

Perform EDA and Visualization of Twitter Sentiments in Social Media

Overview of the TWITTER Dataset

This is an entity-level sentiment analysis dataset of twitter. Given a message and an entity, the task is to judge the sentiment of the message about the entity. There are three classes in this dataset: Positive, Negative and Neutral. We regard messages that are not relevant to the entity (i.e. Irrelevant) as Neutral.

```
In [ ]: ▶ import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
In [ ]: ▶ import warnings
warnings.filterwarnings("ignore")
```

```
In [ ]: ▶ Twitter= pd.read_csv('twitter_training.csv')
```

```
In [ ]: ▶ Twitter
```

```
Out[5]:
```

	2401	Borderlands	Positive	im getting on borderlands and i will murder you all ,
0	2401	Borderlands	Positive	I am coming to the borders and I will kill you...
1	2401	Borderlands	Positive	im getting on borderlands and i will kill you ...
2	2401	Borderlands	Positive	im coming on borderlands and i will murder you...
3	2401	Borderlands	Positive	im getting on borderlands 2 and i will murder ...
4	2401	Borderlands	Positive	im getting into borderlands and i can murder y...
...
74676	9200	Nvidia	Positive	Just realized that the Windows partition of my...
74677	9200	Nvidia	Positive	Just realized that my Mac window partition is ...
74678	9200	Nvidia	Positive	Just realized the windows partition of my Mac ...
74679	9200	Nvidia	Positive	Just realized between the windows partition of...
74680	9200	Nvidia	Positive	Just like the windows partition of my Mac is l...

74681 rows × 4 columns

In []: `Twitter.head()`

Out[6]:

	2401	Borderlands	Positive	im getting on borderlands and i will murder you all ,
0	2401	Borderlands	Positive	I am coming to the borders and I will kill you...
1	2401	Borderlands	Positive	im getting on borderlands and i will kill you ...
2	2401	Borderlands	Positive	im coming on borderlands and i will murder you...
3	2401	Borderlands	Positive	im getting on borderlands 2 and i will murder ...
4	2401	Borderlands	Positive	im getting into borderlands and i can murder y...

In []: `Twitter.tail()`

Out[7]:

	2401	Borderlands	Positive	im getting on borderlands and i will murder you all ,
74676	9200	Nvidia	Positive	Just realized that the Windows partition of my...
74677	9200	Nvidia	Positive	Just realized that my Mac window partition is ...
74678	9200	Nvidia	Positive	Just realized the windows partition of my Mac ...
74679	9200	Nvidia	Positive	Just realized between the windows partition of...
74680	9200	Nvidia	Positive	Just like the windows partition of my Mac is l...

In []: `Twitter.shape`

Out[8]: (74681, 4)

In []: `Twitter.describe()`

Out[9]:

	2401
count	74681.000000
mean	6432.640149
std	3740.423819
min	1.000000
25%	3195.000000
50%	6422.000000
75%	9601.000000
max	13200.000000

In []: ▶ Twitter.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 74681 entries, 0 to 74680
Data columns (total 4 columns):
 #   Column                                Non-Null Count
---  ---
 0   2401                                74681 non-null
 1   int64
 1   Borderlands                        74681 non-null
 1   object
 2   Positive                          74681 non-null
 1   object
 3   im getting on borderlands and i will murder you all , 73995 non-null
 1   object
dtypes: int64(1), object(3)
memory usage: 2.3+ MB
```

In []: ▶ Twitter.isnull().sum()

