In [1]:
```
import numpy as np
import pandas as pd
```

In [26]:
```
df = pd.read_csv("dataset_Facebook.csv", delimiter=";")
df.head()
```

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	Life Impress by pe who liked
0	139441	Photo	2	12	4	3	0.0	2752	5091	178	109	159	
1	139441	Status	2	12	3	10	0.0	10460	19057	1457	1361	1674	1
2	139441	Photo	3	12	3	3	0.0	2413	4373	177	113	154	
3	139441	Photo	2	12	2	10	1.0	50128	87991	2211	790	1119	6
4	139441	Photo	2	12	2	3	0.0	7244	13594	671	410	580	

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

Out[4]: (500, 19)

In [42]:
```
subset1 = df[df["Type"] == "Photo"]
subset1.head()
```

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	Lifetime Post Consumers	Lifetime Post Consumptions	Life Impress by pe who liked I
0	139441	Photo	2	12	4	3	0.0	2752	5091	178	109	159	
2	139441	Photo	3	12	3	3	0.0	2413	4373	177	113	154	
3	139441	Photo	2	12	2	10	1.0	50128	87991	2211	790	1119	6
4	139441	Photo	2	12	2	3	0.0	7244	13594	671	410	580	
6	139441	Photo	3	12	1	3	1.0	11692	19479	481	265	364	1

In [38]:
```
subset2 =   df[df['comment']>100]
subset2
```

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	Lifetime Post Consumers	Lifetime Post Consumptions	L Impre by wh like
168	135428	Photo	1	9	3	10	0.0	41984	68290	3370	2420	4074	
244	130791	Photo	2	7	3	5	1.0	180480	319133	8072	4010	6242	
288	126141	Photo	1	6	4	12	0.0	20896	29062	1418	1038	2048	
460	92507	Photo	3	2	1	13	0.0	55520	665792	4544	3586	6624	

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

Out[35]: (426, 19)

In [36]:
```
subset2.shape
```

Out[36]: (4, 19)

In [37]:
```
merged_data = pd.merge(subset1, subset2)
merged_data
```

Out[37]:

	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	Life Impress by pe who liked I
0	135428	Photo	1	9	3	10	0.0	41984	68290	3370	2420	4074	3
1	130791	Photo	2	7	3	5	1.0	180480	319133	8072	4010	6242	10
2	126141	Photo	1	6	4	12	0.0	20896	29062	1418	1038	2048	1
3	92507	Photo	3	2	1	13	0.0	55520	665792	4544	3586	6624	64

In [13]:
```
merged_data.shape
```

Out[13]: (4, 19)

In [14]:
```
df.isna().sum()
```

Out[14]:
```
Page total likes                                                  0
Type                                                              0
Category                                                          0
Post Month                                                        0
Post Weekday                                                      0
Post Hour                                                         0
Paid                                                              1
Lifetime Post Total Reach                                         0
Lifetime Post Total Impressions                                   0
Lifetime Engaged Users                                            0
Lifetime Post Consumers                                           0
Lifetime Post Consumptions                                        0
Lifetime Post Impressions by people who have liked your Page      0
Lifetime Post reach by people who like your Page                  0
Lifetime People who have liked your Page and engaged with your post    0
comment                                                           0
like                                                              1
share                                                             4
Total Interactions                                                0
dtype: int64
```

In [15]:
```
print("Missing Values: ", df.isnull().sum().sum())
```

```
Missing Values:  6
```

In [16]:
```
df['Paid'].fillna(df['Paid'].mean(), inplace=True)
df['like'].fillna(df['like'].mean(), inplace=True)
df['share'].fillna(df['share'].mean(), inplace=True)
```

In [17]:
```
print("DataFrame after handling missing values: ", df.isna().sum())
```

```
DataFrame after handling missing values:  Page total likes
0
Type                                                              0
Category                                                          0
Post Month                                                        0
Post Weekday                                                      0
Post Hour                                                         0
Paid                                                              0
Lifetime Post Total Reach                                         0
Lifetime Post Total Impressions                                   0
Lifetime Engaged Users                                            0
Lifetime Post Consumers                                           0
Lifetime Post Consumptions                                        0
Lifetime Post Impressions by people who have liked your Page      0
Lifetime Post reach by people who like your Page                  0
Lifetime People who have liked your Page and engaged with your post    0
comment                                                           0
like                                                              0
share                                                             0
Total Interactions                                                0
dtype: int64
```

In [41]:
```
sorted_data = df.sort_values(by='Total Interactions', ascending=False)
sorted_data.head()
```

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	Lifetime Post Consumers	Lifetime Post Consumptions	L Impre by wh like
244	130791	Photo	2	7	3	5	1.0	180480	319133	8072	4010	6242	
379	111620	Photo	3	4	1	14	1.0	105632	147918	3984	2254	3391	
168	135428	Photo	1	9	3	10	0.0	41984	68290	3370	2420	4074	
460	92507	Photo	3	2	1	13	0.0	55520	665792	4544	3586	6624	
349	117764	Photo	3	5	5	13	0.0	81856	124753	3000	1637	2718	

In [21]:
```
transposed_data = df.T
```

In [40]:
```
transposed_data.head()
```

Out[40]:

	0	1	2	3	4	5	6	7	8	9	...	490	491	492	493
Page total likes	139441	139441	139441	139441	139441	139441	139441	139441	139441	139441	...	85979	85979	85979	85093
Type	Photo	Status	Photo	Photo	Photo	Status	Photo	Photo	Status	Photo	...	Photo	Photo	Link	Photo
Category	2	2	3	2	2	2	3	3	2	3	...	3	3	1	3
Post Month	12	12	12	12	12	12	12	12	12	12	...	1	1	1	1
Post Weekday	4	3	3	2	2	1	1	7	7	6	...	6	6	5	1

5 rows × 500 columns

In [23]:
```
id_vars = ['Page total likes', 'Type', 'Category', 'Post Month', 'Post Weekday', 'Post Hour']
value_vars = ['comment', 'like', 'share', 'Total Interactions']
```

In [24]:
```
melted_data = pd.melt(df, id_vars=id_vars, value_vars=value_vars, var_name='Metric', value_name='Value'
```

In [43]:
```
melted_data.head()
```

Out[43]:

	Page total likes	Type	Category	Post Month	Post Weekday	Post Hour	Metric	Value
0	139441	Photo	2	12	4	3	comment	4.0
1	139441	Status	2	12	3	10	comment	5.0
2	139441	Photo	3	12	3	3	comment	0.0
3	139441	Photo	2	12	2	10	comment	58.0
4	139441	Photo	2	12	2	3	comment	19.0