tiktok

November 27, 2024

[4]:
```
import os
cwd=os.getcwd()
print(cwd)
```

```
C:\Users\user
```

[6]:
```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
```

[7]:
```
df=pd.read_csv("Tiktok dataset.csv")
print(df)
```

```
   POST TITLE    VIEWS  LIKES  SHARES  COMMENTS                    HASHTAGS  \
0       POST 1    7300    101       2       1.0                     #capcut
1       POST 2    8100     97       3       3.0                     #capcut
2       POST 3    9500    133       4       2.0                     #capcut
3       POST 4   12300    288       5       6.0                  #lovemyself
4       POST 5   13600    248       6       1.0                     #capcut
5       POST 6   12000    219       4       4.0                     #capcut
6       POST 7   12100    163       4       2.0                   #costatrip
7       POST 8   12700      1       4       1.0                  #mcmidoclub
8       POST 9   14200    218       6       5.0                   #kidudumtu
9      POST 10   66500   2327      55      21.0                 #bombastieye
10     POST 11   29700    577       4       2.0                     #parents
11     POST 12   25700    321       3       2.0                      #wambea
12     POST 13   37100    929      16       3.0                  #stayloyal
13     POST 14   35400    693      13       6.0               #kabiveronicah
14     POST 15   34200    645      12      10.0                   #pullingup
15     POST 16  182700   7575      53       NaN              #you22challenge
16     POST 17  153200   5831      39       NaN                        #ur22
17     POST 18   87700   2668      24       NaN         #capcut#kabiveronicah
18     POST 19   44200   1001      16       9.0                  #kilimanjaro
19     POST 20  431400  26700     232     275.0         #DianaB#kabiveronicah
20     POST 21   62300   1933      21      15.0                     #capcut
21     POST 22   41700    702      10       7.0               #kabiveronicah
22     POST 23  147300   6955      78     124.0               #kabiveronicah
23     POST 24  200200  10900     112     137.0               #kabiveronicah
24     POST 25  109300   5977      62     126.0               #kabiveronicah
```

25	POST 26	49700	3139	40	128.0	#kabiveronicah#selflove
26	POST 27	88200	5940	41	136.0	#lazimanimake#kabiveronicah
27	POST 28	53400	1947	37	58.0	#kabiveronicah
28	POST 29	143700	8300	93	177.0	#kabiveronicah
29	POST 30	62600	2979	28	NaN	#kabiveronicah
30	POST 31	13600	653	8	22.0	#kabiveronicah
31	POST 32	40000	3081	25	74.0	#capcut
32	POST 33	15900	725	8	18.0	#capcut
33	POST 34	41000	2123	34	52.0	#foryou#kabiveronicah
34	POST 35	22200	1063	9	23.0	#trending#kabiveronicah

	DATE
0	31/1/2023
1	1/2/2023
2	1/2/2023
3	28/2/2023
4	12/3/2023
5	12/3/2023
6	12/3/2023
7	12/3/2023
8	12/3/2023
9	16/03/2023
10	16/3/2023
11	16/3/2023
12	18/3/2023
13	2/4/2023
14	2/4/20223
15	2/4/2023
16	3/4/20223
17	3/4/2023
18	3/4/20223
19	3/4/2023
20	14/4/2023
21	29/4/2023
22	1/6/2023
23	2/6/2023
24	19/6/2023
25	28/6/2023
26	5/7/2023
27	7/7/2023
28	9/7/2023
29	9/7/2023
30	27/7/2023
31	30/7/2023
32	31/7/2023
33	5/8/2023
34	13/8/2023

[8]:
```
df.isnull().sum()
```

[8]:
```
POST TITLE     0
VIEWS          0
LIKES          0
SHARES         0
COMMENTS       4
HASHTAGS       0
DATE           0
dtype: int64
```

