

# Data Analysis Assessment

## Christina(Yiting) Zhang

- What is the typical engagement rate we can expect? What's the likelihood that we can achieve a 15% engagement rate?
- Does day of the week and time of posting affect engagement rates?
- How are our game titles doing in terms of social performance? Is there a specific game we should focus more on or less?
- What media type performs the best?
- What is our best performing campaign?
- Define out a posting strategy for our social channels based on your discoveries.
- What suggestions would you give to the social media team if they want to expand their presence (e.g. if our CSGO youtube channel is doing well should we expand to TikTok?)

## Importing Data

```
In [7]: import pandas as pd
import numpy as np
from matplotlib import pyplot as plt
import seaborn as sns
import os

In [3]: df = pd.read_csv('social_data.csv')
df.head()
```

	Published Date	Account	Account Type	Campaign Name	Total Impressions	Total Engagements	Media Type
0	03-31-2023 19:55	CSGO	TWITTER	N/A	0	0	Text
1	03-31-2023 19:49	CSGO	TWITTER	N/A	0	0	Text
2	03-31-2023 19:49	CSGO	TWITTER	N/A	0	0	Text
3	03-31-2023 19:49	CSGO	TWITTER	N/A	0	0	Text
4	03-31-2023 19:43	CSGO	TWITTER	Community Engagement	9517	1215	Video

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

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3479 entries, 0 to 3478
Data columns (total 7 columns):
#   Column                Non-Null Count  Dtype
---  ---                ---
0   Published Date         3479 non-null   object
1   Account                3479 non-null   object
2   Account Type           3479 non-null   object
3   Campaign Name          3479 non-null   object
4   Total Impressions      3479 non-null   int64
5   Total Engagements      3479 non-null   int64
6   Media Type             3479 non-null   object
dtypes: int64(2), object(5)
memory usage: 190.4+ KB

In [4]: df.describe()
```

	Total Impressions	Total Engagements
count	3479.000000	3479.000000
mean	10972.453579	947.390629
std	29577.108063	3468.628038
min	0.000000	0.000000
25%	167.000000	0.000000
50%	1884.000000	39.000000
75%	12467.000000	276.500000
max	753227.000000	95062.000000

```
In [12]: df['Account Type'].unique()

array(['TWITTER', 'FBPAGE', 'INSTAGRAM', 'TIKTOK_BUSINESS', 'YOUTUBE',
       'LINKEDIN_COMPANY'], dtype=object)
```

## 1. What is the typical engagement rate we can expect? What's the likelihood that we can achieve a 15% engagement rate?

First we should look at the engagements and see if there is any pattern.

```
In [24]: fig = px.box(df, y = "Total Engagements",
                  height=500,width=600,
                  title="Distribution of Total Engagement")
fig.show()
```

A typical engagement we can expect is around 947 across all social media platforms.

```
In [19]: df.groupby('Account Type')['Total Engagements'].describe()
```

	count	mean	std	min	25%	50%	75%	max
FBPAGE	585.0	4405.613675	7317.270756	0.0	0.00	1687.0	5994.0	95062.0
INSTAGRAM	588.0	143.005102	538.534299	0.0	0.00	0.00	6.0	5472.0
LINKEDIN_COMPANY	22.0	38.590909	37.700144	0.0	16.25	28.5	41.5	133.0
TIKTOK_BUSINESS	113.0	166.946903	939.566065	0.0	11.00	40.0	83.0	9680.0
TWITTER	1951.0	3058.71348	974.696962	0.0	9.00	51.0	223.5	12348.0
YOUTUBE	220.0	46.945455	299.680190	0.0	0.00	9.5	35.0	4388.0

We expect most engagement from Facebook and little engagement from LinkedIn and Youtube.

Now let's calculate the 'Engagement Rate'.

- Engagement rate = Engagement / Impression \* 100

```
In [25]: df['Engagement Rate'] = df['Total Engagements'] / df['Total Impressions'] * 100
df.head()
```

	Published Date	Account	Account Type	Campaign Name	Total Impressions	Total Engagements	Media Type	Engagement Rate
0	03-31-2023 19:55	CSGO	TWITTER	N/A	0	0	Text	NaN
1	03-31-2023 19:49	CSGO	TWITTER	N/A	0	0	Text	NaN
2	03-31-2023 19:49	CSGO	TWITTER	N/A	0	0	Text	NaN
3	03-31-2023 19:49	CSGO	TWITTER	N/A	0	0	Text	NaN
4	03-31-2023 19:43	CSGO	TWITTER	Community Engagement	9517	1215	Video	12.766628

```
In [27]: def replaceNullByZero(col):
    """
    Replace NaN and infinity values by zeros
    """
    col_value = col[0]
    col_value = np.nan_to_num(col_value, nan=0, posinf=0, neginf=0)
    return col_value
```

```
In [32]: df['Engagement Rate'] = df[['Engagement Rate']].apply(replaceNullByZero,axis=1)
df.head(100)
```

	Published Date	Account	Account Type	Campaign Name	Total Impressions	Total Engagements	Media Type	Engagement Rate
0	03-31-2023 19:55	CSGO	TWITTER	N/A	0	0	Text	0.000000
1	03-31-2023 19:49	CSGO	TWITTER	N/A	0	0	Text	0.000000
2	03-31-2023 19:49	CSGO	TWITTER	N/A	0	0	Text	0.000000
3	03-31-2023 19:49	CSGO	TWITTER	N/A	0	0	Text	0.000000
4	03-31-2023 19:43	CSGO	TWITTER	Community Engagement	9517	1215	Video	12.766628
...	...	...	...	...	...	...	...	...
95	03-30-2023 11:03	General	INSTAGRAM	N/A	0	0	Text	0.000000
96	03-30-2023 11:03	General	INSTAGRAM	Evil Exhibited	7028	100	Video	1.422880
97	03-30-2023 11:00	DOTA2	TWITTER	N/A	1451	47	Text	3.239145
98	03-30-2023 11:00	DOTA2	TWITTER	N/A	0	0	Video	0.000000
99	03-30-2023 11:00	General	FBPAGE	Evil Exhibited	9273	306	Video	3.299903

100 rows x 8 columns

```
In [38]: df['Engagement Rate'].describe()
```

	count	mean	std	min	25%	50%	75%	max
count	3479.000000							
mean	31.833090							
std	1593.821839							
min	0.000000							
25%	0.000000							
50%	2.041949							
75%	5.084670							
max	94000.000000							

Name: Engagement Rate, dtype: float64

From the table above, we can see that a typical engagement rate is 31.8% across all social media platforms.

```
In [31]: df.groupby('Account Type')['Engagement Rate'].describe()
```

	count	mean	std	min	25%	50%	75%	max
FBPAGE	585.0	175.564847	3896.331657	0.0	0.000000	8.536612	17.680797	94000.000000
INSTAGRAM	588.0	0.666769	1.567378	0.0	0.000000	0.000000	0.958168	13.273076
LINKEDIN_COMPANY	22.0	1.795018	0.915007	0.0	0.000000	1.335601	3.960691	
TIKTOK_BUSINESS	113.0	5.316753	2.971130	0.0	0.218302	4.487179	6.951872	16.835700
TWITTER	1951.0	3.248994	3.713681	0.0	0.077716	2.248996	4.496582	60.000000
YOUTUBE	220.0	3.048881	4.102267	0.0	0.000000	2.569124	4.335485	37.931034

From the table above, the mean engagement rate for Facebook is questionable since the percentage is greater than 100, which means that for some of the posts, the engagements are actually greater than impressions.

As for other social media platforms, TikTok has the highest engagement rate around 5.3% and Instagram has the lowest engagement rate of 0.67%.

We should remove the outliers to have a better picture of a typical engagement rate from Facebook.

```
In [36]: comparison_result = df['Total Engagements'] > df['Total Impressions']
value_greater = df[comparison_result]
# Print the values
value_greater
```

	Published Date	Account	Account Type	Campaign Name	Total Impressions	Total Engagements	Media Type	Engagement Rate
837	03-31-2023 12:37	General	FBPAGE	N/A	1	940	Photo	94000.000000
3216	01-19-2023 06:02	General	FBPAGE	N/A	5	44	Photo	880.000000
3457	01-05-2023 11:55	General	FBPAGE	N/A	300	3775	Photo	1258.666667

There are only 3 posts that has more engagements than impressions, we should remove the outliers and check the engagement rate again.

```
In [37]: # Use boolean indexing to filter out the rows where the condition is True
df_filtered = df[comparison_result]
df_filtered.groupby('Account Type')['Engagement Rate'].describe()
```

	count	mean	std	min	25%	50%	75%	max
FBPAGE	582.0	11.283108	12.151957	0.0	0.000000	8.536606	17.436023	100.000000
INSTAGRAM	588.0	0.666769	1.567378	0.0	0.000000	0.000000	0.958168	13.273076
LINKEDIN_COMPANY	22.0	1.795018	0.915007	0.0	0.000000	1.335601	3.960691	
TIKTOK_BUSINESS	113.0	5.316753	2.971130	0.0	0.077716	2.248996	4.496582	60.000000
TWITTER	1951.0	3.248994	3.713681	0.0	0.077716	2.248996	4.496582	60.000000
YOUTUBE	220.0	3.048881	4.102267	0.0	0.000000	2.569124	4.335485	37.931034

```
In [36]: df_filtered['Engagement Rate'].describe()
```

	count	mean	std	min	25%	50%	75%	max
count	3476.000000							
mean	6.202720							
std	6.736290							
min	0.000000							
25%	0.000000							
50%	2.039731							
75%	5.000000							
max	100.000000							

Name: Engagement Rate, dtype: float64

After removing the outliers, we have a clearer picture of the engagement rate from Facebook, 11.3%, which is the highest rate among across all platforms.

Takeaway:

- Since the highest mean engagement rate we are seeing is 11.3% from Facebook, after removing the outliers, it is unlikely we will see a 15% mean engagement rate across all platforms.

## 2. Does day of the week and time of posting affect engagement rates

Let's convert 'Published Date' to 'datetime' object. Day of the week:

- Tuesday: 0
- Monday: 1
- Wednesday: 2
- Thursday: 3
- Friday: 4
- Saturday: 5
- Sunday: 6

```
In [38]: df['Published Date'] = pd.to_datetime(df['Published Date'])
#day of the week
df['Day of the week'] = df['Published Date'].apply(lambda x: x.weekday())
#time
df['time'] = df['Published Date'].apply(lambda x: x.strftime("%H:%M"))
df['time'] = pd.to_datetime(df['time'])
df.head()
```

```
C:\Users\User\AppData\Local\Temp\ipykernel_3448\22488064636.py:6: UserWarning:
Could not infer format, so each element will be parsed individually, falling back to 'dateutil'. To ensure parsing is consistent and as-expected, please specify a format.
```

```
Out[38]:
```

	Published Date	Account	Account Type	Campaign Name	Total Impressions	Media Type	Engagement Rate	Day of the week	time
0	2023-03-31 19:55:00	CSGO	TWITTER	N/A	0	0	Text	0.000000	4 2023-05-28 19:55:00
1	2023-03-31 19:49:00	CSGO	TWITTER	N/A	0	0	Text	0.000000	4 2023-05-28 19:49:00
2	2023-03-31 19:49:00	CSGO	TWITTER	N/A	0	0	Text	0.000000	4 2023-05-28 19:49:00
3	2023-03-31 19:49:00	CSGO	TWITTER	N/A	0	0	Text	0.000000	4 2023-05-28 19:49:00
4	2023-03-31 19:43:00	CSGO	TWITTER	Community Engagement	9517	1215	Video	12.766628	4 2023-05-28 19:43:00

## 2-a Day of the Week

```
In [43]: #removing outliers first
df_filtered = df[comparison_result]
df_filtered.groupby('Day of the week')['Engagement Rate'].describe()
```

	count	mean	std	min	25%	50%	75%	max
Day of the week								
0	234.0	3.907170	5.494242	0.0	0.479198	1.914950	4.258937	37.931034
1	561.0	5.055946	7.629040	0.0	0.000000	2.456992	5.962577	56.480028
2	575.0	3.322831	5.462101	0.0	0.000000	1.771126	4.311582	58.858403
3	704.0	4.915714	8.349649	0.0	0.000000	2.304830	5.450978	100.000000
4	707.0	4.195042	5.896210	0.0	0.505443	2.251704	5.096991	44.713813
5	319.0	3.068647	4.540120	0.0	0.000000	1.409869	4.009694	31.076575
6	376.0	4.297761	7.139118	0.0	0.000000	1.741883	4.746114	55.765629

```
In [61]: # Calculate the mean engagement rate per day of the week
mean_engagement = df_filtered.groupby('Day of the week')['Engagement Rate'].mean().reset_index()
```

```
fig = px.bar(mean_engagement,
             x="Day of the week",
             y="Engagement Rate",
             width=500,
             height=300)
```

```
fig.update_layout(
    xaxis_title="Day of the Week",
    yaxis_title="Mean Engagement Rate",
    margin={"r": 0, "l": 0, "b": 5, "t": 50})
fig.show()
```

Tuesday has the highest average engagement rate and Thursday has the highest engagement rate of 100%.

## 2-b Time

We can create a variable that put times into different time interval and investigate it further.

```
In [179]: df_filtered['time'] = df_filtered['Published Date'].apply(lambda x: x.strftime("%H:%M"))

def categorize_time(x):
    """
    Categorize time into 8 intervals of 3 hours apart
    """
    hour = int(x.split(':')[0])
    interval = hour // 3 # Divide the hour by 3 to get the interval number
    interval_start = interval * 3 # Calculate the start hour of the interval
    interval_end = interval_start + 3 # Calculate the end hour of the interval
    interval_label = f"{interval_start:02d}-{interval_end:02d}" # Format the label with leading zeros
    return interval_label
```

```
df_filtered['Time of day'] = df_filtered['time'].apply(lambda x: categorize_time(x))
```

```
Out[188]: df_filtered.groupby('Time of day')['Engagement Rate'].describe()
```

	count	mean	std	min	25%	50%	75%	max
Time of day								
00-03	2.0	0.420078	0.594080	0.0	0.210039	0.420078	0.630117	0.840156
03-06	46.0	8.201175	13.796341	0.0	1.861152	3.457819	9.278722	56.480028
06-09	377.0	6.594548	8.958138	0.0	0.000000	1.377095	7.513030	100.000000
09-12	1020.0	3.965005	5.933866	0.0	0.000000	2.056412	4.936015	44.713813
12-15	999.0	3.839246	6.224821	0.0	0.000000	1.844854	4.575111	57.889590
15-18	813.0	4.237278	6.555623	0.0	0.486454	2.033454	4.818436	60.000000
18-21	183.0	3.716742	6.064221	0.0	0.000000	1.749664	4.396422	51.509997
21-24	36.0	2.182405	4.528815	0.0	0.000000	0.386954	1.824023	21.649485

```
Out[190]: df_filtered.groupby('Time of day')['Engagement Rate'].mean().reset_index()
```

	Time of day	Engagement Rate
0	00-03	0.420078
1	03-06	8.201175
2	06-09	6.594548
3	09-12	3.965005
4	12-15	3.839246
5	15-18	4.237278
6	18-21	3.716742
7	21-24	2.182405

```
In [281]: df_filtered.groupby(['DayOfWeek', 'TimeOfDay'])['EngagementRate'].mean().reset_index()
```

	DayOfWeek	TimeOfDay	EngagementRate
0	0	00-03	NaN
1	0	03-06	NaN
2	0	06-09	3.234250
3	0	09-12	4.053950
4	0	12-15	3.948494
5	0	15-18	2.882218
6	0	18-21	2.432243
7	0	21-24	1.077597
8	1	00-03	0.840156
9	1	03-06	38.029515
10	1	06-09	4.342660
11	1	09-12	4.169044
12	1	12-15	4.731023
13	1	15-18	6.606080
14			



```
[137]: # Calculate the mean engagement rate
mean_engagement = df_filtered.groupby(['Account', 'Account Type'])['Engagement Rate'].mean().reset_index()

fig = px.bar(mean_engagement,
              x="Account Type",
              y="Engagement Rate",
              facet_col="Account",
              width=600,
              height=400)

for axis in fig.axes._yaxes():
    axis.title = None

fig.update_layout(
    yaxis_title="Mean Engagement Rate"
)

fig.show()
```

Content creators and Valorant generally post videos on Youtube platform only. The general Evil Geniuses account posts all types of social media content across all social media platform.

```
In [124]: df_filtered.groupby(['Media Type', 'Account Type'])[['Total Engagements', 'Total Impressions', 'Engagement Rate']].mean().reset_index()

Out[124]:
```

	Account	Total Engagements	Total Impressions	Engagement Rate
0	CSGO	341.985185	8570.066667	3.030547
1	Content Creators	26.169811	852.566038	4.075632
2	DOTA2	154.408468	2315.474471	4.298295
3	General	1341.732751	14817.227511	4.341212
4	Valorant	155.16667	382.983333	3.024861

The engagement rates are similar across the game titles. The general Evil Geniuses account has more engagements and impressions than other titles, followed by CSGO and DOTA2.

#### Takeaway:

General, DOTA2 and CSGO are doing fairly well. If we want to improve the overall performance of the game titles, we should focus on improving the impressions of Valorant and posts from Content Creators.

In addition, having more detailed information on each game title would give us a better perspectives of the overall social media activities. For example, if we can access data on each game title of Evil Geniuses in combination with daily increases of new users on each game platform, we can calculate the conversion rate and investigate further.

#### 4. What media type performs the best?

```
In [127]: df_filtered.groupby(['Media Type', 'Account Type'])[['Total Engagements', 'Total Impressions', 'Engagement Rate']].mean().reset_index()

Out[127]:
```

	Media Type	Account Type	Total Engagements	Total Impressions	Engagement Rate
0	Album	FBPAGE	2.000000	5.000000	10.000000
1	Carousel	INSTAGRAM	726.777778	17854.333333	3.784294
2	Link	FBPAGE	3.714286	458.285714	0.057891
3	Link	INSTAGRAM	0.000000	0.000000	0.000000
4	Link	LINKEDIN_COMPANY	64.666667	2584.000000	2.310642
5	Link	TWITTER	28.323944	3074.605634	2.056435
6	Link	YOUTUBE	0.000000	0.000000	0.000000
7	Mixed	TWITTER	2732.800000	36996.800000	10.770357
8	Photo	FBPAGE	5240.313043	28593.101449	13.647598
9	Photo	INSTAGRAM	1.538462	4753.035461	0.493752
10	Photo	LINKEDIN_COMPANY	41.538462	1960.923077	1.934351
11	Photo	TWITTER	372.761511	14950.670602	2.932845
12	Text	FBPAGE	1526.082707	9051.872180	3.927818
13	Text	INSTAGRAM	0.000000	0.000000	0.000000
14	Text	LINKEDIN_COMPANY	11.000000	1660.750000	0.613156
15	Text	TWITTER	67.973312	2975.919937	2.689368
16	Text	YOUTUBE	0.000000	0.000000	0.000000
17	Video	FBPAGE	6530.081395	39510.313953	15.059735
18	Video	INSTAGRAM	193.720000	10365.070000	1.141756
19	Video	LINKEDIN_COMPANY	35.500000	1457.000000	2.479645
20	Video	TIKTOK_BUSINESS	166.946903	3045.654867	5.316753
21	Video	TWITTER	587.867008	11607.253197	4.965840
22	Video	YOUTUBE	59.071743	1362.302857	3.832879

If we take into account of the social media platforms, we can see that Album is only posted on Facebook, Carousel is only posted on Instagram, and Mixed is only posted on Twitter.

Since album is essentially a collage of photos, we can put album into the same category as photo.

```
In [132]: # Convert 'album' to 'photo' in the 'media type' column
df_filtered['Media Type'] = df_filtered['Media Type'].replace('Album', 'Photo')

C:\Users\user\AppData\Local\Temp\ipykernel_3448\3090891946.py:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

In [133]: df_filtered.groupby(['Media Type', 'Account Type'])[['Total Engagements', 'Total Impressions', 'Engagement Rate']].mean().reset_index()

Out[133]:
```

	Media Type	Total Engagements	Total Impressions	Engagement Rate
0	Carousel	726.777778	17854.333333	3.784294
1	Link	247.010638	2473.031915	1.635631
2	Mixed	2732.800000	36996.800000	10.770357
3	Photo	1450.695506	16025.289068	4.954619
4	Text	270.672527	3413.409890	2.453918
5	Video	888.780765	10956.373320	4.930493

```
In [134]: # Calculate the mean engagement rate
mean_engagement = df_filtered.groupby(['Media Type'])['Engagement Rate'].mean().reset_index()

fig = px.bar(mean_engagement,
              x="Media Type",
              y="Engagement Rate",
              width=600,
              height=400)

fig.update_layout(
    yaxis_title="Mean Engagement Rate"
)

fig.show()
```

```
In [135]: df_filtered.groupby(['Media Type', 'Account Type'])[['Total Engagements', 'Total Impressions', 'Engagement Rate']].mean().reset_index()

Out[135]:
```

	Media Type	Account Type	Total Engagements	Total Impressions	Engagement Rate
0	Carousel	INSTAGRAM	726.777778	17854.333333	3.784294
1	Link	FBPAGE	3.714286	458.285714	0.057891
2	Link	INSTAGRAM	0.000000	0.000000	0.000000
3	Link	LINKEDIN_COMPANY	64.666667	2584.000000	2.310642
4	Link	TWITTER	28.323944	3074.605634	2.056435
5	Link	YOUTUBE	0.000000	0.000000	0.000000
6	Mixed	TWITTER	2732.800000	36996.800000	10.770357
7	Photo	FBPAGE	5180.275072	28265.444126	13.605735
8	Photo	INSTAGRAM	137.595745	4753.035461	0.459752
9	Photo	LINKEDIN_COMPANY	41.538462	1960.923077	1.934351
10	Photo	TWITTER	372.761511	14950.670602	2.932845
11	Text	FBPAGE	1526.082707	9051.872180	3.927818
12	Text	INSTAGRAM	0.000000	0.000000	0.000000
13	Text	LINKEDIN_COMPANY	11.000000	1660.750000	0.613156
14	Text	TWITTER	67.973312	2975.919937	2.689368
15	Text	YOUTUBE	0.000000	0.000000	0.000000
16	Video	FBPAGE	6530.081395	39510.313953	15.059735
17	Video	INSTAGRAM	193.720000	10365.070000	1.141756
18	Video	LINKEDIN_COMPANY	35.500000	1457.000000	2.479645
19	Video	TIKTOK_BUSINESS	166.946903	3045.654867	5.316753
20	Video	TWITTER	587.867008	11607.253197	4.965840
21	Video	YOUTUBE	59.071743	1362.302857	3.832879

```
In [138]: df_filtered.groupby(['Media Type'])['Engagement Rate'].describe().reset_index()

Out[138]:
```

	Media Type	count	mean	std	min	25%	50%	75%	max
0	Carousel	9.0	3.784294	2.389164	1.226810	2.542834	3.032566	3.902148	9.272955
1	Link	94.0	1.635631	6.637472	0.000000	0.000000	0.000000	8.010285	40.000000
2	Mixed	5.0	10.770357	5.574731	6.533617	6.605380	7.247822	15.07464	18.295905
3	Photo	1491.0	4.954619	8.262124	0.000000	0.000000	1.949306	5.604760	100.000000
4	Text	910.0	2.453918	3.942132	0.000000	0.000000	1.122761	3.476488	37.657809
5	Video	967.0	4.903493	5.818951	0.000000	1.423765	3.190552	6.241128	44.993383

If we use Engagement Rate as the metrics, Mixed media type performs the best among all media types, and Video and Photo has similar performance. In terms of total impressions, mixed media type still has the highest views, followed by Carousel and Photo.

#### Takeaway:

- Photo, Video, and Text are the most common types of posts and ads on social media platform.
- Mixed media type is only posted on Twitter, so we can infer that Twitter has a high engagement rate. Thus, posting mixed media type on Twitter is an efficient way to boost the company's social media present.
- Carousel is only posted on Instagram. Similarly, posting Carousel media type is more efficient than other media type posts on Instagram.
- Link on Youtube and Text on Instagram have little effect on social media performance.
- Link works best on LinkedIn, although the engagement rate is still fairly low.
- Photo, Video, Text works best on Facebook and has an engagement rate better than any other social media platform.
- Overall, if the company intended on increasing its engagement on social media, Facebook should be the go-to platform.

#### 5. What is our best performing campaign?

```
In [142]: df_filtered.groupby(['Campaign Name'])['Engagement Rate'].describe().reset_index()

Out[142]:
```

	Campaign Name	count	mean	std	min	25%	50%	75%	max
0	Community Engagement	1411.0	5.634547	7.575413	0.0	1.144285	2.882206	6.627286	58.858403
1	Evergreen	163.0	3.467474	2.487111	0.0	1.661541	3.090909	4.413168	13.649289
2	Evil Exhibited	420.0	2.681813	2.548381	0.0	0.854577	2.064265	3.856838	15.988244
3	N/A	1482.0	3.351382	6.777150	0.0	0.000000	0.000000	4.195538	100.000000

```
In [145]: # Calculate the mean engagement rate
mean_engagement = df_filtered.groupby(['Campaign Name'])['Engagement Rate'].mean().reset_index()

fig = px.bar(mean_engagement,
              x="Campaign Name",
              y="Engagement Rate",
              width=600,
              height=400)

fig.update_layout(
    yaxis_title="Mean Engagement Rate"
)

fig.show()
```

The best performing campaign is Community Engagement, followed by Evergreen.

```
In [146]: df_filtered.groupby(['Campaign Name', 'Account Type'])['Engagement Rate'].describe().reset_index()

Out[146]:
```

	Campaign Name	Account Type	count	mean	std	min	25%	50%	75%	max
0	Community Engagement	FBPAGE	301.0	14.905180	11.058025	0.000000	6.284133	12.971957	20.639778	58.858403
1	Community Engagement	INSTAGRAM	159.0	1.208653	2.484503	0.000000	0.000000	0.000000	1.366499	13.273076
2	Community Engagement	LINKEDIN_COMPANY	18.0	1.860454	0.779330	0.686591	1.398350	1.689915	2.311451	3.960691
3	Community Engagement	TIKTOK_BUSINESS	13.0	6.422724	2.324590	3.973510	4.683270	5.732484	7.241911	11.067708
4	Community Engagement	TWITTER	899.0	3.468913	3.333065	0.000000	1.269614	2.369668	4.625469	34.639889
5	Community Engagement	YOUTUBE	21.0	7.455909	9.598989	0.625835	2.681292	4.334365	5.555556	37.931034
6	Evergreen	FBPAGE	1.0	13.098000	NaN	13.098000	13.098000	13.098000	13.098000	13.098000
7	Evergreen	INSTAGRAM	64.0	1.744471	0.877822	1.056168	1.139780	1.513705	1.968753	5.219921
8	Evergreen	TIKTOK_BUSINESS	63.0	4.832811	2.523657	1.236089	3.067077	4.237288	6.604817	13.649289
9	Evergreen	TWITTER	1.0	11.770774	NaN	11.770774	11.770774	11.770774	11.770774	11.770774
10	Evergreen	YOUTUBE	34.0	3.653416	1.636373	0.000000	3.097628	3.727058	4.500504	8.333333
11	Evil Exhibited	FBPAGE	34.0	3.513447	2.349649	0.625835	1.778778	3.243064	4.880881	10.298713
12	Evil Exhibited	INSTAGRAM	72.0	0.815349	1.121810	0.000000	0.000000	0.000000	1.425745	7.094595
13	Evil Exhibited	LINKEDIN_COMPANY	1.0	3.213268	NaN	3.213268	3.213268	3.213268	3.213268	3.213268
14	Evil Exhibited	TIKTOK_BUSINESS	22.0	3.770242	2.327084	0.000000	1.399593	2.345297	4.032413	15.988244
15	Evil Exhibited	TWITTER	230.0	2.961700	2.427204	0.000000	1.139953	2.345297	4.032413	15.988244
16	Evil Exhibited	YOUTUBE	61.0	2.964750	3.523174	0.000000	0.000000	2.013423	5.151915	14.285714
17	N/A	FBPAGE	246.0	8.407131	13.048530	0.000000	0.000000	0.000000	15.039087	100.000000
18	N/A	INSTAGRAM	293.0	0.100795	6.620454	0.000000	0.000000	0.000000	0.000000	7.207207
19	N/A	LINKEDIN_COMPANY	3.0	0.929658	1.208698	0.000000	0.264679	0.492958	1.394486	2.290015
20	N/A	TIKTOK_BUSINESS	15.0	8.659015	4.192338	0.000000	5.803398	8.058252	11.481317	16.833700
21	N/A	TWITTER	821.0	3.078287	4.333225	0.000000	0.000000	1.862197	4.530201	60.000000
22	N/A	YOUTUBE	104.0	2.010710	2.151343	0.000000	0.000000	1.716897	3.315322	7.629108

The Community Engagement campaign is mostly advertised on Twitter and Facebook. The Evergreen campaign is mostly advertised on Instagram and TikTok. Evil Exhibited Campaign is mostly advertised Twitter, Instagram and Youtube.

#### Takeaway:

- Community Engagement is the most successful campaign with a mean engagement rate of 5.6%, in which the campaign has the highest engagement rate on Facebook even though the campaign is more widely advertised on Twitter.
- Evergreen campaign is mostly advertised on Facebook, with a mean engagement rate of 13%.
- Evil Exhibited campaign is mostly advertised on Twitter, but the engagement rate is higher on TikTok.

#### Further exploration

```
In [284]: df_filtered.groupby(['Account Type'])[['Total Engagements', 'Total Impressions', 'Engagement Rate']].mean().reset_index()

Out[284]:
```

	Account Type	Total Engagements	Total Impressions	Engagement Rate
0	FBPAGE	4420.144330	24867.426117	11.283108
1	INSTAGRAM	143.005102	6078.331633	0.666769
2	LINKEDIN_COMPANY	38.530909	1945.500000	1.795018
3	TIKTOK_BUSINESS	169.946903	3045.654867	5.316753
4	TWITTER	309.871348	9995.179908	3.248994
5	YOUTUBE	46.945455	1083.650000	3.048881

```
In [150]: # Calculate the mean engagement rate
mean_engagement = df_filtered.groupby(['Account Type'])['Engagement Rate'].mean().reset_index()

# Sort the DataFrame by 'Engagement Rate' in descending order
mean_engagement = mean_engagement.sort_values('Engagement Rate', ascending=False)

# Plot the bar graph
fig = px.bar(mean_engagement, x="Account Type", y="Engagement Rate", width=600, height=400)
fig.update_layout(yaxis_title="Mean Engagement Rate")
fig.show()
```

In addition to the questions above, it is also beneficial to look into the social media performance on each social media platform. If we use engagement rate as a metric, Facebook has the highest mean engagement rate, followed by TikTok, Twitter and Youtube. At the same time, Facebook engagement rate also has a large variance in comparison to other social media types.

## 7. What suggestions would you give to the social media team if they want to expand their presence (e.g. if our CSGO youtube channel is doing well should we expand to TikTok)?

Ultimately, we would need more data to make the decisions, such as conversion rates from different platform. But based on the data we have so far, here are some suggestions:

1. Analyze Audience and Platform Fit:
  - Research the demographics and user behavior of the target audience on TikTok. Understand if the demographics align with the brand's target audience for the CSGO YouTube channel.
  - Evaluate if the content format and style that performs well on YouTube would resonate with TikTok users. Consider the differences in content duration, editing style, and engagement patterns on TikTok.
  - Assess if the content from the CSGO YouTube channel can be repurposed or adapted to suit the TikTok platform.
2. Explore TikTok Potential:
  - Conduct a competitive analysis to see if other CSGO or gaming-related accounts are successful on TikTok. Assess the engagement levels and the types of content that perform well within the CSGO community on TikTok.
  - Experiment with creating TikTok content that aligns with the interests and preferences of the TikTok gaming community.
  - Leverage TikTok's unique features, such as music, effects, challenges, and trends, to engage with the audience and showcase the brand's personality.