

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 [4]: import pandas as pd
import numpy as np
df = pd.read_csv(r"C:\Users\yasha\Desktop\Ashish\sem 6\DSBDA\datasets\dataset_11\facebook_data.csv")
```

Out[4]:

| | 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 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 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



In [5]: `df.describe()`

Out[5]:

| | Page total likes | Category | Post Month | Post Weekday | Post Hour | Paid | Lifetime Pos Total Reacl |
|--------------|---------------------|------------|---------------|-----------------|------------|------------|-----------------------------|
| count | 500.000000 | 500.000000 | 500.000000 | 500.000000 | 500.000000 | 499.000000 | 500.000000 |
| mean | 123194.176000 | 1.880000 | 7.038000 | 4.150000 | 7.840000 | 0.278557 | 13903.360000 |
| std | 16272.813214 | 0.852675 | 3.307936 | 2.030701 | 4.368589 | 0.448739 | 22740.787800 |
| min | 81370.000000 | 1.000000 | 1.000000 | 1.000000 | 1.000000 | 0.000000 | 238.000000 |
| 25% | 112676.000000 | 1.000000 | 4.000000 | 2.000000 | 3.000000 | 0.000000 | 3315.000000 |
| 50% | 129600.000000 | 2.000000 | 7.000000 | 4.000000 | 9.000000 | 0.000000 | 5281.000000 |
| 75% | 136393.000000 | 3.000000 | 10.000000 | 6.000000 | 11.000000 | 1.000000 | 13168.000000 |
| max | 139441.000000 | 3.000000 | 12.000000 | 7.000000 | 23.000000 | 1.000000 | 180480.000000 |

In [6]: `df.shape`

Out[6]: (500, 19)

creating the subsets

In [8]: `# subset 1`

```
df1 = df[['Page total likes', 'Category', 'Post Weekday']].loc[0:15]  
df1
```

Out[8]:

| | Page total likes | Category | Post Weekday |
|----|------------------|----------|--------------|
| 0 | 139441 | 2 | 4 |
| 1 | 139441 | 2 | 3 |
| 2 | 139441 | 3 | 3 |
| 3 | 139441 | 2 | 2 |
| 4 | 139441 | 2 | 2 |
| 5 | 139441 | 2 | 1 |
| 6 | 139441 | 3 | 1 |
| 7 | 139441 | 3 | 7 |
| 8 | 139441 | 2 | 7 |
| 9 | 139441 | 3 | 6 |
| 10 | 139441 | 2 | 5 |
| 11 | 139441 | 2 | 5 |
| 12 | 139441 | 2 | 5 |
| 13 | 139441 | 2 | 5 |
| 14 | 138414 | 2 | 4 |
| 15 | 138414 | 2 | 3 |

In [9]: `# subset 2`

```
df2 = df[['Page total likes', 'Category', 'Post Weekday']].loc[16:30]  
df2
```

Out[9]:

| | Page total likes | Category | Post Weekday |
|----|------------------|----------|--------------|
| 16 | 138414 | 3 | 3 |
| 17 | 138414 | 1 | 2 |
| 18 | 138414 | 3 | 2 |
| 19 | 138414 | 3 | 1 |
| 20 | 138414 | 2 | 1 |
| 21 | 138414 | 1 | 7 |
| 22 | 138414 | 1 | 7 |
| 23 | 138414 | 3 | 7 |
| 24 | 138414 | 2 | 6 |
| 25 | 138458 | 2 | 6 |
| 26 | 138458 | 2 | 5 |
| 27 | 138458 | 3 | 5 |
| 28 | 138895 | 2 | 5 |
| 29 | 138895 | 1 | 4 |
| 30 | 138895 | 2 | 4 |

In [10]: `# subset 3`

```
df3 = df[['Page total likes', 'Category', 'Post Weekday']].loc[31:50]  
df3
```

Out[10]:

| | Page total likes | Category | Post Weekday |
|----|------------------|----------|--------------|
| 31 | 138895 | 2 | 3 |
| 32 | 138895 | 3 | 3 |
| 33 | 138895 | 3 | 2 |
| 34 | 138895 | 1 | 2 |
| 35 | 138895 | 2 | 1 |
| 36 | 138895 | 3 | 1 |
| 37 | 138895 | 1 | 7 |
| 38 | 138895 | 2 | 7 |
| 39 | 138895 | 1 | 7 |
| 40 | 138895 | 2 | 6 |
| 41 | 138895 | 1 | 6 |
| 42 | 138353 | 1 | 5 |
| 43 | 138353 | 1 | 5 |
| 44 | 138353 | 1 | 4 |
| 45 | 138353 | 1 | 4 |
| 46 | 138353 | 1 | 3 |
| 47 | 138353 | 1 | 3 |
| 48 | 138353 | 1 | 2 |
| 49 | 138353 | 1 | 2 |
| 50 | 138353 | 2 | 1 |

Merge data

```
In [12]: mergeing=pd.concat([df1,df2,df3])  
mergeing
```

Out[12]:

| | Page total likes | Category | Post Weekday |
|----|------------------|----------|--------------|
| 0 | 139441 | 2 | 4 |
| 1 | 139441 | 2 | 3 |
| 2 | 139441 | 3 | 3 |
| 3 | 139441 | 2 | 2 |
| 4 | 139441 | 2 | 2 |
| 5 | 139441 | 2 | 1 |
| 6 | 139441 | 3 | 1 |
| 7 | 139441 | 3 | 7 |
| 8 | 139441 | 2 | 7 |
| 9 | 139441 | 3 | 6 |
| 10 | 139441 | 2 | 5 |
| 11 | 139441 | 2 | 5 |
| 12 | 139441 | 2 | 5 |
| 13 | 139441 | 2 | 5 |
| 14 | 138414 | 2 | 4 |
| 15 | 138414 | 2 | 3 |
| 16 | 138414 | 3 | 3 |
| 17 | 138414 | 1 | 2 |
| 18 | 138414 | 3 | 2 |
| 19 | 138414 | 3 | 1 |
| 20 | 138414 | 2 | 1 |
| 21 | 138414 | 1 | 7 |
| 22 | 138414 | 1 | 7 |
| 23 | 138414 | 3 | 7 |
| 24 | 138414 | 2 | 6 |
| 25 | 138458 | 2 | 6 |
| 26 | 138458 | 2 | 5 |
| 27 | 138458 | 3 | 5 |
| 28 | 138895 | 2 | 5 |
| 29 | 138895 | 1 | 4 |
| 30 | 138895 | 2 | 4 |
| 31 | 138895 | 2 | 3 |
| 32 | 138895 | 3 | 3 |
| 33 | 138895 | 3 | 2 |
| 34 | 138895 | 1 | 2 |
| 35 | 138895 | 2 | 1 |

| | Page total likes | Category | Post Weekday |
|----|------------------|----------|--------------|
| 36 | 138895 | 3 | 1 |
| 37 | 138895 | 1 | 7 |
| 38 | 138895 | 2 | 7 |
| 39 | 138895 | 1 | 7 |
| 40 | 138895 | 2 | 6 |
| 41 | 138895 | 1 | 6 |
| 42 | 138353 | 1 | 5 |
| 43 | 138353 | 1 | 5 |
| 44 | 138353 | 1 | 4 |
| 45 | 138353 | 1 | 4 |
| 46 | 138353 | 1 | 3 |
| 47 | 138353 | 1 | 3 |
| 48 | 138353 | 1 | 2 |
| 49 | 138353 | 1 | 2 |
| 50 | 138353 | 2 | 1 |

sort data


```
In [25]: sort_values=df.sort_values('Page total likes',ascending=False)
sort_values
```

Out[25]:

| | 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 |
| 8 | 139441 | Status | | 2 | 12 | 7 | 3 | 0.0 | 11844 | 22538 | 1530 |
| 1 | 139441 | Status | | 2 | 12 | 3 | 10 | 0.0 | 10460 | 19057 | 1457 |
| 12 | 139441 | Photo | | 2 | 12 | 5 | 10 | 0.0 | 2847 | 5133 | 193 |
| 11 | 139441 | Photo | | 2 | 12 | 5 | 10 | 0.0 | 3112 | 5590 | 208 |
| ... | ... | ... | | ... | ... | ... | ... | ... | ... | ... | |
| 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



transpose data

```
In [26]: df.transpose()
```

Out[26]:

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | ... | 490 | |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|------|-------|---|
| 139441 | 139441 | 139441 | 139441 | 139441 | 139441 | 139441 | 139441 | 139441 | 139441 | 139441 | ... | 85979 | 8 |
| | Photo | Status | Photo | Photo | Photo | Status | Photo | Photo | Status | Photo | ... | Photo | F |
| | 2 | 2 | 3 | 2 | 2 | 2 | 3 | 3 | 2 | 3 | ... | 3 | |
| | 12 | 12 | 12 | 12 | 12 | 12 | 12 | 12 | 12 | 12 | ... | 1 | |
| | 4 | 3 | 3 | 2 | 2 | 1 | 1 | 7 | 7 | 6 | ... | 6 | |
| | 3 | 10 | 3 | 10 | 3 | 9 | 3 | 9 | 3 | 10 | ... | 11 | |
| | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 | 0.0 | ... | 0.0 | |
| 2752 | 10460 | 2413 | 50128 | 7244 | 10472 | 11692 | 13720 | 11844 | 4694 | ... | 5280 | | |
| 5091 | 19057 | 4373 | 87991 | 13594 | 20849 | 19479 | 24137 | 22538 | 8668 | ... | 8703 | 1 | |
| | 178 | 1457 | 177 | 2211 | 671 | 1191 | 481 | 537 | 1530 | 280 | ... | 951 | |
| | 109 | 1361 | 113 | 790 | 410 | 1073 | 265 | 232 | 1407 | 183 | ... | 911 | |
| | 159 | 1674 | 154 | 1119 | 580 | 1389 | 364 | 305 | 1692 | 250 | ... | 1237 | |
| 3078 | 11710 | 2812 | 61027 | 6228 | 16034 | 15432 | 19728 | 15220 | 4309 | ... | 5757 | | |
| 1640 | 6112 | 1503 | 32048 | 3200 | 7852 | 9328 | 11056 | 7912 | 2324 | ... | 3300 | | |
| | 119 | 1108 | 132 | 1386 | 396 | 1016 | 379 | 422 | 1250 | 199 | ... | 431 | |
| | 4 | 5 | 0 | 58 | 19 | 1 | 3 | 0 | 0 | 3 | ... | 1 | |
| 79.0 | 130.0 | 66.0 | 1572.0 | 325.0 | 152.0 | 249.0 | 325.0 | 161.0 | 113.0 | ... | 79.0 | 1 | |
| 17.0 | 29.0 | 14.0 | 147.0 | 49.0 | 33.0 | 27.0 | 14.0 | 31.0 | 26.0 | ... | 30.0 | | |
| 100 | 164 | 80 | 1777 | 393 | 186 | 279 | 339 | 192 | 142 | ... | 110 | | |

columns

shape and reshape the data

```
In [27]: #shaping of the data  
df.shape
```

```
Out[27]: (500, 19)
```

reshaping the data

```
In [33]: pivot_table=pd.pivot_table(df,index=['Type','Category'], values='like')  
pivot_table
```

```
Out[33]:
```

| | | like |
|--------|----------|------------|
| Type | Category | |
| Link | 1 | 75.650000 |
| | 2 | 32.000000 |
| | 3 | 68.000000 |
| Photo | 1 | 126.000000 |
| | 2 | 235.857143 |
| | 3 | 219.753333 |
| Status | 1 | 136.333333 |
| | 2 | 182.552632 |
| | 3 | 151.500000 |
| Video | 1 | 231.428571 |

```
In [ ]:
```

```
In [ ]:
```