```python
import pandas as pd
df=pd.read_csv('/content/twitter_validation.csv',header=None)
df
```

	0	1	2	3
0	3364	Facebook	Irrelevant	I mentioned on Facebook that I was struggling ...
1	352	Amazon	Neutral	BBC News - Amazon boss Jeff Bezos rejects clai...
2	8312	Microsoft	Negative	@Microsoft Why do I pay for WORD when it funct...
3	4371	CS-GO	Negative	CSGO matchmaking is so full of closet hacking,...
4	4433	Google	Neutral	Now the President is slapping Americans in the...
...	...	...	...	...
995	4891	GrandTheftAuto(GTA)	Irrelevant	⭐ Toronto is the arts and culture capital of ...
				tHIS IS ACTUALLY A GOOD MOVE

```python
df.columns=['id','location','target','text']
df
```

	id	location	target	text
0	3364	Facebook	Irrelevant	I mentioned on Facebook that I was struggling ...
1	352	Amazon	Neutral	BBC News - Amazon boss Jeff Bezos rejects clai...
2	8312	Microsoft	Negative	@Microsoft Why do I pay for WORD when it funct...
3	4371	CS-GO	Negative	CSGO matchmaking is so full of closet hacking,...
4	4433	Google	Neutral	Now the President is slapping Americans in the...
...	...	...	...	...
995	4891	GrandTheftAuto(GTA)	Irrelevant	⭐ Toronto is the arts and culture capital of ...
				tHIS IS ACTUALLY A GOOD MOVE

```python
df.isna().sum()
```

```
id          0
location    0
target      0
text        0
dtype: int64
```

```python
df['target'].value_counts()
```
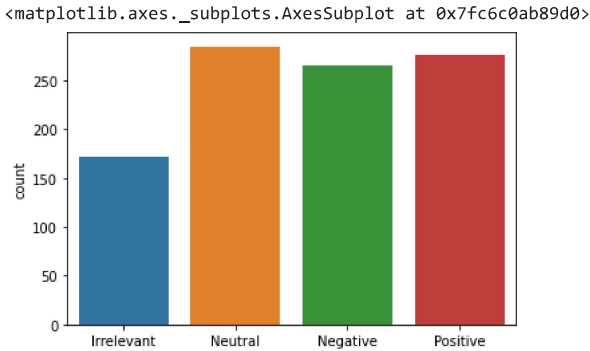
```
Neutral       285
Positive      277
Negative      266
Irrelevant    172
Name: target, dtype: int64
```

```python
df['target'].unique()
```

```
array(['Irrelevant', 'Neutral', 'Negative', 'Positive'], dtype=object)
```

```python
#plot
import seaborn as sns
sns.countplot(x="target",data=df)
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7fc6c0ab89d0>
```



```
#
df.drop(df.index[(df['target']=='Irrelevant')],axis=0,inplace=True)
df
```

	id	location	target	text
1	352	Amazon	Neutral	BBC News - Amazon boss Jeff Bezos rejects clai...
2	8312	Microsoft	Negative	@Microsoft Why do I pay for WORD when it funct...
3	4371	CS-GO	Negative	CSGO matchmaking is so full of closet hacking,...
4	4433	Google	Neutral	Now the President is slapping Americans in the...
5	6273	FIFA	Negative	Hi @EAHelp I've had Madeleine McCann in my cel...
...	...	...	...	...
993	314	Amazon	Negative	Please explain how this is possible! How can t...
				Good on Sony. As much as I want to see

```
df.reset_index(drop=True,inplace=True)
df
```

	id	location	target	text
0	352	Amazon	Neutral	BBC News - Amazon boss Jeff Bezos rejects clai...
1	8312	Microsoft	Negative	@Microsoft Why do I pay for WORD when it funct...
2	4371	CS-GO	Negative	CSGO matchmaking is so full of closet hacking,...
3	4433	Google	Neutral	Now the President is slapping Americans in the...
4	6273	FIFA	Negative	Hi @EAHelp I've had Madeleine McCann in my cel...
...	...	...	...	...
823	314	Amazon	Negative	Please explain how this is possible! How can t...
				Good on Sony. As much as I want to see

```
df.drop(['id','location'],axis=1,inplace=True)

df
```

	target	text
0	Neutral	BBC News - Amazon boss Jeff Bezos rejects clai...
1	Negative	@Microsoft Why do I pay for WORD when it funct...
2	Negative	CSGO matchmaking is so full of closet hacking,...
3	Neutral	Now the President is slapping Americans in the...
4	Negative	Hi @EAHelp I've had Madeleine McCann in my cel...
...	...	...
823	Negative	Please explain how this is possible! How can t...
824	Positive	Good on Sony. As much as I want to see the new...

```python
df["target"]=df['target'].map({'Positive':1,'Negative':-1,'Neutral':0})
```

```python
df
```

	target	text
0	0	BBC News - Amazon boss Jeff Bezos rejects clai...
1	-1	@Microsoft Why do I pay for WORD when it funct...
2	-1	CSGO matchmaking is so full of closet hacking,...
3	0	Now the President is slapping Americans in the...
4	-1	Hi @EAHelp I've had Madeleine McCann in my cel...
...	...	...
823	-1	Please explain how this is possible! How can t...
824	1	Good on Sony. As much as I want to see the new...
825	1	Today sucked so it's time to drink wine n play...
826	1	Bought a fraction of Microsoft today. Small wins.
827	0	Johnson & Johnson to stop selling talc baby po...

```python
tweets=df.text
tweets
```

```
0      BBC News - Amazon boss Jeff Bezos rejects clai...
1      @Microsoft Why do I pay for WORD when it funct...
2      CSGO matchmaking is so full of closet hacking,...
3      Now the President is slapping Americans in the...
4      Hi @EAHelp I've had Madeleine McCann in my cel...
                             ...
823    Please explain how this is possible! How can t...
824    Good on Sony. As much as I want to see the new...
825    Today sucked so it's time to drink wine n play...
826    Bought a fraction of Microsoft today. Small wins.
827    Johnson & Johnson to stop selling talc baby po...
Name: text, Length: 828, dtype: object
```

```python
#to remove specialchara
tweets=tweets.str.replace('[^a-zA-Z0-9]+',' ')# ^ is used to say that except that and+ is take everyt
tweets
```

```
<ipython-input-60-f429aa0fdde0>:2: FutureWarning: The default value of regex will change from True to False in a future version.
  tweets=tweets.str.replace('[^a-zA-Z0-9]+',' ')# ^ is used to say that except that and+ is take everything together
0      BBC News Amazon boss Jeff Bezos rejects claims...
1       Microsoft Why do I pay for WORD when it funct...
2      CSGO matchmaking is so full of closet hacking ...
3      Now the President is slapping Americans in the...
4      Hi EAHelp I ve had Madeleine McCann in my cell...
                             ...
823    Please explain how this is possible How can th...
824    Good on Sony As much as I want to see the new ...
825    Today sucked so it s time to drink wine n play...
826     Bought a fraction of Microsoft today Small wins
827    Johnson Johnson to stop selling talc baby powd...
Name: text, Length: 828, dtype: object
```

```python
from nltk.stem import SnowballStemmer# to remove tail
stemmer=SnowballStemmer('english')
```

```
from nltk import TweetTokenizer# one word tokenize
tk=TweetTokenizer()
tweets=tweets.apply(lambda x:[stemmer.stem(i.lower()) for i in tk.tokenize(x)]).apply(lambda x:' '.jc
tweets
```

```
0      bbc news amazon boss jeff bezo reject claim co...
1      microsoft whi do i pay for word when it functi...
2      csgo matchmak is so full of closet hack it s a...
3      now the presid is slap american in the face th...
4      hi eahelp i ve had madelein mccann in my cella...
                             ...
823    pleas explain how this is possibl how can they...
824    good on soni as much as i want to see the new ...
825    today suck so it s time to drink wine n play b...
826       bought a fraction of microsoft today small win
827    johnson johnson to stop sell talc babi powder ...
Name: text, Length: 828, dtype: object
```

Double-click (or enter) to edit

Colab paid products - Cancel contracts here