In [ ]:
```
//Akanksha Indalkar TE-IT-B 37024
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 [1]:
```
import os
os.chdir('desktop')
```

In [2]:
```
import os
os.getcwd()
```

Out[2]:
```
'C:\\Users\\AKANKSHA\\desktop'
```

In [3]:
```
import pandas as pd
import numpy as np
```

In [25]:
```
df = pd.read_csv(r"dataset_Facebook.csv",sep=';')
```

In [26]:
```
df
```

Out[26]:

	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
0	139441	Photo	2	12	4	3	0.0	2752	5091	178	109	159	3078	1640	119
1	139441	Status	2	12	3	10	0.0	10460	19057	1457	1361	1674	11710	6112	1108
2	139441	Photo	3	12	3	3	0.0	2413	4373	177	113	154	2812	1503	132

3	139441	Photo	2	12	2	10	1.0	50128	87991	2211	790	1119	61027	32048	1386
4	139441	Photo	2	12	2	3	0.0	7244	13594	671	410	580	6228	3200	396
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
495	85093	Photo	3	1	7	2	0.0	4684	7536	733	708	985	4750	2876	392
496	81370	Photo	2	1	5	8	0.0	3480	6229	537	508	687	3961	2104	301
497	81370	Photo	1	1	5	2	0.0	3778	7216	625	572	795	4742	2388	363
498	81370	Photo	3	1	4	11	0.0	4156	7564	626	574	832	4534	2452	370
499	81370	Photo	2	1	4	4	NaN	4188	7292	564	524	743	3861	2200	316

500 rows × 19 columns

In [27]:
```
#describe
df.describe()
```

Out[27]:

	Page total likes	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
count	500.000000	500.000000	500.000000	500.000000	500.000000	499.000000	500.00000	5.000000e+02	500.000000	500.000000	500.000000
mean	123194.176000	1.880000	7.038000	4.150000	7.840000	0.278557	13903.36000	2.958595e+04	920.344000	798.772000	1415.130000
std	16272.813214	0.852675	3.307936	2.030701	4.368589	0.448739	22740.78789	7.680325e+04	985.016636	882.505013	2000.594118
min	81370.000000	1.000000	1.000000	1.000000	1.000000	0.000000	238.00000	5.700000e+02	9.000000	9.000000	9.000000
25%	112676.000000	1.000000	4.000000	2.000000	3.000000	0.000000	3315.00000	5.694750e+03	393.750000	332.500000	509.250000
50%	129600.000000	2.000000	7.000000	4.000000	9.000000	0.000000	5281.00000	9.051000e+03	625.500000	551.500000	851.000000
75%	136393.000000	3.000000	10.000000	6.000000	11.000000	1.000000	13168.00000	2.208550e+04	1062.000000	955.500000	1463.000000
max	139441.000000	3.000000	12.000000	7.000000	23.000000	1.000000	180480.00000	1.110282e+06	11452.000000	11328.000000	19779.000000

In [28]:
```
df.shape
```

Out[28]:
```
(500, 19)
```

In [31]:
```
#creating subsets
#subset-1
df1=df[['Page total likes','Category','Post Month','Post Weekday']].iloc[0:15]
df1
```

Out[31]:

	Page total likes	Category	Post Month	Post Weekday
0	139441	2	12	4
1	139441	2	12	3
2	139441	3	12	3
3	139441	2	12	2
4	139441	2	12	2
5	139441	2	12	1
6	139441	3	12	1
7	139441	3	12	7
8	139441	2	12	7
9	139441	3	12	6
10	139441	2	12	5
11	139441	2	12	5
12	139441	2	12	5
13	139441	2	12	5
14	138414	2	12	4

In [32]:
```
#subset-2
df2=df[['Page total likes','Category','Post Month','Post Weekday']].iloc[16:30]
```

```
df2
```

Out[32]:

	Page total likes	Category	Post Month	Post Weekday
16	138414	3	12	3
17	138414	1	12	2
18	138414	3	12	2
19	138414	3	12	1
20	138414	2	12	1
21	138414	1	12	7
22	138414	1	12	7
23	138414	3	12	7
24	138414	2	12	6
25	138458	2	12	6
26	138458	2	12	5
27	138458	3	12	5
28	138895	2	12	5
29	138895	1	12	4

In [33]:

```
#subset-3
df3=df[['Page total likes','Category','Post Month','Post Weekday']].iloc[31:50]
df3
```

Out[33]:

	Page total likes	Category	Post Month	Post Weekday
31	138895	2	12	3
32	138895	3	12	3
33	138895	3	12	2
34	138895	1	12	2
35	138895	2	12	1

	Page total likes	Category	Post Month	Post Weekday
36	138895	3	12	1
37	138895	1	12	7
38	138895	2	12	7
39	138895	1	12	7
40	138895	2	12	6
41	138895	1	12	6
42	138353	1	12	5
43	138353	1	12	5
44	138353	1	12	4
45	138353	1	12	4
46	138353	1	12	3
47	138353	1	12	3
48	138353	1	12	2
49	138353	1	12	2

In [34]:
```
#merging subsets
merging=pd.concat([df1,df2,df3])
merging
```

Out[34]:

	Page total likes	Category	Post Month	Post Weekday
0	139441	2	12	4
1	139441	2	12	3
2	139441	3	12	3
3	139441	2	12	2
4	139441	2	12	2
5	139441	2	12	1

	Page total likes	Category	Post Month	Post Weekday
6	139441	3	12	1
7	139441	3	12	7
8	139441	2	12	7
9	139441	3	12	6
10	139441	2	12	5
11	139441	2	12	5
12	139441	2	12	5
13	139441	2	12	5
14	138414	2	12	4
16	138414	3	12	3
17	138414	1	12	2
18	138414	3	12	2
19	138414	3	12	1
20	138414	2	12	1
21	138414	1	12	7
22	138414	1	12	7
23	138414	3	12	7
24	138414	2	12	6
25	138458	2	12	6
26	138458	2	12	5
27	138458	3	12	5
28	138895	2	12	5
29	138895	1	12	4
31	138895	2	12	3

	Page total likes	Category	Post Month	Post Weekday
32	138895	3	12	3
33	138895	3	12	2
34	138895	1	12	2
35	138895	2	12	1
36	138895	3	12	1
37	138895	1	12	7
38	138895	2	12	7
39	138895	1	12	7
40	138895	2	12	6
41	138895	1	12	6
42	138353	1	12	5
43	138353	1	12	5
44	138353	1	12	4
45	138353	1	12	4
46	138353	1	12	3
47	138353	1	12	3
48	138353	1	12	2
49	138353	1	12	2

In [40]:
```
#Sort data
sort_values = df.sort_values('Page total likes',ascending=False)
sort_values
```

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	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
0	139441	Photo	2	12	4	3	0.0	2752	5091	178	109	159	3078	1640	119
8	139441	Status	2	12	7	3	0.0	11844	22538	1530	1407	1692	15220	7912	1250
1	139441	Status	2	12	3	10	0.0	10460	19057	1457	1361	1674	11710	6112	1108
12	139441	Photo	2	12	5	10	0.0	2847	5133	193	115	133	3779	2072	152
11	139441	Photo	2	12	5	10	0.0	3112	5590	208	127	145	3887	2174	165
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
495	85093	Photo	3	1	7	2	0.0	4684	7536	733	708	985	4750	2876	392
496	81370	Photo	2	1	5	8	0.0	3480	6229	537	508	687	3961	2104	301
497	81370	Photo	1	1	5	2	0.0	3778	7216	625	572	795	4742	2388	363
498	81370	Photo	3	1	4	11	0.0	4156	7564	626	574	832	4534	2452	370
499	81370	Photo	2	1	4	4	NaN	4188	7292	564	524	743	3861	2200	316

500 rows × 19 columns

In [41]:
```
#Transposinng data
df.transpose()
```

Out[41]:

	0	1	2	3	4	5	6	7	8	9	...	490	491	492	493	494	495	496
Page total likes	139441	139441	139441	139441	139441	139441	139441	139441	139441	139441	...	85979	85979	85979	85093	85093	85093	81370

	0	1	2	3	4	5	6	7	8	9	...	490	491	492	493	494	495	496
Type	Photo	Status	Photo	Photo	Photo	Status	Photo	Photo	Status	Photo	...	Photo	Photo	Link	Photo	Photo	Photo	Photo
Category	2	2	3	2	2	2	3	3	2	3	...	3	3	1	3	3	3	2
Post Month	12	12	12	12	12	12	12	12	12	12	...	1	1	1	1	1	1	1
Post Weekday	4	3	3	2	2	1	1	7	7	6	...	6	6	5	1	7	7	5
Post Hour	3	10	3	10	3	9	3	9	3	10	...	11	3	11	2	10	2	8
Paid	0	0	0	1	0	0	1	1	0	0	...	0	1	0	0	0	0	0
Lifetime Post Total Reach	2752	10460	2413	50128	7244	10472	11692	13720	11844	4694	...	5280	6184	45920	8412	5400	4684	3480
Lifetime Post Total Impressions	5091	19057	4373	87991	13594	20849	19479	24137	22538	8668	...	8703	10228	5808	13960	9218	7536	6229
Lifetime Engaged Users	178	1457	177	2211	671	1191	481	537	1530	280	...	951	956	753	1179	810	733	537
Lifetime Post Consumers	109	1361	113	790	410	1073	265	232	1407	183	...	911	901	655	1111	756	708	508
Lifetime Post Consumptions	159	1674	154	1119	580	1389	364	305	1692	250	...	1237	1140	763	1632	1003	985	687
Lifetime Post Impressions by people who have liked your Page	3078	11710	2812	61027	6228	16034	15432	19728	15220	4309	...	5757	6085	15766	8632	5654	4750	3961
Lifetime Post reach by people who like your Page	1640	6112	1503	32048	3200	7852	9328	11056	7912	2324	...	3300	3502	10720	5348	3230	2876	2104
Lifetime People who have liked your Page and	119	1108	132	1386	396	1016	379	422	1250	199	...	431	437	220	699	422	392	301

	0	1	2	3	4	5	6	7	8	9	...	490	491	492	493	494	495	496
engaged with your post																		
comment	4	5	0	58	19	1	3	0	0	3	...	1	1	0	17	10	5	0
like	79	130	66	1572	325	152	249	325	161	113	...	79	105	128	185	125	53	53
share	17	29	14	147	49	33	27	14	31	26	...	30	46	9	55	41	26	22
Total Interactions	100	164	80	1777	393	186	279	339	192	142	...	110	152	137	257	176	84	75

19 rows × 500 columns

In [42]:
```
#shaping
df.shape
```

Out[42]:
```
(500, 19)
```

In [45]:
```
#reshaping
pivot_table=pd.pivot_table(df,index=['Type','Category'],values='comment')
pivot_table
```

Out[45]:

		comment
Type	**Category**	
Link	**1**	2.900000
	2	2.000000
	3	2.000000
Photo	**1**	5.897297
	2	11.692308
	3	6.913333
Status	**1**	4.333333
	2	9.921053

		comment
Type	**Category**	
	3	2.750000
Video	**1**	12.285714

In [46]:
```
#reshaping into array
reshaping_arr=np.array([1,2,3,4,5,6,7,8,9,10])
reshaping_arr.reshape(5,2)
```

Out[46]:
```
array([[ 1,  2],
       [ 3,  4],
       [ 5,  6],
       [ 7,  8],
       [ 9, 10]])
```

In [ ]: