1)CREATE AN ACCOUNT IN TWEETER APP. Get app

api\_key = "\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*"

api\_secret = "\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*"

access\_token = "\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*"

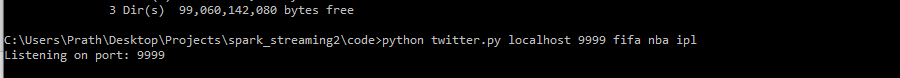
access\_token\_secret = "\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*"

2)WRITE PROGRAM TO CONNECT TO TWITTER STREAMING API USING TWEEPY.

3) RUN THE PYTHON PROGRAM.

**Twitter.py localhost 9999 fifa nba ipl**

**Twitter.py localhost 9999 fifa nba corona**



4) WRITE SPARK PROGRAM TO FILTER AND PROCESS DATA (Count hashtag in twitter to find the overall trend.)

4.1) Need to connect to a host and port.

4.2) Use spark session to read streaming data

4.3) Extract the message and split the message to get individual hash tag

5) Submit the spark job

**spark-submit count\_hash\_tag.py localhost 9999**

