Data Collection

In [1]:
```
import pandas as pd
import matplotlib.pyplot as plt
```

In [2]:
```
#imported the csv file
df = pd.read_csv("C:/Users/Trupti/Downloads/sentimentdataset.csv")
df.head()
```

Out[2]:

	Unnamed: 0.1	Unnamed: 0	Text	Sentiment	Timestamp	User	Platform	Hashtags	Retweets
0	0	0	Enjoying a beautiful day at the park! ...	Positive	2023-01-15 12:30:00	User123	Twitter	#Nature #Park	15.0
1	1	1	Traffic was terrible this morning. ...	Negative	2023-01-15 08:45:00	CommuterX	Twitter	#Traffic #Morning	5.0
2	2	2	Just finished an amazing workout! 💪 ...	Positive	2023-01-15 15:45:00	FitnessFan	Instagram	#Fitness #Workout	20.0
3	3	3	Excited about the upcoming weekend getaway! ...	Positive	2023-01-15 18:20:00	AdventureX	Facebook	#Travel #Adventure	8.0
4	4	4	Trying out a new recipe for dinner tonight. ...	Neutral	2023-01-15 19:55:00	ChefCook	Instagram	#Cooking #Food	12.0

In [5]:
```
#shape is used for getting rows and columns present in dataset,here 732 rows ar pres
df.shape
```

Out[5]:
```
(732, 15)
```

In [6]:
```
#head method is used to get specific count of rows
df.head(1)
```

Out[6]:

	Unnamed: 0.1	Unnamed: 0	Text	Sentiment	Timestamp	User	Platform	Hashtags	Retweets	Likes
0	0	0	Enjoying a beautiful day at the park! ...	Positive	2023-01-15 12:30:00	User123	Twitter	#Nature #Park	15.0	30.0

In [3]:
```
print(df['User']) #This is giving the user's name
```

```
0                                User123
1                                CommuterX
2                                FitnessFan
3                                AdventureX
4                                ChefCook
                        ...
727          ScienceProjectSuccessHighSchool
728               BirthdayPartyJoyHighSchool
729      CharityFundraisingTriumphHighSchool
730       MulticulturalFestivalJoyHighSchool
731       VirtualTalentShowSuccessHighSchool
Name: User, Length: 732, dtype: object
```

In [8]:
```
df.info
```

Out[8]:
```
<bound method DataFrame.info of      Unnamed: 0.1  Unnamed: 0  \
0               0           0
1               1           1
2               2           2
3               3           3
4               4           4
..            ...         ...
727           728         732
728           729         733
729           730         734
730           731         735
731           732         736

                                                  Text    Sentiment  \
0     Enjoying a beautiful day at the park!        ...   Positive
1     Traffic was terrible this morning.           ...   Negative
2     Just finished an amazing workout! 💪           ...   Positive
3     Excited about the upcoming weekend getaway!  ...   Positive
4     Trying out a new recipe for dinner tonight.  ...   Neutral
..                                                 ...         ...
727  Collaborating on a science project that receiv...       Happy
728  Attending a surprise birthday party organized ...       Happy
729  Successfully fundraising for a school charity ...       Happy
730  Participating in a multicultural festival, cel...       Happy
731  Organizing a virtual talent show during challe...       Happy

                Timestamp                                    User     Platform  \
0     2023-01-15 12:30:00                          User123          Twitter
1     2023-01-15 08:45:00                          CommuterX        Twitter
2     2023-01-15 15:45:00                          FitnessFan      Instagram
3     2023-01-15 18:20:00                          AdventureX       Facebook
4     2023-01-15 19:55:00                          ChefCook        Instagram
..                    ...                                 ...         ...
727   2017-08-18 18:20:00       ScienceProjectSuccessHighSchool     Facebook
728   2018-06-22 14:15:00            BirthdayPartyJoyHighSchool    Instagram
729   2019-04-05 17:30:00    CharityFundraisingTriumphHighSchool      Twitter
730   2020-02-29 20:45:00     MulticulturalFestivalJoyHighSchool     Facebook
731   2020-11-15 15:15:00     VirtualTalentShowSuccessHighSchool    Instagram

                                          Hashtags  Retweets  Likes  \
0        #Nature #Park                                  15.0   30.0
1        #Traffic #Morning                               5.0   10.0
2        #Fitness #Workout                              20.0   40.0
```

```
3        #Travel #Adventure                                  8.0   15.0
4        #Cooking #Food                                     12.0   25.0
..                                                   ...     ...    ...
727        #ScienceFairWinner #HighSchoolScience           20.0   39.0
728    #SurpriseCelebration #HighSchoolFriendship          25.0   48.0
729      #CommunityGiving #HighSchoolPhilanthropy          22.0   42.0
730        #CulturalCelebration #HighSchoolUnity           21.0   43.0
731   #VirtualEntertainment #HighSchoolPositivity          24.0   47.0

          Country  Year  Month  Day  Hour
0       USA        2023      1   15    12
1       Canada     2023      1   15     8
2     USA          2023      1   15    15
3       UK         2023      1   15    18
4      Australia   2023      1   15    19
..             ...   ...    ...  ...   ...
727            UK  2017      8   18    18
728           USA  2018      6   22    14
729        Canada  2019      4    5    17
730            UK  2020      2   29    20
731           USA  2020     11   15    15

[732 rows x 15 columns]>
```

In [9]:
```
df.head(1)
```

Out[9]:

	Unnamed: 0.1	Unnamed: 0	Text	Sentiment	Timestamp	User	Platform	Hashtags	Retweets	Likes
0	0	0	Enjoying a beautiful day at the park! ...	Positive	2023-01-15 12:30:00	User123	Twitter	#Nature #Park	15.0	30.0

In [10]:
```
#Giving names to the columns
df.columns=["NotReQ","Id","Text","Sentiment","Timestamp","User","Platform","Hashtags
            "Month","Day","Hour"]
```

In [11]:
```
df.head(1)
```

Out[11]:

	NotReQ	Id	Text	Sentiment	Timestamp	User	Platform	Hashtags	Retweets	Likes	Country
0	0	0	Enjoying a beautiful day at the park! ...	Positive	2023-01-15 12:30:00	User123	Twitter	#Nature #Park	15.0	30.0	USA

Data Cleaning

In [12]:
```
#Droping columns which are not required
df.drop(columns='NotReQ',inplace=True)
```

In [13]:
```
df.head(1)
```

Out[13]:

	Id	Text	Sentiment	Timestamp	User	Platform	Hashtags	Retweets	Likes	Country	Year	M
0	0	Enjoying a beautiful day at the park! ...	Positive	2023-01-15 12:30:00	User123	Twitter	#Nature #Park	15.0	30.0	USA	2023	

In [14]:
```
#dtypes gives datatype of each column
df.dtypes
```

Out[14]:
```
Id             int64
Text          object
Sentiment     object
Timestamp     object
User          object
Platform      object
Hashtags      object
Retweets     float64
Likes        float64
Country       object
Year           int64
Month          int64
Day            int64
Hour           int64
dtype: object
```

In [15]:
```
#converting the datatype of timestamp object to datetime.
df['Timestamp'] = pd.to_datetime(df['Timestamp'])
df['Day'] = df['Timestamp'].dt.day
df['Month'] = df['Timestamp'].dt.month
df['Year'] = df['Timestamp'].dt.year
```

In [16]:
```
df.head(1)
```

Out[16]:

	Id	Text	Sentiment	Timestamp	User	Platform	Hashtags	Retweets	Likes	Country	Year	M
0	0	Enjoying a beautiful day at the park! ...	Positive	2023-01-15 12:30:00	User123	Twitter	#Nature #Park	15.0	30.0	USA	2023	

In [17]:
```
#strip() is used to remove the whitespaces
df['Text']= df['Text'].str.strip()
df['Sentiment']= df['Sentiment'].str.strip()
df['User']= df['User'].str.strip()
df['Platform']= df['Platform'].str.strip()
df['Hashtags']= df['Hashtags'].str.strip()
df['Country']= df['Country'].str.strip()
```

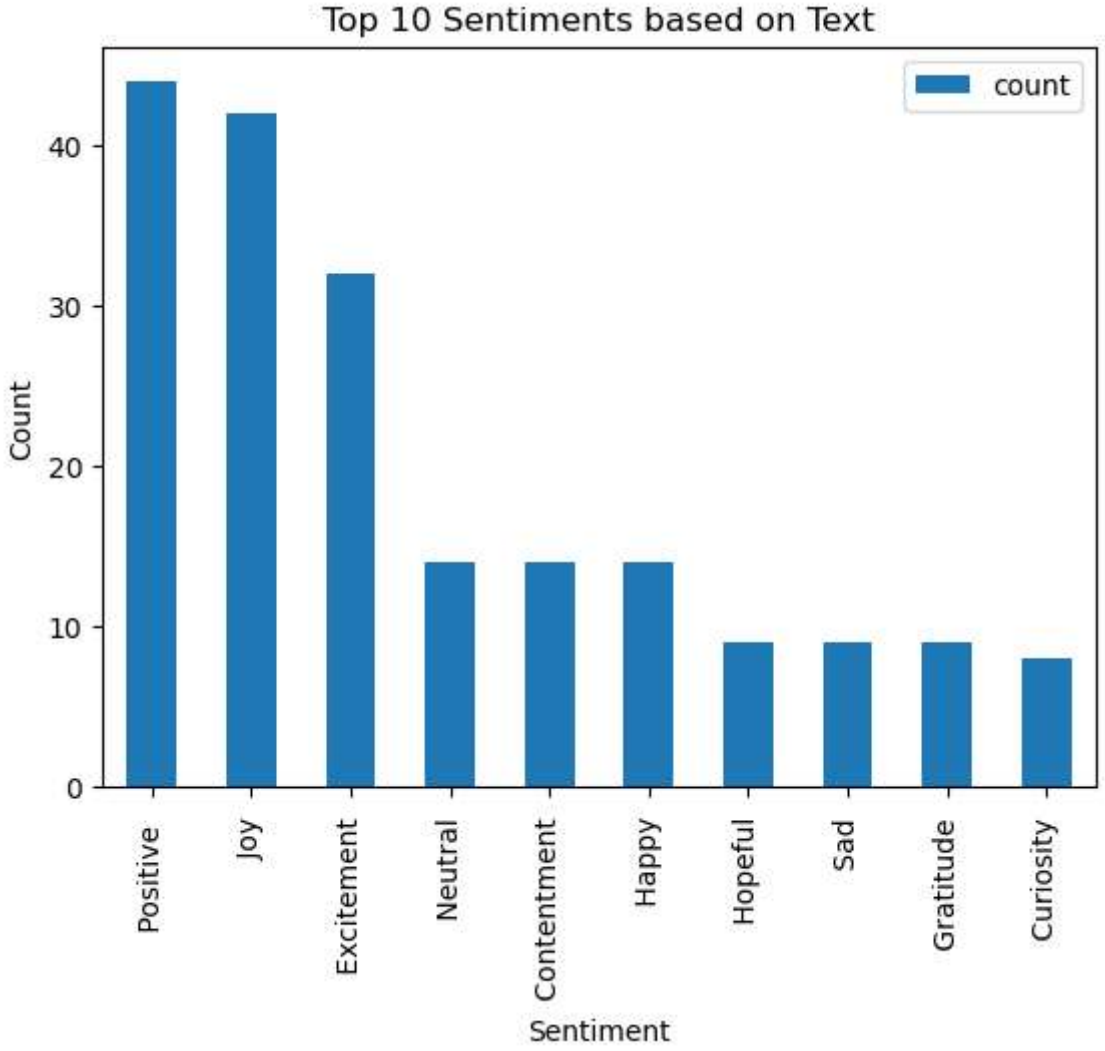
In [18]:
```
df.head(1)
```

Out[18]:

	Id	Text	Sentiment	Timestamp	User	Platform	Hashtags	Retweets	Likes	Country	Year	M
0	0	Enjoying a beautiful day at the park!	Positive	2023-01-15 12:30:00	User123	Twitter	#Nature #Park	15.0	30.0	USA	2023	

Data Visualization

In [18]:
```
df['Sentiment'].value_counts().nlargest(10).plot(kind='bar')
plt.title('Top 10 Sentiments based on Text')
plt.xlabel('Sentiment')
plt.ylabel('Count')
plt.legend()
plt.show()
```



In [20]:
```
"""
Conclusion:

Bar chart indicates that Positive sentiment is having maximum count and after that j
having maximum count in dataset.
Nostalgia sentiment is occuring least times.
Neutral and Gratitudr sentiment is having same count.
"""
```

Out[20]:
```
'\nBar chart indicates that Positive sentiment is having maximum count and after th
at joy sentiment is \nhaving maximum count in dataset.\nNostalgia sentiment is occu
ring least times.\nNeutral and Gratitudr sentiment is having same count.\n'
```

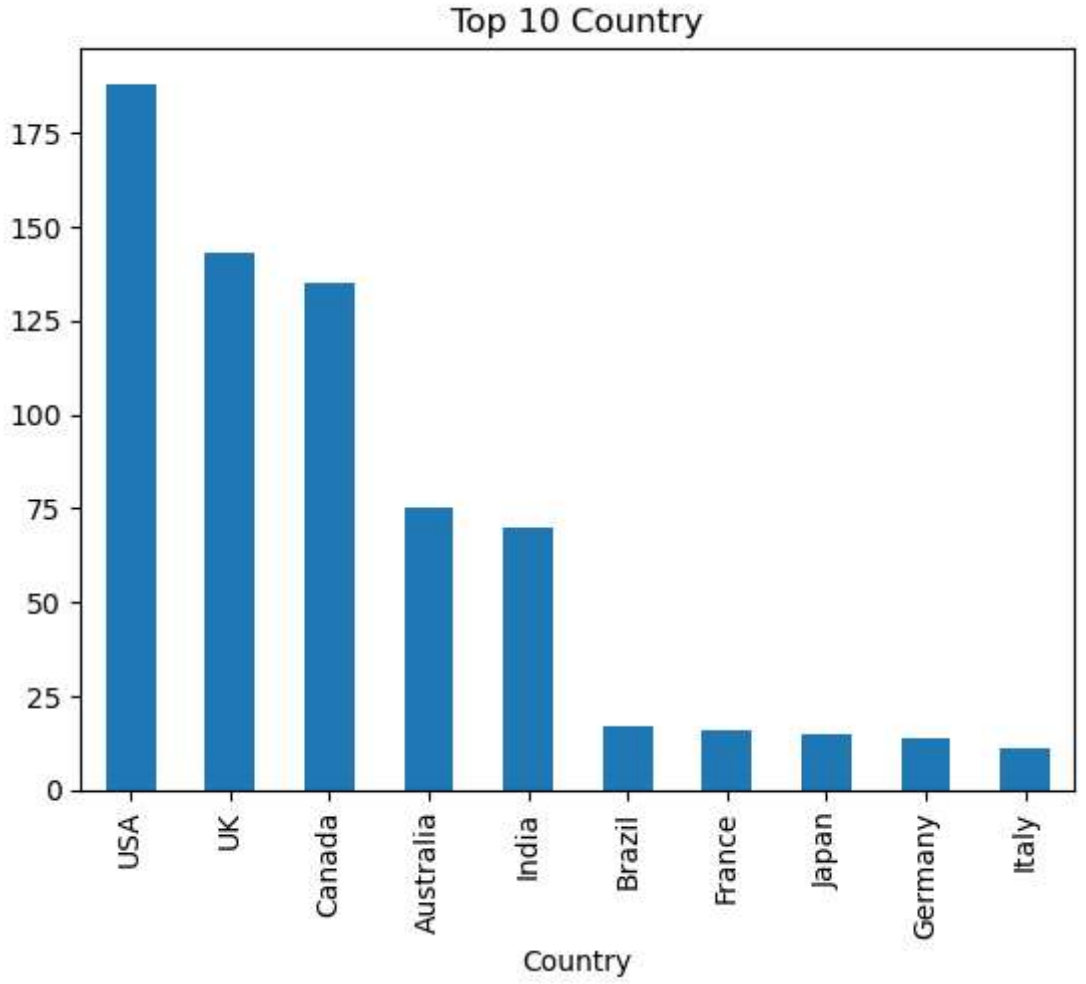
In [21]:
```
df['Platform'].value_counts()
```

Out[21]:
```
Platform
Instagram    258
Twitter      243
Facebook     231
Name: count, dtype: int64
```

In [22]:
```
"""
By value_counts(),I have analysed that instagram has occured 258 times,twitter 248 t
People shared their sentiments max times on Instagram.
"""
```

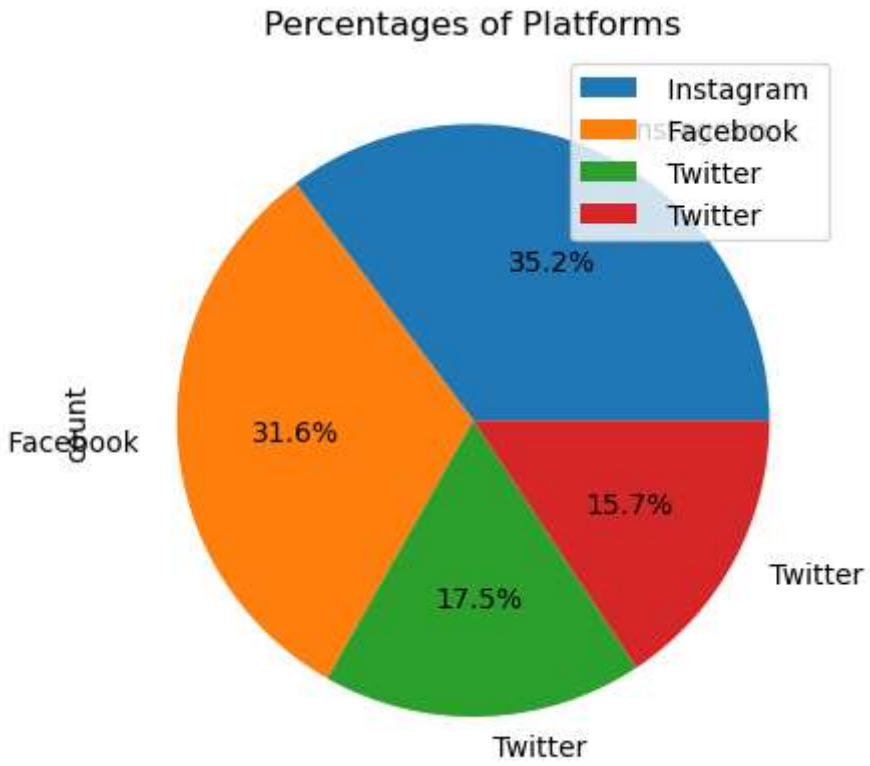
Out[22]:
```
'\nvalue_counts() gives a series containing counts of unique values.\nInstagram pla
tform is occuring 258 times in dataset means people shared their sentiments max tim
es on Instagram.\nThen they put their emotions 248 times on twitter and on facebook
231 times.\n'
```

In [23]:
```
df['Country'].value_counts().nlargest(10).plot(kind='bar')
plt.title('Top 10 Country')
plt.show()
```



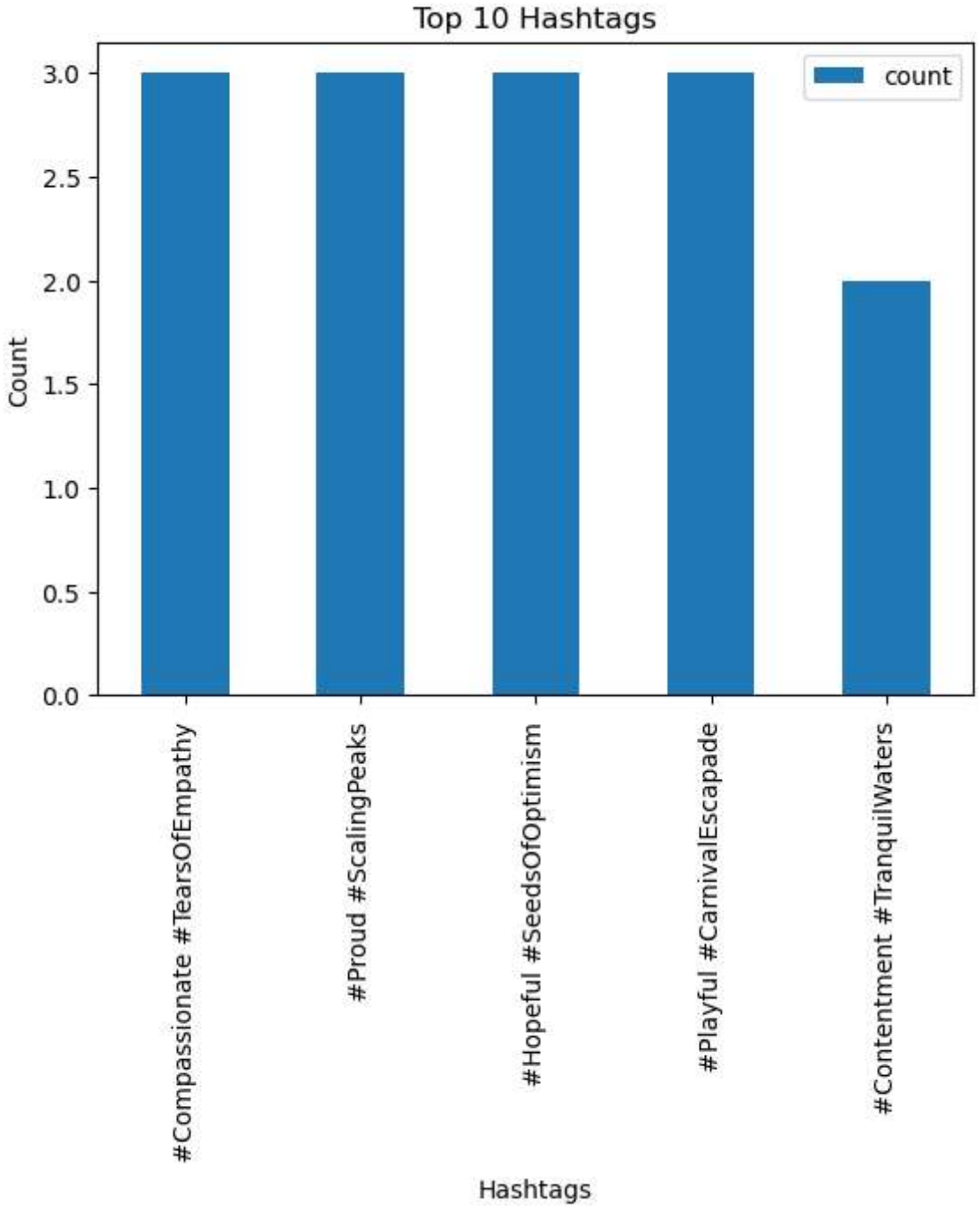
In [ ]:
```
"""
Conclusion:
USA is one of the country top most country how is using platform for sharing sentime
Italy is on the last position.
"""
```

In [5]:
```
df['Platform'].value_counts().plot(kind='pie',autopct='%1.1f%%')
plt.title('Percentages of Platforms')
plt.legend()
plt.show()
```



In [ ]:
```
"""
Conclusion:
By this pie chart we can conclude that the Instagram is highest used platform then f
"""
```

In [17]:
```
df['Hashtags'].value_counts().nlargest(10).plot(kind='bar')
plt.title('Top 10 Hashtags')
plt.xlabel('Hashtags')
plt.ylabel('Count')
plt.legend()
plt.show()
```

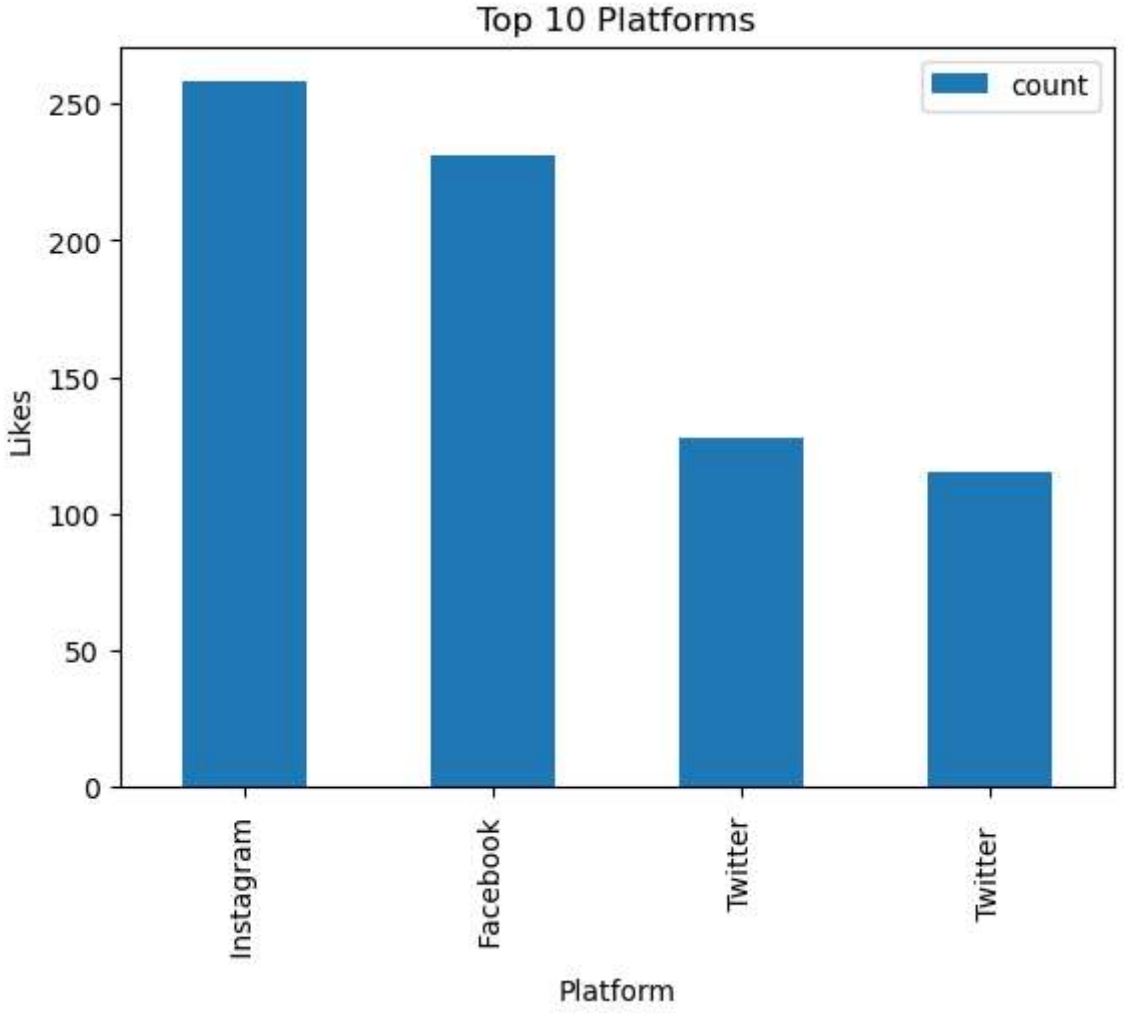

Top 10 Hashtags
count
Count
3.0
2.5
2.0
1.5
1.0
0.5
0.0
#Compassionate #TearsOfEmpathy
#Proud #ScalingPeaks
#Hopeful #SeedsOfOptimism
#Playful #CarnivalEscapade
#Contentment #TranquilWaters
Hashtags

In [19]:
```
df.head(1)
```

Out[19]:

	Unnamed: 0.1	Unnamed: 0	Text	Sentiment	Timestamp	User	Platform	Hashtags	Retweets	Likes
0	0	0	Enjoying a beautiful day at the park! ...	Positive	2023-01-15 12:30:00	User123	Twitter	#Nature #Park	15.0	30.0

In [20]:
```
df['Platform'].value_counts().nlargest(10).plot(kind='bar')
plt.title('Top 10 Platforms')
plt.xlabel('Platform')
plt.ylabel('Likes')
plt.legend()
plt.show()
```



In [ ]:
```
"""
Conclusion: By this graph we can conclude that instagram has maximum number of likes
"""
```

In [ ]:
```

```