[9]:
```
column_to_fill=['COMMENTS']
```

[10]:
```
df[column_to_fill]=df[column_to_fill].fillna(0)
print(df)
```

```
   POST TITLE    VIEWS  LIKES  SHARES  COMMENTS                    HASHTAGS  \
0       POST 1    7300    101       2       1.0                     #capcut
1       POST 2    8100     97       3       3.0                     #capcut
2       POST 3    9500    133       4       2.0                     #capcut
3       POST 4   12300    288       5       6.0                  #lovemyself
4       POST 5   13600    248       6       1.0                     #capcut
5       POST 6   12000    219       4       4.0                     #capcut
6       POST 7   12100    163       4       2.0                   #costatrip
7       POST 8   12700      1       4       1.0                  #mcmidoclub
8       POST 9   14200    218       6       5.0                   #kidudumtu
9      POST 10   66500   2327      55      21.0                 #bombastieye
10     POST 11   29700    577       4       2.0                     #parents
11     POST 12   25700    321       3       2.0                     #wambea
12     POST 13   37100    929      16       3.0                   #stayloyal
13     POST 14   35400    693      13       6.0               #kabiveronicah
14     POST 15   34200    645      12      10.0                   #pullingup
15     POST 16  182700   7575      53       0.0              #you22challenge
16     POST 17  153200   5831      39       0.0                        #ur22
17     POST 18   87700   2668      24       0.0        #capcut#kabiveronicah
18     POST 19   44200   1001      16       9.0                #kilimanjaro
19     POST 20  431400  26700     232     275.0        #DianaB#kabiveronicah
20     POST 21   62300   1933      21      15.0                     #capcut
21     POST 22   41700    702      10       7.0               #kabiveronicah
22     POST 23  147300   6955      78     124.0               #kabiveronicah
23     POST 24  200200  10900     112     137.0               #kabiveronicah
24     POST 25  109300   5977      62     126.0               #kabiveronicah
25     POST 26   49700   3139      40     128.0     #kabiveronicah#selflove
26     POST 27   88200   5940      41     136.0  #lazimanimake#kabiveronicah
27     POST 28   53400   1947      37      58.0               #kabiveronicah
28     POST 29  143700   8300      93     177.0               #kabiveronicah
29     POST 30   62600   2979      28       0.0               #kabiveronicah
30     POST 31   13600    653       8      22.0               #kabiveronicah
```

```
31     POST 32   40000   3081       25      74.0                      #capcut
32     POST 33   15900    725        8      18.0                      #capcut
33     POST 34   41000   2123       34      52.0        #foryou#kabiveronicah
34     POST 35   22200   1063        9      23.0      #trending#kabiveronicah

          DATE
0    31/1/2023
1     1/2/2023
2     1/2/2023
3    28/2/2023
4    12/3/2023
5    12/3/2023
6    12/3/2023
7    12/3/2023
8    12/3/2023
9   16/03/2023
10   16/3/2023
11   16/3/2023
12   18/3/2023
13    2/4/2023
14   2/4/20223
15    2/4/2023
16   3/4/20223
17    3/4/2023
18   3/4/20223
19    3/4/2023
20   14/4/2023
21   29/4/2023
22    1/6/2023
23    2/6/2023
24   19/6/2023
25   28/6/2023
26    5/7/2023
27    7/7/2023
28    9/7/2023
29    9/7/2023
30   27/7/2023
31   30/7/2023
32   31/7/2023
33    5/8/2023
34   13/8/2023
```

[11]:
```
df.describe()
```

[11]:

	VIEWS	LIKES	SHARES	COMMENTS
count	35.000000	35.000000	35.000000	35.000000
mean	66305.714286	3061.485714	31.742857	41.428571

```
std      82410.821473   4976.919732   44.258162   65.464533
min       7300.000000      1.000000    2.000000    0.000000
25%      13900.000000    304.500000    5.500000    2.000000
50%      40000.000000   1001.000000   16.000000    7.000000
75%      77100.000000   3110.000000   39.500000   55.000000
max     431400.000000  26700.000000  232.000000  275.000000
```

[12]:
```
print(df.dtypes)
```

```
POST TITLE      object
VIEWS            int64
LIKES            int64
SHARES           int64
COMMENTS       float64
HASHTAGS        object
DATE            object
dtype: object
```

[13]:
```
numerical_df=df.select_dtypes(include=['number'])
print(numerical_df)
```

```
     VIEWS  LIKES  SHARES  COMMENTS
0     7300    101       2       1.0
1     8100     97       3       3.0
2     9500    133       4       2.0
3    12300    288       5       6.0
4    13600    248       6       1.0
5    12000    219       4       4.0
6    12100    163       4       2.0
7    12700      1       4       1.0
8    14200    218       6       5.0
9    66500   2327      55      21.0
10   29700    577       4       2.0
11   25700    321       3       2.0
12   37100    929      16       3.0
13   35400    693      13       6.0
14   34200    645      12      10.0
15  182700   7575      53       0.0
16  153200   5831      39       0.0
17   87700   2668      24       0.0
18   44200   1001      16       9.0
19  431400  26700     232     275.0
20   62300   1933      21      15.0
21   41700    702      10       7.0
22  147300   6955      78     124.0
23  200200  10900     112     137.0
24  109300   5977      62     126.0
25   49700   3139      40     128.0
```

```
26   88200   5940      41     136.0
27   53400   1947      37      58.0
28  143700   8300      93     177.0
29   62600   2979      28       0.0
30   13600    653       8      22.0
31   40000   3081      25      74.0
32   15900    725       8      18.0
33   41000   2123      34      52.0
34   22200   1063       9      23.0
```

[16]:
```
correlation=numerical_df.corr()
print(correlation)
```

```
              VIEWS     LIKES    SHARES  COMMENTS
VIEWS      1.000000  0.980604  0.956370  0.758461
LIKES      0.980604  1.000000  0.977334  0.840072
SHARES     0.956370  0.977334  1.000000  0.874086
COMMENTS   0.758461  0.840072  0.874086  1.000000
```

[20]:
```
grouped=df.groupby(['HASHTAGS'])[['VIEWS','LIKES', 'COMMENTS']].agg({'VIEWS':
  ↪'mean','LIKES':'mean', 'COMMENTS':'mean'})
print(grouped)
```

```
                                      VIEWS         LIKES  COMMENTS
HASHTAGS
#DianaB#kabiveronicah          431400.000000  26700.000000    275.00
#bombastieye                    66500.000000   2327.000000     21.00
#capcut                         21087.500000    817.125000     14.75
#capcut#kabiveronicah           87700.000000   2668.000000      0.00
#costatrip                      12100.000000    163.000000      2.00
#foryou#kabiveronicah           41000.000000   2123.000000     52.00
#kabiveronicah                  89688.888889   4345.111111     73.00
#kabiveronicah#selflove         49700.000000   3139.000000    128.00
#kidudumtu                      14200.000000    218.000000      5.00
#kilimanjaro                    44200.000000   1001.000000      9.00
#lazimanimake#kabiveronicah     88200.000000   5940.000000    136.00
#lovemyself                     12300.000000    288.000000      6.00
#mcmidoclub                     12700.000000      1.000000      1.00
#parents                        29700.000000    577.000000      2.00
#pullingup                      34200.000000    645.000000     10.00
#stayloyal                      37100.000000    929.000000      3.00
#trending#kabiveronicah         22200.000000   1063.000000     23.00
#ur22                           153200.000000   5831.000000      0.00
#wambea                         25700.000000    321.000000      2.00
#you22challenge                 182700.000000   7575.000000      0.00
```

