CS5540

PRINCIPLES OF BIG DATA MANAGEMENT

(FALL 2019)



BY

Chakradhar Chinnam (16291617)

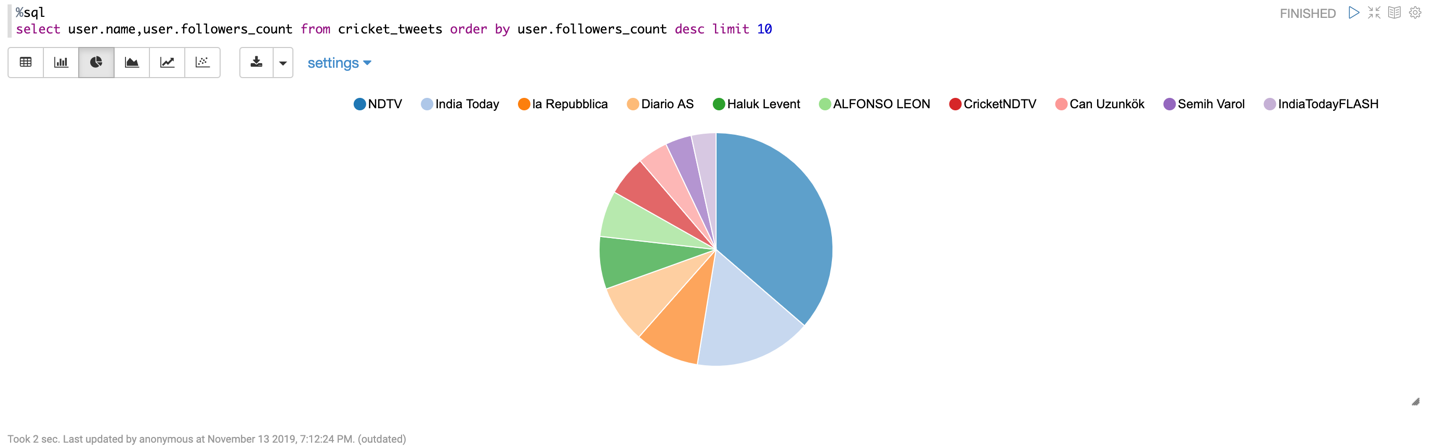
Sai Aravind Reddy Mannem (16292968)

Bommireddy Pravallika (16292279)

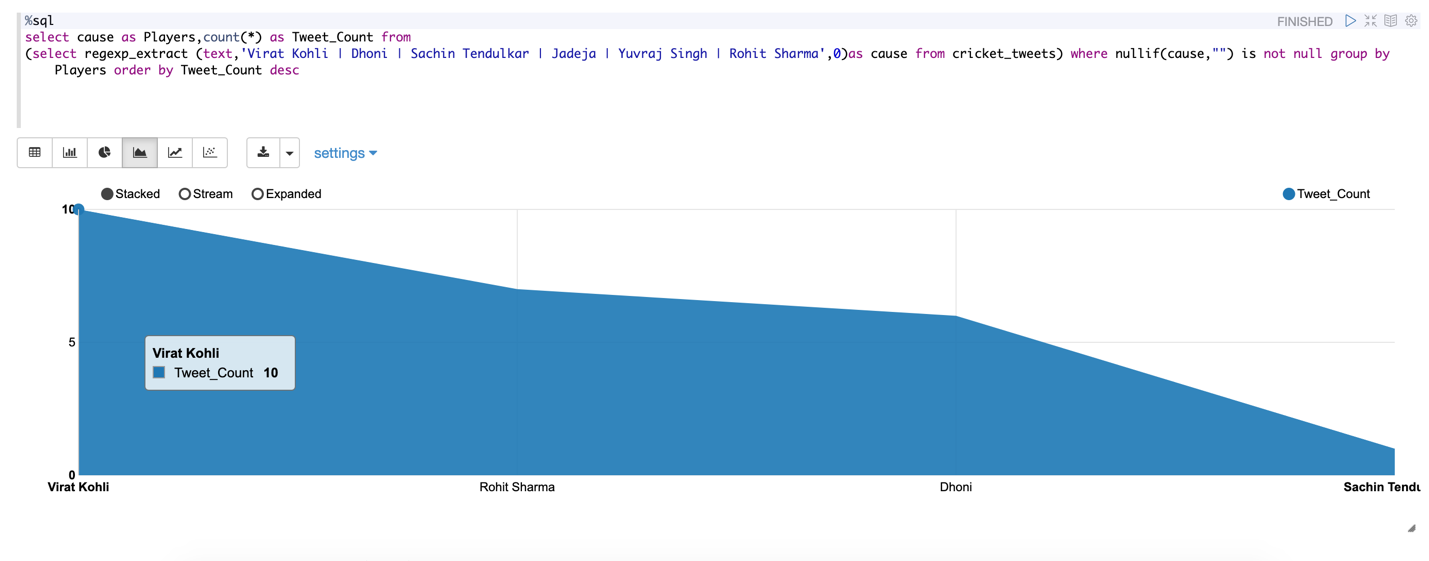
OBJECTIVE:

* To collect tweets related to Cricket.
* To analyze the collected tweets and develop interesting analysis on them.
* Create Visualizations on the analysis done.

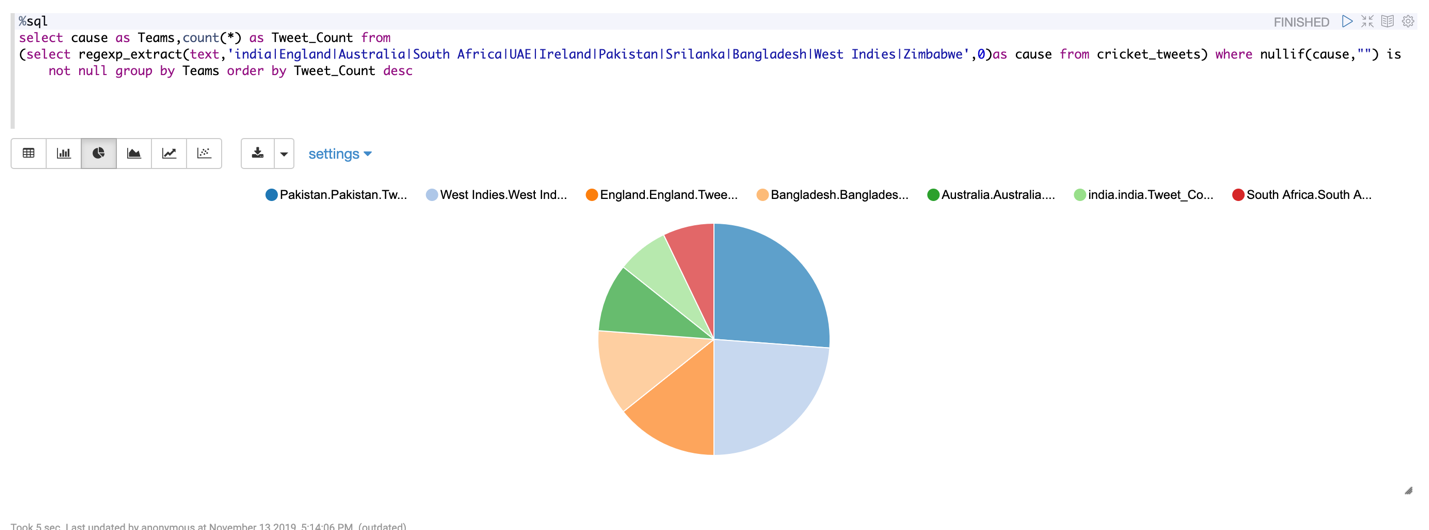
1.selecting tweets depending on the followers count of the user.



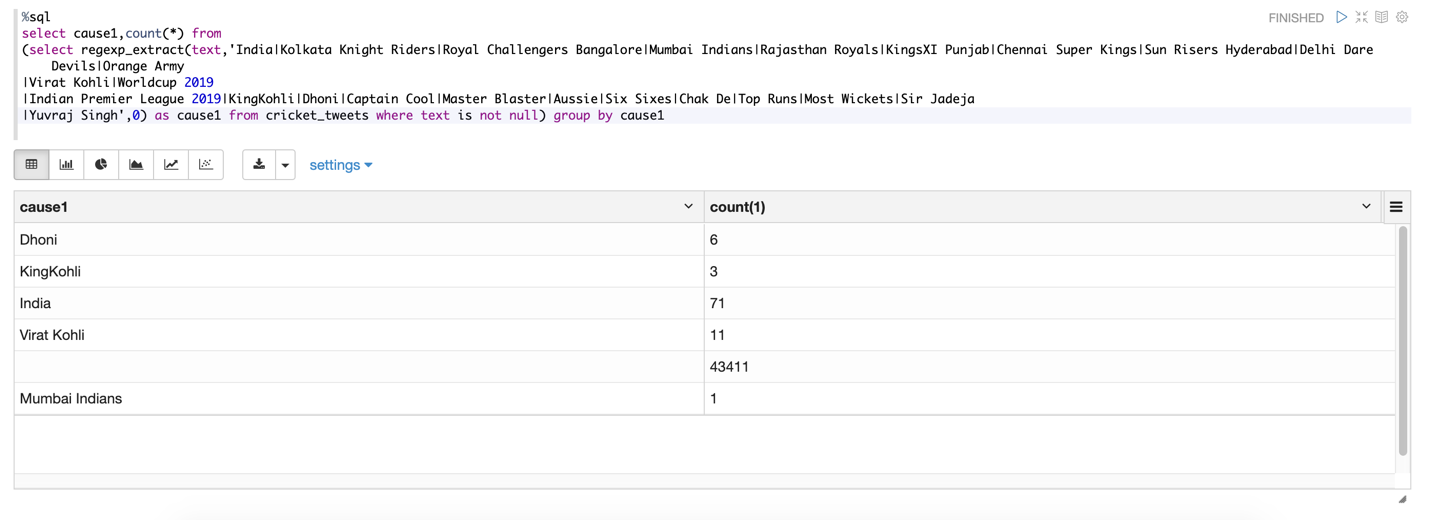
2. Tweets selected on the basis of player names.



3. Tweets on cricket Playing Nations.



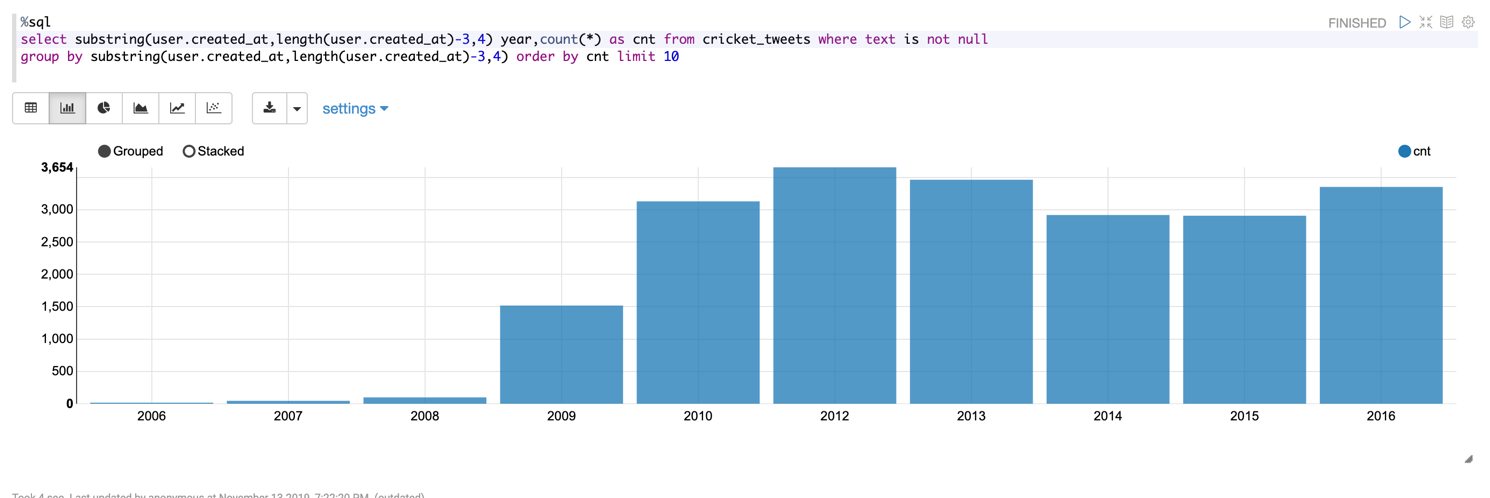
4. Tweets on different players and on teams of IPL.

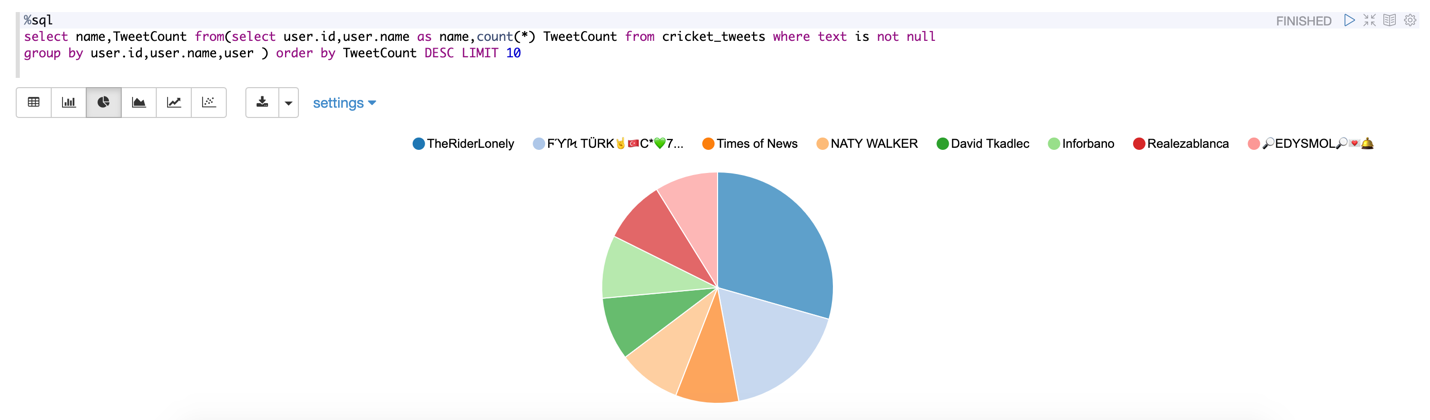


5. Tweets selected on the basis of distinct languages.

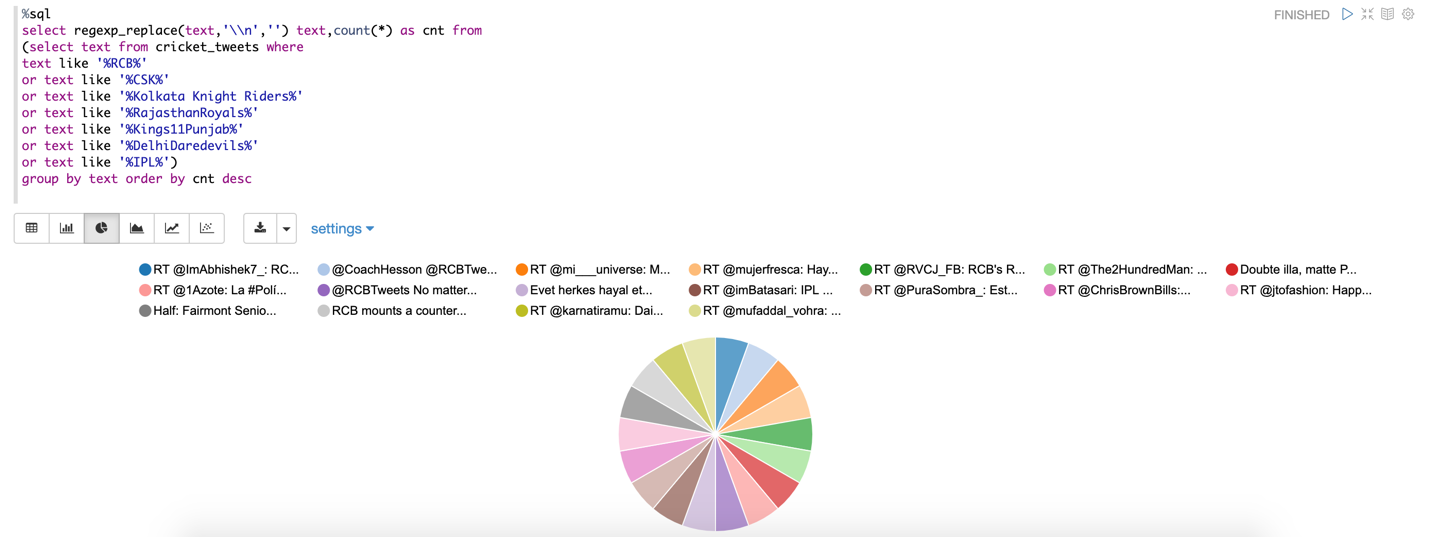


6. Tweets on different years.

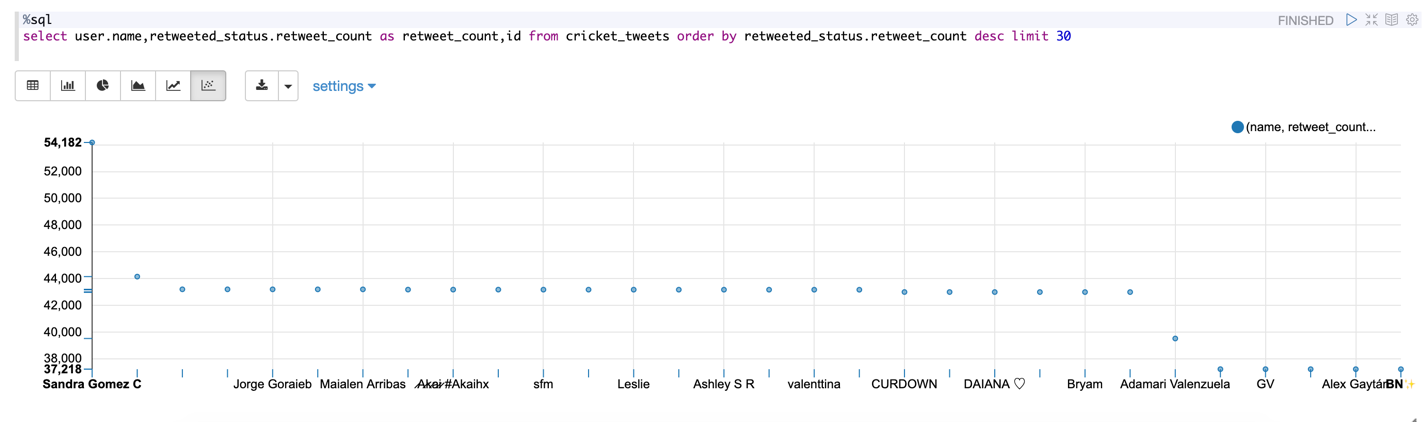


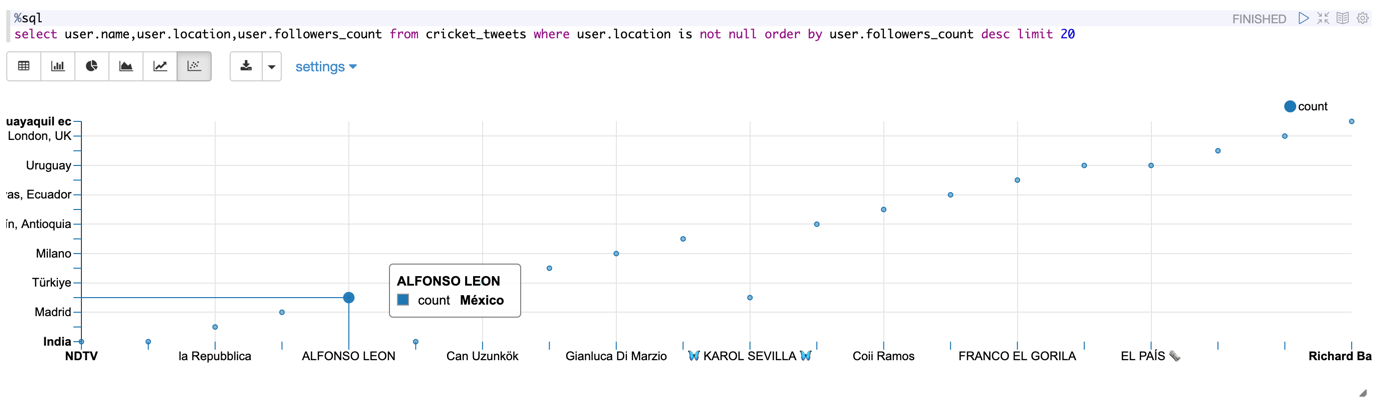
7. Number of Tweets based on the user. 

8. Tweets having the team names tweeted by the user.

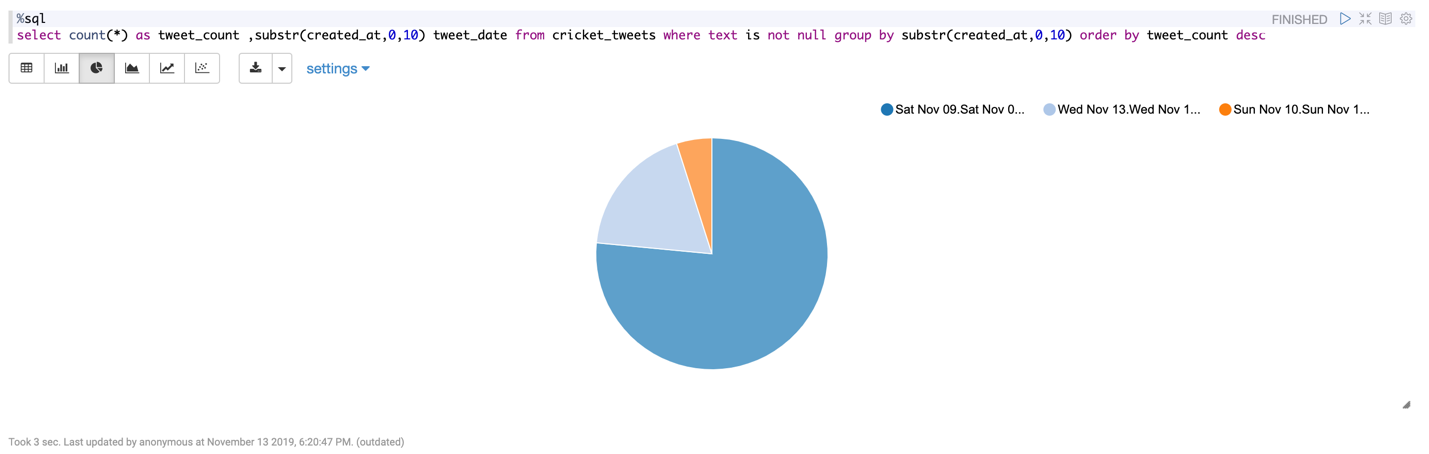


9. Tweets selected on the basis of retweet count.



10. Tweets selected on the basis of location. 

11. Tweets tweeted on different dates.

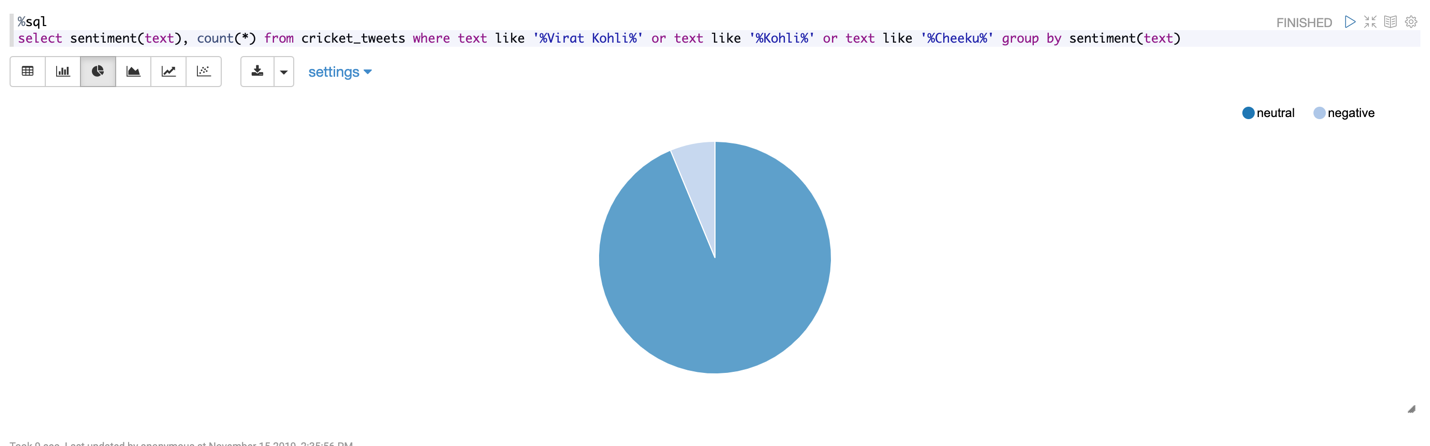


12.Twitter users who are verified.



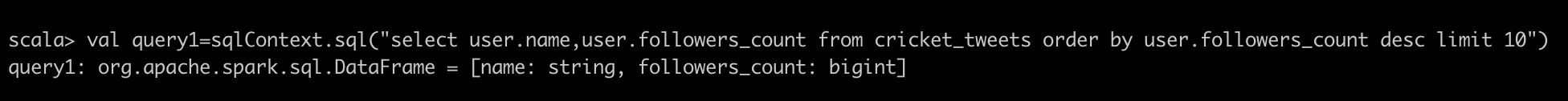
13. sentiment analysis.

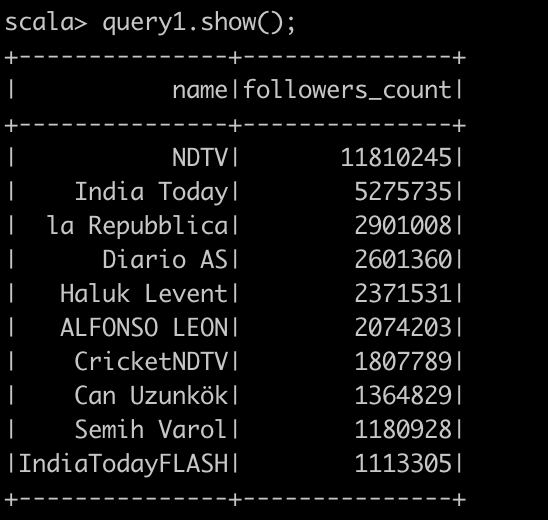


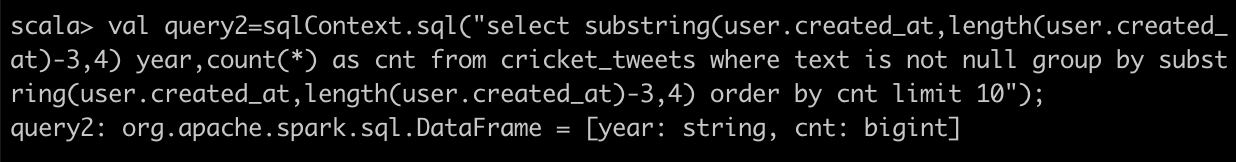


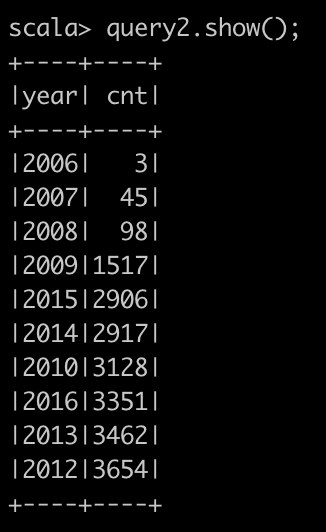
Manual Testing:

The queries are tested manually first whether it is giving the correct results or not.For instance let’s take the query of selecting the tweets depending on the followers count of the user.





Let’s take the query for tweets tweeted yearly. 



Testing with Spyder tool:

Let’s test the queries with spyder tool.

