	...	...	...	...	...
183	134879	Video	1	9	2	10	0.0	30624	56950	2080
55	138329	Video	1	11	6	2	1.0	16416	31950	459
71	137893	Video	1	11	5	3	1.0	100768	220447	2101
74	137893	Video	1	11	3	11	0.0	13544	30235	517
29	138895	Video	1	12	4	11	1.0	36208	61262	1141

500 rows × 19 columns

In [29]:

```
st2=df.sort_values(by='Post Weekday',ascending=False)
st2
```

Out[29]:

	Page total likes	Type	Category	Post Month	Post Weekday	Post Hour	Paid	Lifetime Post Total Reach	Lifetime Post Total Impressions	Lifetime Engaged Users
250	129600	Photo	1	7	7	6	1.0	5848	9068	622
428	100732	Photo	1	3	7	15	0.0	5132	9067	398
430	100732	Link	1	3	7	14	0.0	2933	5144	24
431	100732	Photo	1	3	7	12	0.0	4094	7469	206
380	111620	Photo	1	4	7	14	0.0	128064	251269	1539
...	...	...	...	...	...	...	...	...	...	...
220	131956	Photo	2	8	1	4	0.0	2540	4372	389
219	131956	Photo	3	8	1	12	0.0	5746	9874	769
106	137020	Photo	3	10	1	11	0.0	3674	7221	452
107	136736	Status	2	10	1	4	0.0	9504	19556	1132
262	128032	Photo	2	7	1	3	0.0	3330	5461	513

500 rows × 19 columns

In [30]:

```
#Transpose the dataset
tp=df.transpose()
```

In [31]:

```
tp.shape
```

Out[31]:

```
(19, 500)
```

In [32]:

```
#Reshape the dataset
rs=pd.melt(df,id_vars=['Type'],value_vars=['Post Month'],var_name=['Post Month'])
```

In [33]:

```
rs
```

Out[33]:

	Type	Post Month	value
0	Photo	Post Month	12
1	Status	Post Month	12
2	Photo	Post Month	12
3	Photo	Post Month	12
4	Photo	Post Month	12
...	...	...	...
495	Photo	Post Month	1
496	Photo	Post Month	1
497	Photo	Post Month	1
498	Photo	Post Month	1
499	Photo	Post Month	1

500 rows × 3 columns

In [34]:

```
rs1=pd.melt(df,id_vars=['Type'],value_vars=['Post Month','Category'])
```

In [35]:

```
rs1
```

Out[35]:

	Type	variable	value
0	Photo	Post Month	12
1	Status	Post Month	12
2	Photo	Post Month	12
3	Photo	Post Month	12
4	Photo	Post Month	12
...	...	...	...
995	Photo	Category	3
996	Photo	Category	2
997	Photo	Category	1
998	Photo	Category	3
999	Photo	Category	2

1000 rows × 3 columns

In [ ]:

```
```