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 libraries and reading dataset

In [27]:
```
import pandas as pd
import numpy as np
```

In [28]:
```
data=pd.read_csv("Facebook.csv")
data
```

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	Cc
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

a) Shape of Data

In [29]:
```
data.shape
```

Out[29]: (500, 19)

b) Displaying first five rows of file

In [30]:
```
data.head()
```

Out[30]:

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

c) Displaying last five rows of file

In [31]:
```
data.tail()
```

Out[31]:

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

d) Missing Values

In [32]:
```
data.isnull()
```

Out[32]:

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

e) Columns in dataset

In [33]:
```
data.columns
```

Out[33]:
```
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')
```

f) Creating subset using iloc

In [34]:
```
subset1=data.iloc[:,:8]
subset1
```

Out[34]:

	Page total likes	Type	Category	Post Month	Post Weekday	Post Hour	Paid	Lifetime Post Total Reach
0	139441	Photo	2	12	4	3	0.0	2752
1	139441	Status	2	12	3	10	0.0	10460
2	139441	Photo	3	12	3	3	0.0	2413
3	139441	Photo	2	12	2	10	1.0	50128
4	139441	Photo	2	12	2	3	0.0	7244
...	...	...	...	...	...	...	...	...
495	85093	Photo	3	1	7	2	0.0	4684
496	81370	Photo	2	1	5	8	0.0	3480
497	81370	Photo	1	1	5	2	0.0	3778
498	81370	Photo	3	1	4	11	0.0	4156
499	81370	Photo	2	1	4	4	NaN	4188

500 rows × 8 columns

g) Shape of subset

In [35]:
```
subset1.shape
```

Out[35]:
```
(500, 8)
```

h) Creating a subset Using Indexing operator

In [36]:
```
subset2=data[["Post Month","Post Hour","Paid"]]
subset2
```

Out[36]:

	Post Month	Post Hour	Paid
0	12	3	0.0
1	12	10	0.0
2	12	3	0.0
3	12	10	1.0
4	12	3	0.0
...	...	...	...
495	1	2	0.0
496	1	8	0.0
497	1	2	0.0
498	1	11	0.0
499	1	4	NaN

500 rows × 3 columns

i) Shape of subset

In [37]:
```
subset2.shape
```

Out[37]:
```
(500, 3)
```

j) Subset containing all rows

In [38]:
```
subset3=data.iloc[:,7:14]
subset3
```

Out[38]:

	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
0	2752	5091	178	109	159	3078	1640
1	10460	19057	1457	1361	1674	11710	6112
2	2413	4373	177	113	154	2812	1503
3	50128	87991	2211	790	1119	61027	32048
4	7244	13594	671	410	580	6228	3200
...	...	...	...	...	...	...	...
495	4684	7536	733	708	985	4750	2876
496	3480	6229	537	508	687	3961	2104
497	3778	7216	625	572	795	4742	2388
498	4156	7564	626	574	832	4534	2452
499	4188	7292	564	524	743	3861	2200

500 rows × 7 columns

k) Shape of subset

In [39]:
```
subset3.shape
```

Out[39]:
```
(500, 7)
```

l) Creating subset using loc

In [40]:
```
subset4=data.loc[0:199]
subset4
```

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	Cc
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	
...	...	...	...	...	...	...	...	...	...	...	
195	133594	Photo	2	8	6	10	0.0	5282	8730	703	
196	133594	Photo	1	8	6	8	0.0	1809	3130	399	
197	133594	Photo	2	8	5	13	0.0	1920	3124	365	
198	133451	Photo	1	8	4	9	1.0	1954	3530	356	
199	132817	Photo	3	8	4	10	0.0	33536	64850	1954	

200 rows × 19 columns

m) Shape of subset

In [41]:
```
subset4.shape
```

Out[41]: (200, 19)

n) Subset containing half no. of rows

In [42]:
```
subset5=data.loc[200:]
subset5
```

Out[42]:

	Page total likes	Type	Category	Post Month	Post Weekday	Post Hour	Paid	Lifetime Post Total Reach	Lifetime Post Total Impressions	Lifetime Engaged Users	Cc
200	132817	Photo	2	8	4	3	1.0	4204	7191	498	
201	132817	Photo	1	8	3	9	0.0	3376	6557	428	
202	132817	Status	2	8	3	2	1.0	9236	16054	1151	
203	132817	Photo	3	8	2	10	0.0	72864	205934	946	
204	132817	Photo	3	8	2	3	0.0	3358	5682	394	
...	...	...	...	...	...	...	...	...	...	...	
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	

300 rows × 19 columns

p) Shape of subset

In [43]:
```
subset5.shape
```

Out[43]: (300, 19)

q) Merge subset using merge function

In [44]:
```
merge_subset_c=subset1.merge(subset3)
merge_subset_c
```

Out[44]:

