

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 numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
df = pd.read_csv('dataset_Facebook.csv', sep=';')
df
```

Out[1]:

	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 [2]:

df.head()

Out[2]:

	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

In [3]:

df.describe() #its shows count,mean,std,min

Out[3]:

	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	500.000000	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 [4]:

df.info()

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 500 entries, 0 to 499

Data columns (total 19 columns):

#	Column	Non-Null Count	Dtype
0	Page total likes	500 non-null	int64
1	Type	500 non-null	object
2	Category	500 non-null	int64
3	Post Month	500 non-null	int64
4	Post Weekday	500 non-null	int64
5	Post Hour	500 non-null	int64
6	Paid	499 non-null	float64
7	Lifetime Post Total Reach	500 non-null	int64
8	Lifetime Post Total Impressions	500 non-null	int64
9	Lifetime Engaged Users	500 non-null	int64
10	Lifetime Post Consumers	500 non-null	int64

```

11 Lifetime Post Consumptions          500 non-null    int64
12 Lifetime Post Impressions by people who have liked your Page  500 non-null    int64
13 Lifetime Post reach by people who like your Page              500 non-null    int64
14 Lifetime People who have liked your Page and engaged with your post 500 non-null    int64
15 comment                  500 non-null    int64
16 like                     499 non-null    float64
17 share                    496 non-null    float64
18 Total Interactions       500 non-null    int64
dtypes: float64(3), int64(15), object(1)
memory usage: 74.3+ KB

```

In [5]: `df.shape` *#its show,s size of data set*

Out[5]: (500, 19)

a) Create Subset

In [6]: `temp1 = df.loc[10:25]`
temp1

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	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	
10	139441	Status		2	12	5	10	0.0	21744	42334	4258	4100	4540	37849	18952	3798
11	139441	Photo		2	12	5	10	0.0	3112	5590	208	127	145	3887	2174	165
12	139441	Photo		2	12	5	10	0.0	2847	5133	193	115	133	3779	2072	152
13	139441	Photo		2	12	5	3	0.0	2549	4896	249	134	168	3631	1917	183
14	138414	Photo		2	12	4	5	1.0	22784	39941	887	337	417	34415	19312	684

15	138414	Status	2	12	3	10	0.0	10060	19680	1264	1209	1425	17272	8548	1162
16	138414	Photo	3	12	3	3	0.0	1722	2981	163	123	148	1868	1050	123
17	138414	Photo	1	12	2	12	1.0	53264	111785	1706	1103	1655	92512	39776	1307
18	138414	Status	3	12	2	3	0.0	3930	7509	130	86	112	5009	2410	101
19	138414	Photo	3	12	1	11	0.0	1591	2825	121	88	111	2116	1161	100
20	138414	Photo	2	12	1	3	0.0	2848	5066	200	142	184	3561	1963	157
21	138414	Photo	1	12	7	10	0.0	1384	2467	15	15	20	2196	1172	15
22	138414	Link	1	12	7	10	0.0	3454	6853	118	104	130	6282	3100	106
23	138414	Photo	3	12	7	3	0.0	2723	4888	176	118	143	2964	1621	143
24	138414	Status	2	12	6	10	0.0	8488	15294	1341	1270	1489	9684	5244	995
25	138458	Status	2	12	6	3	0.0	8284	15104	1521	1462	1711	10266	5372	1200

b) Merge Data

```
In [7]: df1 = df[['Type', 'Category', 'comment']].loc[4:17]
df2 = df[['Type', 'Category', 'Paid']].loc[24:30]
df3 = df[['Type', 'Category', 'Paid']].loc[31:35]
#merging two or more datasets
pd.concat([df1, df2, df3])
```

```
Out[7]:
```

	Type	Category	comment	Paid
4	Photo	2	19.0	NaN
5	Status	2	1.0	NaN
6	Photo	3	3.0	NaN
7	Photo	3	0.0	NaN
8	Status	2	0.0	NaN
9	Photo	3	3.0	NaN

	Type	Category	comment	Paid
10	Status	2	0.0	NaN
11	Photo	2	0.0	NaN
12	Photo	2	0.0	NaN
13	Photo	2	5.0	NaN
14	Photo	2	2.0	NaN
15	Status	2	4.0	NaN
16	Photo	3	2.0	NaN
17	Photo	1	15.0	NaN
24	Status	2	NaN	0.0
25	Status	2	NaN	0.0
26	Status	2	NaN	0.0
27	Photo	3	NaN	0.0
28	Photo	2	NaN	0.0
29	Video	1	NaN	1.0
30	Photo	2	NaN	0.0
31	Photo	2	NaN	0.0
32	Photo	3	NaN	0.0
33	Photo	3	NaN	0.0
34	Photo	1	NaN	0.0
35	Photo	2	NaN	0.0

c) Sort Data

