Group B Assignment - 1

1. 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 libraries

In [24]:
```
import pandas as pd
import numpy as np
```

Reading csv file

In [2]:
```
data = pd.read_csv("dataset_Facebook (1).csv")
```

In [3]:
```
data
```

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	C
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	
...	...	...	...	...	...	...	...	...	...	...	
495	85093	Photo	3	1	7	2	0.0	4684	7536	733	
496	81370	Photo	2	1	5	8	0.0	3480	6229	537	
497	81370	Photo	1	1	5	2	0.0	3778	7216	625	
498	81370	Photo	3	1	4	11	0.0	4156	7564	626	
499	81370	Photo	2	1	4	4	NaN	4188	7292	564	

500 rows × 19 columns

Shape , Colums , Head and Tail of data

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

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

In [5]:
```
data.columns
```

Out[5]:
```
Index(['Page total likes', 'Type', 'Category', 'Post Month', 'Post Weekday',
       '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 post',
       'comment', 'like', 'share', 'Total Interactions'],
      dtype='object')
```

In [6]:
```
data.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	Cor
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]:
```
data.tail()
```

Out[7]:

	Page total likes	Type	Category	Post Month	Post Weekday	Post Hour	Paid	Lifetime Post Total Reach	Lifetime Post Total Impressions	Lifetime Engaged Users	Co
495	85093	Photo	3	1	7	2	0.0	4684	7536	733	
496	81370	Photo	2	1	5	8	0.0	3480	6229	537	
497	81370	Photo	1	1	5	2	0.0	3778	7216	625	
498	81370	Photo	3	1	4	11	0.0	4156	7564	626	
499	81370	Photo	2	1	4	4	NaN	4188	7292	564	

Removing null valued rows

In [8]:
```
data = data.dropna()
data
```

Out[8]:

Category	Post Month	Post Weekday	Post Hour	Paid	Lifetime Post Total Reach	Lifetime Post Total Impressions	Lifetime Engaged Users	Lifetime Post Consumers	Lifetime Pc Consumptio
2	12	4	3	0.0	2752	5091	178	109	1
2	12	3	10	0.0	10460	19057	1457	1361	16
3	12	3	3	0.0	2413	4373	177	113	1
2	12	2	10	1.0	50128	87991	2211	790	11
2	12	2	3	0.0	7244	13594	671	410	5
...	...	...	...	...	...	...	...	...	
3	1	7	10	0.0	5400	9218	810	756	10
3	1	7	2	0.0	4684	7536	733	708	9
2	1	5	8	0.0	3480	6229	537	508	6
1	1	5	2	0.0	3778	7216	625	572	7
3	1	4	11	0.0	4156	7564	626	574	8

s

Creating the first column subset

In [27]:
```
column_data1=data[["Type","Category","Post Month","Post Weekday","Post Hour"]]
column_data1
```

Out[27]:

	Type	Category	Post Month	Post Weekday	Post Hour
0	Photo	2	12	4	3
1	Status	2	12	3	10
2	Photo	3	12	3	3
3	Photo	2	12	2	10
4	Photo	2	12	2	3
...	...	...	...	...	...
494	Photo	3	1	7	10
495	Photo	3	1	7	2
496	Photo	2	1	5	8
497	Photo	1	1	5	2
498	Photo	3	1	4	11

495 rows × 5 columns

Creating the second column subset

In [28]:
```
column_data2=data[["Paid","Lifetime Post Total Reach","Lifetime Post Consumers
column_data2
```

Out[28]:

	Paid	Lifetime Post Total Reach	Lifetime Post Consumers
0	0.0	2752	109
1	0.0	10460	1361
2	0.0	2413	113
3	1.0	50128	790
4	0.0	7244	410
...	...	...	...
494	0.0	5400	756
495	0.0	4684	708
496	0.0	3480	508
497	0.0	3778	572
498	0.0	4156	574

495 rows × 3 columns

Creating the first row subset

In [29]:
```
row_data1=data.iloc[[1,2,3,4,5,6,8,9,10,11]]
row_data1
```

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	Co
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	
5	139441	Status	2	12	1	9	0.0	10472	20849	1191	
6	139441	Photo	3	12	1	3	1.0	11692	19479	481	
8	139441	Status	2	12	7	3	0.0	11844	22538	1530	
9	139441	Photo	3	12	6	10	0.0	4694	8668	280	
10	139441	Status	2	12	5	10	0.0	21744	42334	4258	
11	139441	Photo	2	12	5	10	0.0	3112	5590	208	

Creating the second row subset

In [36]:
```
row_data2=data.iloc[[1,2,3,7,11,13,14,15,16,17]]
row_data2
```

Out[36]:

	Page total likes	Type	Category	Post Month	Post Weekday	Post Hour	Paid	Lifetime Post Total Reach	Lifetime Post Total Impressions	Lifetime Engaged Users	Co
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	
7	139441	Photo	3	12	7	9	1.0	13720	24137	537	
11	139441	Photo	2	12	5	10	0.0	3112	5590	208	
13	139441	Photo	2	12	5	3	0.0	2549	4896	249	
14	138414	Photo	2	12	4	5	1.0	22784	39941	887	
15	138414	Status	2	12	3	10	0.0	10060	19680	1264	
16	138414	Photo	3	12	3	3	0.0	1722	2981	163	
17	138414	Photo	1	12	2	12	1.0	53264	111785	1706	

Merging two subsets

In [38]:
```
m1 = row_data1.merge(row_data2)
m1
```

Out[38]:

	Page total likes	Type	Category	Post Month	Post Weekday	Post Hour	Paid	Lifetime Post Total Reach	Lifetime Post Total Impressions	Lifetime Engaged Users	Cor
0	139441	Status	2	12	3	10	0.0	10460	19057	1457	
1	139441	Photo	3	12	3	3	0.0	2413	4373	177	
2	139441	Photo	2	12	2	10	1.0	50128	87991	2211	
3	139441	Photo	2	12	5	10	0.0	3112	5590	208	

Concatenating two subsets

In [40]:
```
c1 = pd.concat([row_data1,row_data2])
c1
```

Out[40]:

	Page total likes	Type	Category	Post Month	Post Weekday	Post Hour	Paid	Lifetime Post Total Reach	Lifetime Post Total Impressions	Lifetime Engaged Users	Co
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	
5	139441	Status	2	12	1	9	0.0	10472	20849	1191	
6	139441	Photo	3	12	1	3	1.0	11692	19479	481	
8	139441	Status	2	12	7	3	0.0	11844	22538	1530	
9	139441	Photo	3	12	6	10	0.0	4694	8668	280	
10	139441	Status	2	12	5	10	0.0	21744	42334	4258	
11	139441	Photo	2	12	5	10	0.0	3112	5590	208	
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	
7	139441	Photo	3	12	7	9	1.0	13720	24137	537	
11	139441	Photo	2	12	5	10	0.0	3112	5590	208	
13	139441	Photo	2	12	5	3	0.0	2549	4896	249	
14	138414	Photo	2	12	4	5	1.0	22784	39941	887	
15	138414	Status	2	12	3	10	0.0	10060	19680	1264	
16	138414	Photo	3	12	3	3	0.0	1722	2981	163	
17	138414	Photo	1	12	2	12	1.0	53264	111785	1706	

Sorting the subset

In [41]:
```
sort1= c1.sort_values(by="Post Weekday")
sort1
```

Out[41]:

	Page total likes	Type	Category	Post Month	Post Weekday	Post Hour	Paid	Lifetime Post Total Reach	Lifetime Post Total Impressions	Lifetime Engaged Users	Co
5	139441	Status	2	12	1	9	0.0	10472	20849	1191	
6	139441	Photo	3	12	1	3	1.0	11692	19479	481	
17	138414	Photo	1	12	2	12	1.0	53264	111785	1706	
3	139441	Photo	2	12	2	10	1.0	50128	87991	2211	
4	139441	Photo	2	12	2	3	0.0	7244	13594	671	
3	139441	Photo	2	12	2	10	1.0	50128	87991	2211	
15	138414	Status	2	12	3	10	0.0	10060	19680	1264	
2	139441	Photo	3	12	3	3	0.0	2413	4373	177	
1	139441	Status	2	12	3	10	0.0	10460	19057	1457	
1	139441	Status	2	12	3	10	0.0	10460	19057	1457	
2	139441	Photo	3	12	3	3	0.0	2413	4373	177	
16	138414	Photo	3	12	3	3	0.0	1722	2981	163	
14	138414	Photo	2	12	4	5	1.0	22784	39941	887	
10	139441	Status	2	12	5	10	0.0	21744	42334	4258	
11	139441	Photo	2	12	5	10	0.0	3112	5590	208	
13	139441	Photo	2	12	5	3	0.0	2549	4896	249	
11	139441	Photo	2	12	5	10	0.0	3112	5590	208	
9	139441	Photo	3	12	6	10	0.0	4694	8668	280	
8	139441	Status	2	12	7	3	0.0	11844	22538	1530	
7	139441	Photo	3	12	7	9	1.0	13720	24137	537	

Transposing a subset

In [44]:
```
transpose=sort1.transpose()
transpose
```

Out[44]:

	5	6	17	3	4	3	15	2	1	
Page total likes	139441	139441	138414	139441	139441	139441	138414	139441	139441	1394
Type	Status	Photo	Photo	Photo	Photo	Photo	Status	Photo	Status	Stat
Category	2	3	1	2	2	2	2	3	2	
Post Month	12	12	12	12	12	12	12	12	12	
Post Weekday	1	1	2	2	2	2	3	3	3	
Post Hour	9	3	12	10	3	10	10	3	10	
Paid	0.0	1.0	1.0	1.0	0.0	1.0	0.0	0.0	0.0	0
Lifetime Post Total Reach	10472	11692	53264	50128	7244	50128	10060	2413	10460	104
Lifetime Post Total Impressions	20849	19479	111785	87991	13594	87991	19680	4373	19057	190
Lifetime Engaged Users	1191	481	1706	2211	671	2211	1264	177	1457	14
Lifetime Post Consumers	1073	265	1103	790	410	790	1209	113	1361	13
Lifetime Post Consumptions	1389	364	1655	1119	580	1119	1425	154	1674	16
Lifetime Post Impressions by people who have liked your Page	16034	15432	92512	61027	6228	61027	17272	2812	11710	117
Lifetime Post reach by people who like your Page	7852	9328	39776	32048	3200	32048	8548	1503	6112	61
Lifetime People who have liked your Page and engaged with your post	1016	379	1307	1386	396	1386	1162	132	1108	11
comment	1	3	15	58	19	58	4	0	5	
like	152.0	249.0	678.0	1572.0	325.0	1572.0	86.0	66.0	130.0	130
share	33.0	27.0	20.0	147.0	49.0	147.0	18.0	14.0	29.0	29
Total Interactions	186	279	713	1777	393	1777	108	80	164	1

In [45]:
```
print(sort1.shape)
print(transpose.shape)
```

```
(20, 19)
(19, 20)
```

Reshaping the subset

In [48]:
```
reshape=sort1.melt(id_vars="Type",value_vars=["Lifetime Post Total Reach","Pos
reshape
```

Out[48]:

	Type	variable	value
0	Status	Lifetime Post Total Reach	10472
1	Photo	Lifetime Post Total Reach	11692
2	Photo	Lifetime Post Total Reach	53264
3	Photo	Lifetime Post Total Reach	50128
4	Photo	Lifetime Post Total Reach	7244
5	Photo	Lifetime Post Total Reach	50128
6	Status	Lifetime Post Total Reach	10060
7	Photo	Lifetime Post Total Reach	2413
8	Status	Lifetime Post Total Reach	10460
9	Status	Lifetime Post Total Reach	10460
10	Photo	Lifetime Post Total Reach	2413
11	Photo	Lifetime Post Total Reach	1722
12	Photo	Lifetime Post Total Reach	22784
13	Status	Lifetime Post Total Reach	21744
14	Photo	Lifetime Post Total Reach	3112
15	Photo	Lifetime Post Total Reach	2549
16	Photo	Lifetime Post Total Reach	3112
17	Photo	Lifetime Post Total Reach	4694
18	Status	Lifetime Post Total Reach	11844
19	Photo	Lifetime Post Total Reach	13720
20	Status	Post Weekday	1
21	Photo	Post Weekday	1
22	Photo	Post Weekday	2
23	Photo	Post Weekday	2
24	Photo	Post Weekday	2
25	Photo	Post Weekday	2
26	Status	Post Weekday	3
27	Photo	Post Weekday	3
28	Status	Post Weekday	3
29	Status	Post Weekday	3
30	Photo	Post Weekday	3
31	Photo	Post Weekday	3
32	Photo	Post Weekday	4
33	Status	Post Weekday	5
34	Photo	Post Weekday	5
35	Photo	Post Weekday	5
36	Photo	Post Weekday	5
37	Photo	Post Weekday	6
38	Status	Post Weekday	7

	Type	variable	value
39	Photo	Post Weekday	7