1. What hour of the day does @PrezOno’s tweet the most on average, using every day we have twitter data?  Include a plot of the expected number of tweets for each hour of the day, for those he did tweet.  For example if Ono tweeted once every day at 12:30PM, his expected number of tweets between 12 and 1 would be 1.  If he alternates between 2 and 3 tweets per day, his average would be 2.5.

**Aim**:

Generate an output file using spark that has PrezOno tweet per hour. Use data from this output file to plot average tweets of Prezono per hour.

**Approach:**

* Created a RDD by reading input tweets files from HDFS.
* User defined function is passed in to this RDD. This function detects PrezOno tweets

And creates another RDD containing date and time he tweeted.(This is done bt using flatmap

* Then count the total number of Days by storing days alone into separate RDD. Total tweets of PrezOno per hour is counted by creating another RDD by using reduceByKey.
* Finally average tweets per hour is calculated by dividing total tweets per hour with number of days. This is stored in separate RDD. This RDD is written into output file

Plot for output data: At 5PM (17:00) on average there are more number of tweets.