```
In [8]: df.sort_values(['Category', 'Paid'], ascending=False)
```

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	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
6	139441	Photo	3	12	1	3	1.0	11692	19479	481	265	364	15432	9328	379
7	139441	Photo	3	12	7	9	1.0	13720	24137	537	232	305	19728	11056	422
83	137177	Photo	3	11	5	10	1.0	2938	5799	321	243	379	3668	1848	236
143	136013	Photo	3	10	2	10	1.0	16776	39549	714	440	720	11860	6200	590
167	135428	Photo	3	9	4	2	1.0	10748	19724	892	498	719	13674	7624	704
...
474	91009	Photo	1	2	1	12	0.0	21928	39641	1512	1479	1837	6338	3672	497
477	86909	Link	1	1	6	4	0.0	39600	7927	572	496	581	12522	8176	167
485	86491	Link	1	1	2	2	0.0	5168	8371	66	59	71	7041	3996	58
492	85979	Link	1	1	5	11	0.0	45920	5808	753	655	763	15766	10720	220
497	81370	Photo	1	1	5	2	0.0	3778	7216	625	572	795	4742	2388	363

500 rows × 19 columns



d) Transpose Data

In [9]:

```
df.transpose()
```

Out[9]:

	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
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	0.0	0.0	1.0	0.0	0.0	1.0	1.0	0.0	0.0	...	0.0	1.0	0.0	0.0	0.0	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	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
have liked your Page and engaged with your post																		
comment	4	5	0	58	19	1	3	0	0	3	...	1	1	0	17	10	5	0
like	79.0	130.0	66.0	1572.0	325.0	152.0	249.0	325.0	161.0	113.0	...	79.0	105.0	128.0	185.0	125.0	53.0	53.0
share	17.0	29.0	14.0	147.0	49.0	33.0	27.0	14.0	31.0	26.0	...	30.0	46.0	9.0	55.0	41.0	26.0	22.0
Total Interactions	100	164	80	1777	393	186	279	339	192	142	...	110	152	137	257	176	84	75

19 rows × 500 columns

```
In [10]: pd.pivot_table(df, index=['Type', 'Category', 'Paid'], values=['like', 'comment'])
```

```
Out[10]:
```

			comment	like
Type	Category	Paid		
Link	1	0.0	2.428571	83.785714
		1.0	4.000000	56.666667
	2	0.0	2.000000	32.000000
		0.0	2.000000	68.000000
	3	0.0	5.742424	114.793893
		1.0	6.283019	153.698113
Photo	1	0.0	7.313433	154.985075
		1.0	24.956522	477.739130
	2	0.0	7.205607	218.990654
		1.0	6.186047	221.651163
	3	0.0	3.500000	99.500000
		0.0	3.500000	99.500000
Status	1	0.0	3.500000	99.500000

			comment	like
Type	Category	Paid		
		1.0	6.000000	210.000000
	2	0.0	8.068966	150.655172
		1.0	15.888889	285.333333
	3	0.0	2.750000	151.500000
Video	1	0.0	10.333333	216.000000
		1.0	13.750000	243.000000

e) Shape and reshape

```
In [14]: arr = np.array([1, 2, 3, 4, 5, 6, 7, 8, 9, 10])
arr.reshape(5, 2)
```

```
Out[14]: array([[ 1,  2],
               [ 3,  4],
               [ 5,  6],
               [ 7,  8],
               [ 9, 10]])
```

```
In [ ]: arr.shape
```

```
In [ ]:
```

```
In [16]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 500 entries, 0 to 499
```

```
Data columns (total 19 columns):
```

```
#    Column
```

```
---
```

```
0    Page total likes
```

```
Non-Null Count  Dtype
```

```
-----
```

```
500 non-null    int64
```

1	Type	500	non-null	object
2	Category	500	non-null	int64
3	Post Month	500	non-null	int64
4	Post Weekday	500	non-null	int64
5	Post Hour	500	non-null	int64
6	Paid	499	non-null	float64
7	Lifetime Post Total Reach	500	non-null	int64
8	Lifetime Post Total Impressions	500	non-null	int64
9	Lifetime Engaged Users	500	non-null	int64
10	Lifetime Post Consumers	500	non-null	int64
11	Lifetime Post Consumptions	500	non-null	int64
12	Lifetime Post Impressions by people who have liked your Page	500	non-null	int64
13	Lifetime Post reach by people who like your Page	500	non-null	int64
14	Lifetime People who have liked your Page and engaged with your post	500	non-null	int64
15	comment	500	non-null	int64
16	like	499	non-null	float64
17	share	496	non-null	float64
18	Total Interactions	500	non-null	int64

dtypes: float64(3), int64(15), object(1)

memory usage: 74.3+ KB

In []:

In []:

In []:

In []:

In []:

In []:

In []:

```
a = np.array([[10,20,30,40],[50,60,70,80]])
```

In []:

`a`

In []:

`np.append(a, [[90, 100, 120, 130]], axis=0)`

In [11]:

`b = np.array([10, 20, 30, 40, 50])`

In [12]:

`b`Out[12]: `array([10, 20, 30, 40, 50])`

In [13]:

`b.shape`Out[13]: `(5,)`