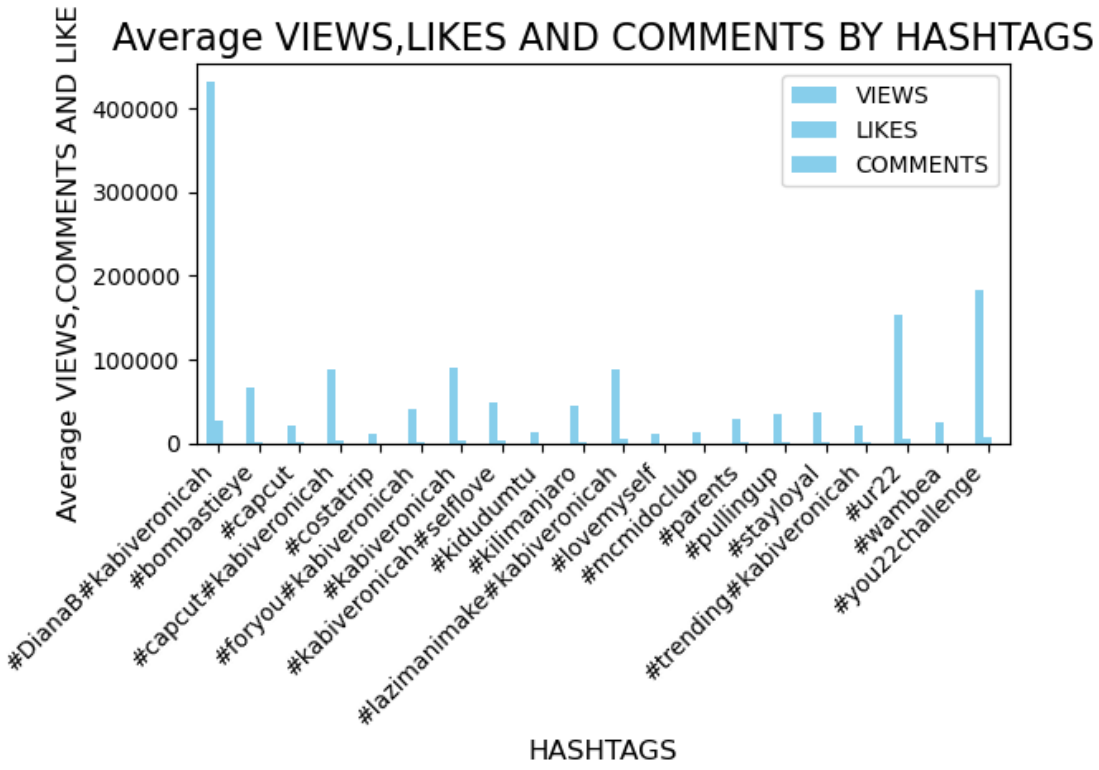
[24]:
```
import matplotlib.pyplot as plt
```

```
plt.figure(figsize=(12, 6))


grouped.plot(kind='bar', color='skyblue', width=0.6)  # Adjust bar width for␣
↪spacing
plt.title("Average VIEWS,LIKES AND COMMENTS BY HASHTAGS", fontsize=16)
plt.xlabel("HASHTAGS", fontsize=12)
plt.ylabel("Average VIEWS,COMMENTS AND LIKES", fontsize=12)


plt.xticks(rotation=45, fontsize=10, ha='right')
plt.tight_layout()  # Adjusts subplot parameters to give space
plt.show()
```

```
<Figure size 1200x600 with 0 Axes>
```



CONCLUSIONS Top-Performing Hashtags:

Hashtags like #kabiveronicah and #you22challenge consistently have the highest averages in views, likes, and comments, indicating strong audience engagement and relevance. Moderate Engagement Hashtags:

Common hashtags like #capcut have high reach (views) but relatively lower engagement (likes and

comments), suggesting they are good for visibility but less effective for deep interaction. Content-Type Preference:

Event-based and challenge-related hashtags, such as #you22challenge, perform well, showing that the audience enjoys trendy and interactive content.

[ ]:

[ ]:

[ ]: