SENTIMENT ANALYSIS

mood

IMPORTING LIBRARIES

```
#importing libraries and reading the files
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

READING CSV FILES

```
a=pd.read_csv(r'C:\Users\Windows\Downloads\twitter_training.csv')
b=pd.read_csv(r'C:\Users\Windows\Downloads\twitter_validation.csv')
```

Setting column names

```
column name=['tweetID','entity','sentiment','tweet content']
a.columns=column name
b.columns=column name
df=pd.concat([a,b],ignore index=False)
df.head()
   tweetID
                entity sentiment \
           Borderlands Positive
0
      2401
1
     2401 Borderlands Positive
2
      2401 Borderlands Positive
3
      2401 Borderlands Positive
     2401 Borderlands Positive
                                      tweet content
O I am coming to the borders and I will kill you...
1 im getting on borderlands and i will kill you ...
2 im coming on borderlands and i will murder you...
  im getting on borderlands 2 and i will murder ...
4 im getting into borderlands and i can murder y...
```

QUICK VIEW OF DATA

```
#list columns in our dataframe
df.columns.tolist()
['tweetID', 'entity', 'sentiment', 'tweet_content']
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 75680 entries, 0 to 998
Data columns (total 4 columns):
                   Non-Null Count Dtype
    Column
0
    tweetID
                   75680 non-null int64
1
    entity
                   75680 non-null object
    sentiment
2
                   75680 non-null
                                   object
3
    tweet content 74994 non-null object
dtypes: int64(1), object(3)
memory usage: 2.9+ MB
```

EDA

Looking for null Values

```
df.isnull().sum().sort_values(ascending = False)

tweet_content 686
tweetID 0
entity 0
sentiment 0
dtype: int64
```

Checking for dupilcates

```
df.duplicated().sum()
np.int64(3216)
```

Dropping null values and duplicated values

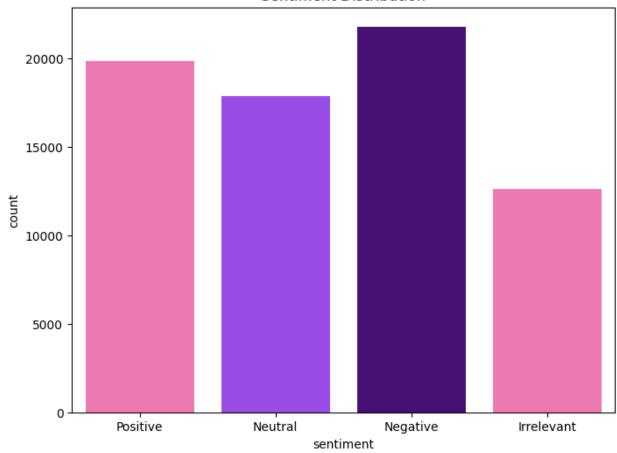
Dropping Unneeded Columns

```
df.drop(columns=['tweetID','tweet content'],inplace=True)
df.head()
       entity sentiment
O Borderlands Positive
1 Borderlands Positive
2 Borderlands Positive
3 Borderlands Positive
4 Borderlands Positive
df.info()
<class 'pandas.core.frame.DataFrame'>
Index: 72138 entries, 0 to 995
Data columns (total 2 columns):
#
    Column Non-Null Count Dtype
     -----
    entity 72138 non-null object
0
    sentiment 72138 non-null object
1
dtypes: object(2)
memory usage: 1.7+ MB
```

VISUALISATION

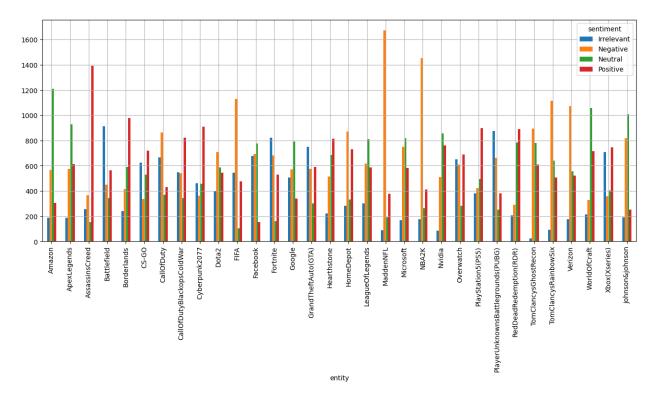
```
custom palette = sns.color palette(["#FF66B2", "#9933FF", "#4B0082"])
plt.figure(figsize=(8,6))
sns.countplot(x='sentiment', data=df, palette=custom_palette)
plt.title('Sentiment Distribution')
plt.show()
C:\Users\Windows\AppData\Local\Temp\ipykernel 11908\2557904149.py:4:
FutureWarning:
Passing `palette` without assigning `hue` is deprecated and will be
removed in v0.14.0. Assign the `x` variable to `hue` and set
`legend=False` for the same effect.
  sns.countplot(x='sentiment', data=df, palette=custom palette)
C:\Users\Windows\AppData\Local\Temp\ipykernel 11908\2557904149.py:4:
UserWarning:
The palette list has fewer values (3) than needed (4) and will cycle,
which may produce an uninterpretable plot.
  sns.countplot(x='sentiment', data=df, palette=custom palette)
```

Sentiment Distribution



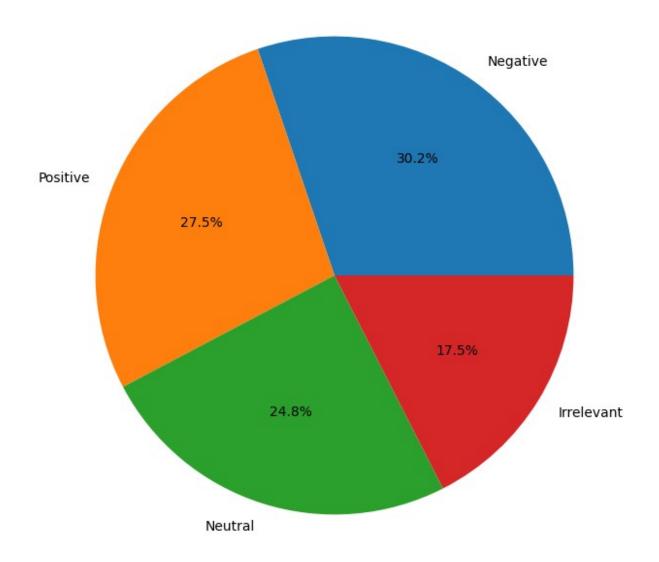
```
reactions_entities = pd.crosstab(df['entity'],df['sentiment'])
reactions_entities.plot(kind='bar', figsize=(16, 6),grid=True)

<Axes: xlabel='entity'>
```



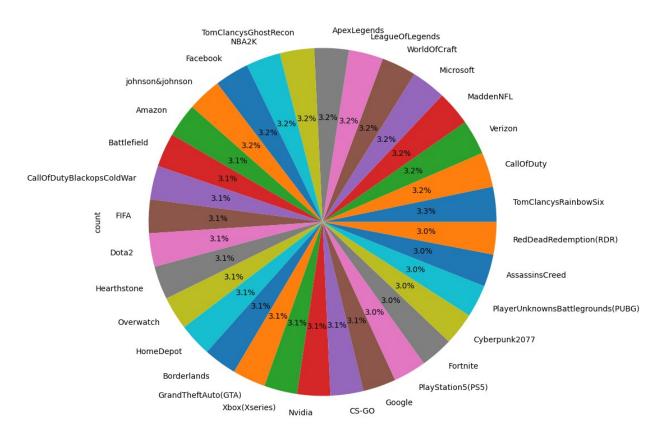
```
df['sentiment'].value_counts().plot.pie(autopct='%1.1f%%',
figsize=(8,8), title='Sentiment Proportion')
plt.ylabel('')
plt.show()
```

Sentiment Proportion



```
entity_content=df['entity'].value_counts()
entity_content.plot(kind='pie', autopct='%1.1f%%', figsize=(10, 12))
plt.title('Distribution of entities')
plt.show()
```

Distribution of entities



From the above figure the highest negative reactions in the MaddenNFL entity The highest irrelevant reactions in the Battlefield entity The highest neutral reactions in the Amazon entity The highest positive reactions in the AssassinsCreed entity

```
pivot_table = df.pivot_table(index='entity', columns='sentiment',
aggfunc='size', fill_value=0)
plt.figure(figsize=(12,8))
sns.heatmap(pivot_table, annot=True, cmap='coolwarm', linewidths=0.5)
plt.title('Heatmap of Sentiments per Entity')
plt.show()
```

	Heatmap of Sentiments per Entity				
Amazon -	1.9e+02	5.7e+02	1.2e+03	3.1e+02	
ApexLegends -	1.9e+02	5.8e+02	9.3e+02	6.1e+02	- 1600
AssassinsCreed -	2.6e+02	3.7e+02	1.6e+02	1.4e+03	
Battlefield -	9.1e+02	4.5e+02	3.4e+02	5.6e+02	
Borderlands -	2.4e+02	4.2e+02	5.9e+02	9.8e+02	
CS-GO -	6.2e+02	3.4e+02	5.3e+02	7.2e+02	- 1400
CallOfDuty -	6.7e+02	8.6e+02	3.7e+02	4.3e+02	
CallOfDutyBlackopsColdWar -	5.5e+02	5.4e+02	3.4e+02	8.2e+02	
Cyberpunk2077 -	4.6e+02	3.6e+02	4.6e+02	9.1e+02	
Dota2 -	4e+02	7.1e+02	5.9e+02	5.4e+02	- 1200
FIFA -	5.4e+02	1.1e+03	1e+02	4.8e+02	
Facebook -	6.8e+02	6.9e+02	7.8e+02	1.5e+02	
Fortnite -	8.2e+02	6.8e+02	1.6e+02	5.3e+02	1000
Google -	5.1e+02	5.7e+02	7.9e+02	3.4e+02	- 1000
GrandTheftAuto(GTA) -	7.5e+02	5.7e+02	3e+02	5.9e+02	
<u></u>	2.2e+02	5.2e+02	6.9e+02	8.1e+02	
Hearthstone - HomeDepot -	2.8e+02	8.7e+02	3.3e+02	7.3e+02	- 800
LeagueOfLegends -	3e+02	6.2e+02	8.1e+02	5.9e+02	800
MaddenNFL -	87	1.7e+03	1.9e+02	3.8e+02	
Microsoft -	1.7e+02	7.5e+02	8.2e+02	5.8e+02	
NBA2K -	1.8e+02	1.5e+03	2.6e+02	4.1e+02	- 600
Nvidia -	86	5.1e+02	8.6e+02	7.6e+02	000
Overwatch -	6.5e+02	6.1e+02	2.8e+02	6.9e+02	
PlayStation5(PS5) -	3.8e+02	4.2e+02	5e+02	9e+02	
PlayerUnknownsBattlegrounds(PUBG) -	8.8e+02	6.6e+02	2.5e+02	3.8e+02	- 400
RedDeadRedemption(RDR) -	2e+02	2.9e+02	7.9e+02	8.9e+02	
TomClancysGhostRecon -	23	8.9e+02	7.8e+02	6.1e+02	
TomClancysRainbowSix -	92	1.1e+03	6.4e+02	5.1e+02	
Verizon -	1.8e+02	1.1e+03	5.6e+02	5.2e+02	- 200
WorldOfCraft -	2.1e+02	3.3e+02	1.1e+03	7.2e+02	
Xbox(Xseries) -	7.1e+02	3.6e+02	4e+02	7.5e+02	
johnson&johnson -	1.9e+02	8.2e+02	1e+03	2.5e+02	
Irrelevant Negative Neutral Positive sentiment					

The overall sentiment analysis reveals a dominant trend where either positive or negative opinions heavily influence the public's perception. If positive, it suggests satisfaction and approval, while a high negative sentiment indicates dissatisfaction or criticism towards the entities. Neutral sentiments reflect a lack of strong emotional engagement. These patterns help gauge general audience behavior and attitudes.OVET

THANK YOU!!!!