```
Out[11]: 2401                                0
Borderlands                                0
Positive                                  0
im getting on borderlands and i will murder you all , 686
dtype: int64
```

Data Cleaning And Exploratory Data Analysis

In []: ▶ Twitter.drop(["2401"],axis = 1,inplace = True)
Twitter.head()

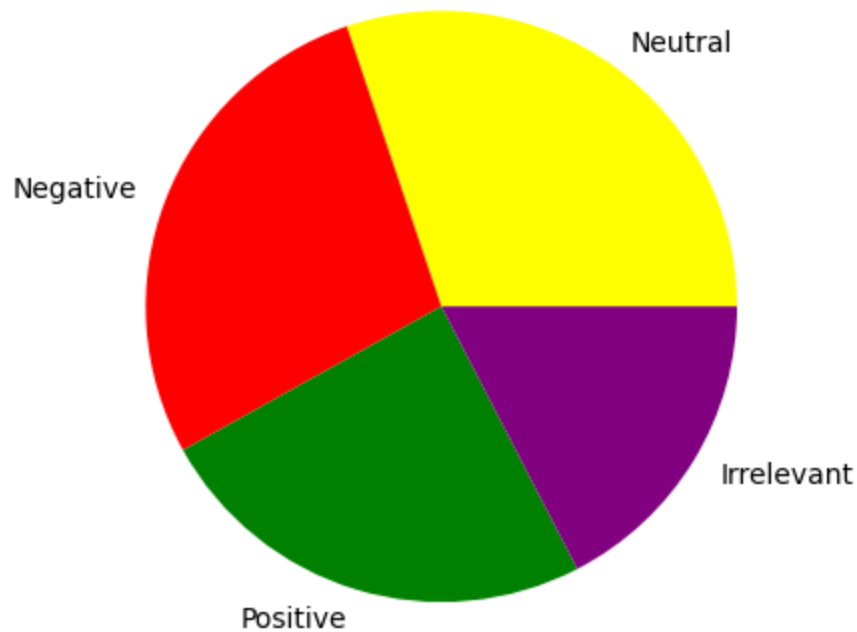
```
Out[12]:
```

	Borderlands	Positive	im getting on borderlands and i will murder you all ,
0	Borderlands	Positive	I am coming to the borders and I will kill you...
1	Borderlands	Positive	im getting on borderlands and i will kill you ...
2	Borderlands	Positive	im coming on borderlands and i will murder you...
3	Borderlands	Positive	im getting on borderlands 2 and i will murder ...
4	Borderlands	Positive	im getting into borderlands and i can murder y...

In []: ▶ Twitter['Positive'].value_counts()

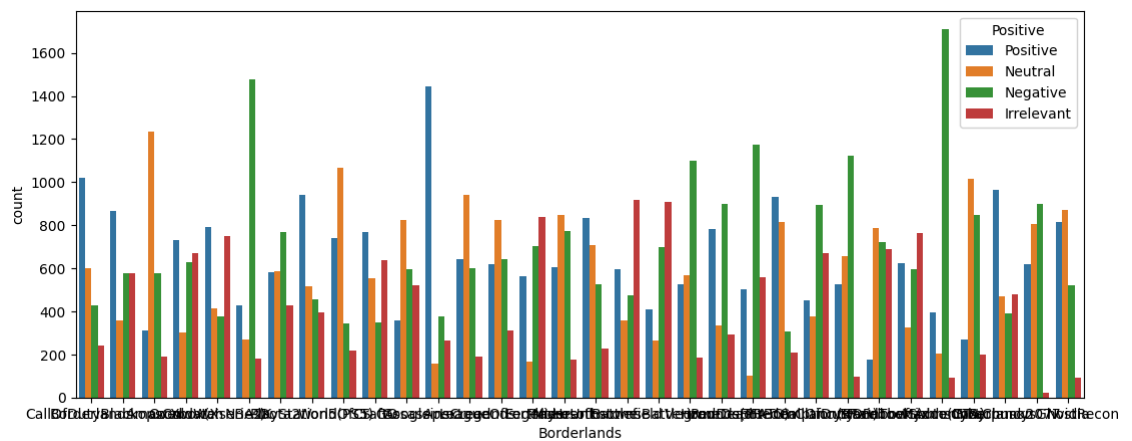
```
Out[25]: Negative      22542
Positive      20831
Neutral       18318
Irrelevant    12990
Name: Positive, dtype: int64
```

```
In [ ]: ▶ plt.pie(Twitter['Positive'].value_counts(), labels=['Neutral', 'Negative', 'Positive', 'Irrelevant'],
plt.show())
```



```
In [ ]: ▶ plt.figure(figsize = (13,5))
sns.countplot(x = "Borderlands",hue = "Positive",data = Twitter,palette = 'magma')
```

Out[29]: <Axes: xlabel='Borderlands', ylabel='count'>



```
In [ ]: ▶
```