	Page total likes	Type	Category	Post Month	Post Weekday	Post Hour	Paid	Lifetime Post Total Reach	Lifetime Post Total Impressions	Lifetime Engaged Users	Cc
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	
...	...	...	...	...	...	...	...	...	...	...	
525	85093	Photo	3	1	7	2	0.0	4684	7536	733	
526	81370	Photo	2	1	5	8	0.0	3480	6229	537	
527	81370	Photo	1	1	5	2	0.0	3778	7216	625	
528	81370	Photo	3	1	4	11	0.0	4156	7564	626	
529	81370	Photo	2	1	4	4	NaN	4188	7292	564	

530 rows × 14 columns

r) Shape of merge subset

In [45]:
```
merge_subset_c.shape
```

Out[45]: (530, 14)

s) Merge subset using concat function

In [46]:
```
merge_subset_r=pd.concat([subset4,subset5])
merge_subset_r
```

Out[46]:

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

t) Transpose of dataset

In [47]:
```
transposeData=data.transpose()
transposeData
```

Out[47]:

	0	1	2	3	4	5	6	7	8	
Page total likes	139441	139441	139441	139441	139441	139441	139441	139441	139441	139
Type	Photo	Status	Photo	Photo	Photo	Status	Photo	Photo	Status	Ph
Category	2	2	3	2	2	2	3	3	2	
Post Month	12	12	12	12	12	12	12	12	12	
Post Weekday	4	3	3	2	2	1	1	7	7	
Post Hour	3	10	3	10	3	9	3	9	3	
Paid	0.0	0.0	0.0	1.0	0.0	0.0	1.0	1.0	0.0	
Lifetime Post Total Reach	2752	10460	2413	50128	7244	10472	11692	13720	11844	4
Lifetime Post Total Impressions	5091	19057	4373	87991	13594	20849	19479	24137	22538	8
Lifetime Engaged Users	178	1457	177	2211	671	1191	481	537	1530	.
Lifetime Post Consumers	109	1361	113	790	410	1073	265	232	1407	
Lifetime Post Consumptions	159	1674	154	1119	580	1389	364	305	1692	.
Lifetime Post Impressions by people who have liked your Page	3078	11710	2812	61027	6228	16034	15432	19728	15220	4
Lifetime Post reach by people who like your Page	1640	6112	1503	32048	3200	7852	9328	11056	7912	2
Lifetime People who have liked your Page and engaged with your post	119	1108	132	1386	396	1016	379	422	1250	
comment	4	5	0	58	19	1	3	0	0	
like	79.0	130.0	66.0	1572.0	325.0	152.0	249.0	325.0	161.0	11
share	17.0	29.0	14.0	147.0	49.0	33.0	27.0	14.0	31.0	2
Total Interactions	100	164	80	1777	393	186	279	339	192	

19 rows × 500 columns

u) shape of transpose

In [48]:
```
transposeData.shape
```

Out[48]: (19, 500)

v) Sorting the dataset

In [49]:
```
sortedData=data.sort_values(by="Post Month")
sortedData
```

Out[49]:

	Page total likes	Type	Category	Post Month	Post Weekday	Post Hour	Paid	Lifetime Post Total Reach	Lifetime Post Total Impressions	Lifetime Engaged Users	Co
499	81370	Photo	2	1	4	4	NaN	4188	7292	564	
475	86909	Photo	3	1	6	16	1.0	5754	9238	1179	
476	86909	Photo	1	1	6	10	1.0	37088	10966	2728	
477	86909	Link	1	1	6	4	0.0	39600	7927	572	
478	86909	Photo	3	1	5	13	0.0	5536	8745	1141	
...	...	...	...	...	...	...	...	...	...	...	
29	138895	Video	1	12	4	11	1.0	36208	61262	1141	
28	138895	Photo	2	12	5	3	0.0	9560	18264	973	
27	138458	Photo	3	12	5	3	0.0	2478	4306	212	
38	138895	Status	2	12	7	9	0.0	17576	33058	5352	
0	139441	Photo	2	12	4	3	0.0	2752	5091	178	

500 rows × 19 columns

w) reshape the dataset

In [50]:
```
reshapeData=data.melt(id_vars=["Type","Category"],value_vars=["Post Month","Post Ho
reshapeData
```

Out[50]:

	Type	Category	variable	value
0	Photo	2	Post Month	12
1	Status	2	Post Month	12
2	Photo	3	Post Month	12
3	Photo	2	Post Month	12
4	Photo	2	Post Month	12
...	...	...	...	...
995	Photo	3	Post Hour	2
996	Photo	2	Post Hour	8
997	Photo	1	Post Hour	2
998	Photo	3	Post Hour	11
999	Photo	2	Post Hour	4

1000 rows × 4 columns

x) shape of the reshaped dataset

In [51]:
```
reshapeData.shape
```

Out[51]:
```
(1000, 4)
```