

# A Twitter Trend Analyzer With AI - Analyse Bollywood Trends



# By-

# SHIVAM KUMAR GIRI



## Python Text Processing

## PROBLEM STATEMENT

1) Submit your Python code to analyze twitter trends. Your code will cover:

a) Generate a word cloud based on the tweets

b) Show the relative popularity of the hashtags

c) Show which celebrity had the maximum share of voice. An indicative list of celebrities you can run this analysis for: Alia Bhatt, Shah Rukh Khan, Salman Khan, Aamir Khan, Anushka Sharma, Ranbir Kapoor, Hrithik, Deepika Padukone

2) Submit a detailed report explaining the steps you followed and clearly showing your outputs including all the graphs and charts

## ABSTRACT

Twitter is a huge social media platform. More than that, tweets allow us to gain insights into the online public behaviour. As such, analysing Twitter has become a crucial source of information for brands and agencies

## LIBRARIES LANGUAGE USED

<b>pandas</b>	For Create Data Frames
<b>re</b>	For regular Expressions
<b>matplotlib.pyplot</b>	For plotting Graaphs
<b>wordcloud</b>	For word-cloud Access

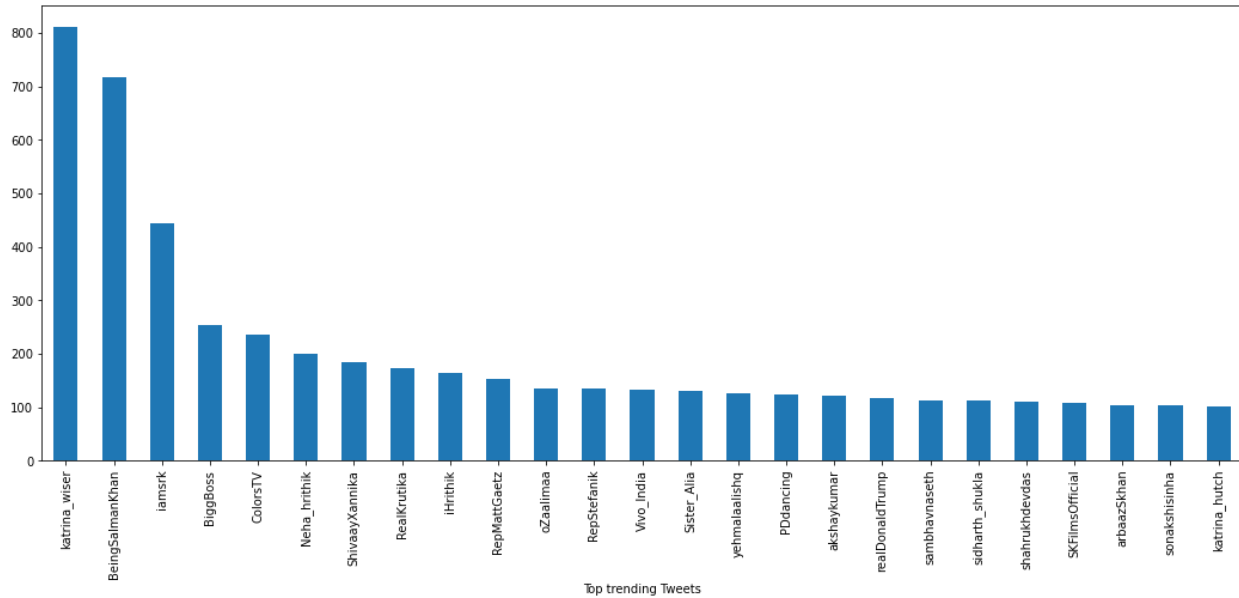
### Fetches the hastags from the file.

***Step 5: Display the count in decending (popular to unpopular)***

```
katrina_wiser    811
BeingSalmanKhan  717
iamsrk           444
BiggBoss         253
ColorsTV         237
Neha_hrithik     200
ShivaayXannika   184
RealKrutika      173
iHrithik         165
RepMattGaetz     154
oZaalimaa        136
RepStefanik      135
Vivo_India       133
Sister_Alia      131
yehmalaalishq    126
PDdancing        125
akshaykumar      122
realDonaldTrump  118
sambhavnaseth    113
sidharth_shukla  112
shahrukhdevdas   111
SKFilmsOfficial  109
arbaazSkhan      104
sonakshisinha    104
katrina_hutch    102
Name: tags, dtype: int64
```

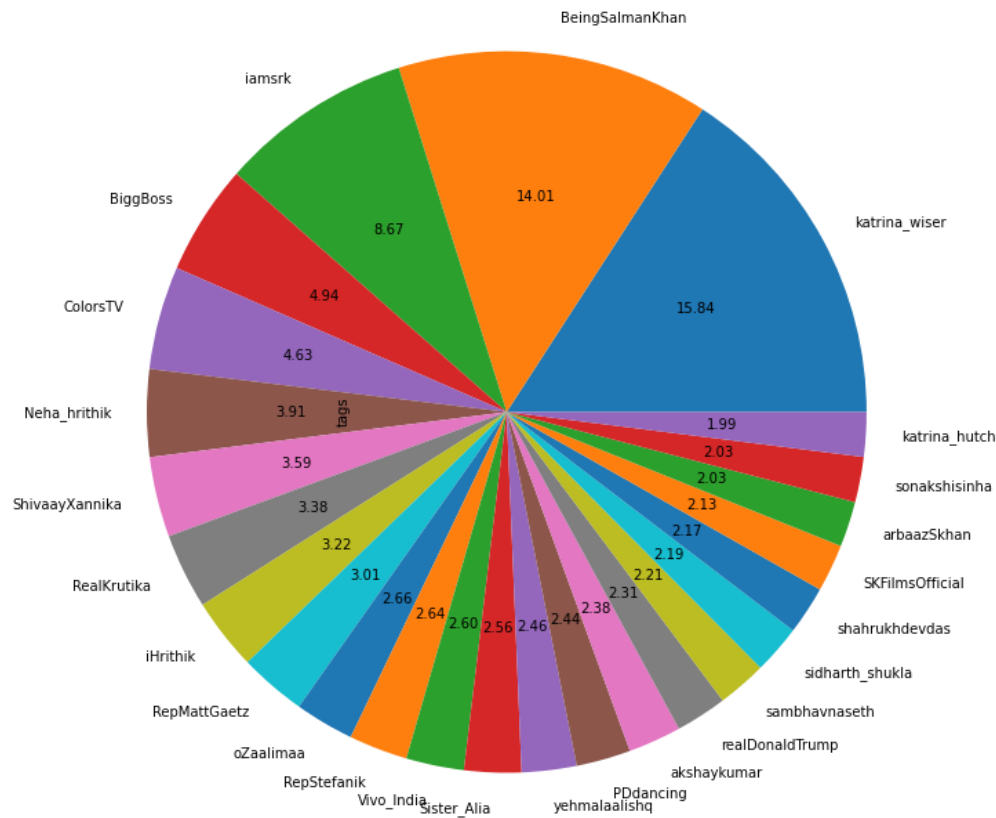
***Step 6: Creating a bar Graph on relative popularity of Tags***

**Create a bar graph with matplotlib**



## Step 7: Creating a pie Chart on relative popularity of Tags

### Creating a Pie Chart with matplotlib Library



### Step 8: Optimize the Trends as per Bollywood Actors/Actress

	tags	number	trend
0	katrina_wiser	811	Katrina Kaif
1	BeingSalmanKhan	717	Salmaan Khan
2	iamsrk	444	Shahrukh Khan
3	BiggBoss	253	Salmaan Khan
4	ColorsTV	237	Others
5	Neha_hrithik	200	Neha Hritik
6	ShivaayXannika	184	Others
7	RealKrutika	173	Kritika Yadadv
8	iHrithik	165	Hritik Roshan
9	RepMattGaetz	154	Others
10	oZaalimaa	136	Shahrukh Khan
11	RepStefanik	135	Others
12	Vivo_India	133	Others
13	Sister_Alia	131	Alia Bhatt
14	yehmalaalishq	126	Others
15	PDdancing	125	Others
16	akshaykumar	122	Akshay Kumar
17	realDonaldTrump	118	Others
18	sambhavnaseth	113	Sambhavana Seth
19	sidharth_shukla	112	Sidharth Shukla
20	shahrukhdevdas	111	Shahrukh Khan
21	SKFilmsOfficial	109	Salmaan Khan
22	arbaazSkhan	104	Arbaaz Khan
23	sonakshisinha	104	Sonakshi Sinha
24	katrina_hutch	102	Katrina kaif

### Step 9: Group the Trends as per Bollywood Actors/Actress

Used the Grouping of the dataframe from the pandas and sum operation to aggregate the sums

trend	
Akshay Kumar	122
Alia Bhatt	131
Arbaaz Khan	104
Hritik Roshan	165
Katrina Kaif	811
Katrina kaif	102
Kritika Yadadv	173
Neha Hritik	200
Others	1212
Salmaan Khan	1079
Sambhavana Seth	113
Shahrukh Khan	691
Sidharth Shukla	112
Sonakshi Sinha	104

Actor	Percentage
Katrina Kaif	15.84
Hritik Roshan	3.22
Arbaaz Khan	2.03
Alia Bhatt	2.56
Akshay Kumar	2.38
Sonakshi Sinha	2.03
Sidharth Shukla	2.19
Shahrukh Khan	13.50
Sambhavana Seth	2.21
Salmaan Khan	21.08
Others	23.68
Neha Hritik	3.91
Kritika Yadav	3.38

Salman Khan shared the maximum popularity over the other Indian Bollywood celebrities as he shares the 21.08% Shares in the twitter. This Includes bigboss and Dabangg Tweets too which are created immensely in twitter. Hence, he topped the list followed by Katrina Kaif with 15.84% of the popularity shares followed by Shahrukh Khan 13.50% . Here is List of Popularity

- Xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx**End**xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx