# Text Mining

**Business Objective**: Extract tweets for any user and Perform sentimental analysis on the tweets

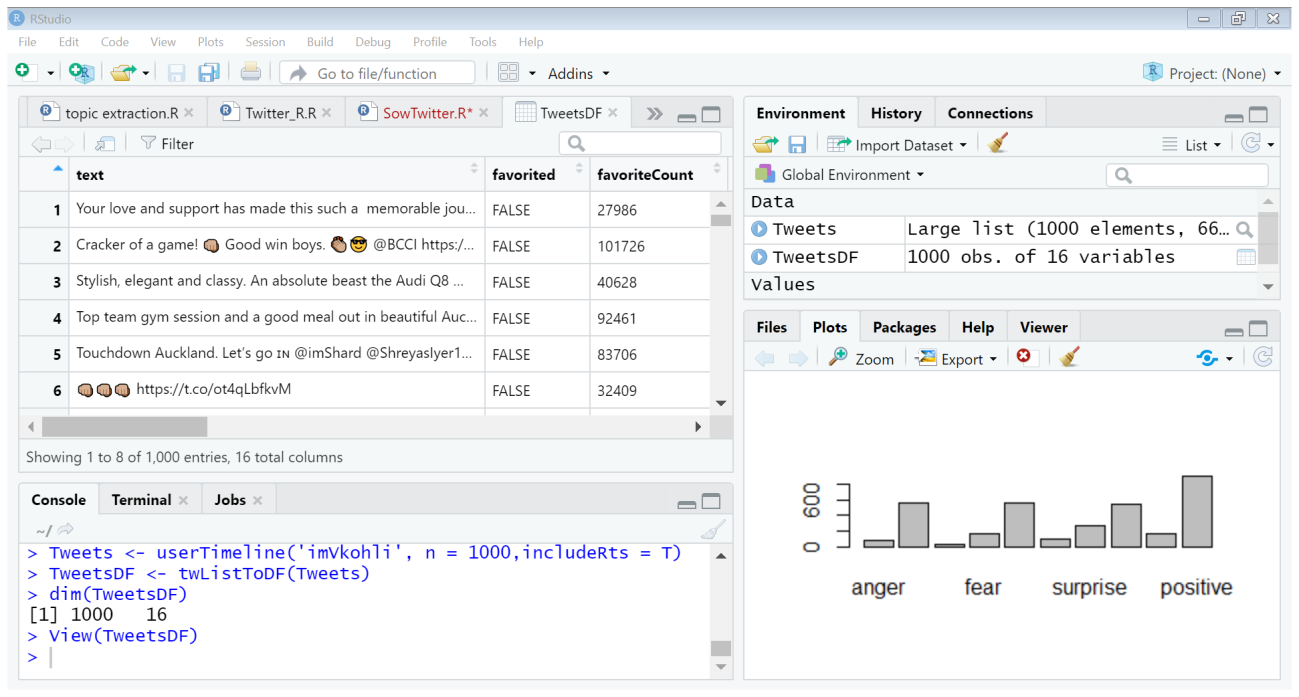
Step 1: Extract tweets for any user so I have extracted Virat Kohli tweeter data.

Step 2: Created a developer access on tweeter and generated the API key, Consumer key, Secret token.

consumerKey='2u1YoldtZAt9mJt4rvYLBA6rr', # Consumer Key (API Key)

consumerSecret='rPQZwMP07uc2sE8I7cwFP8Xi83rzg1Ll9XaU5UBx2VVyYQOSMy', #Consumer Secret (API Secret)

Step 3:Extrated the 1000 tweets of Virat with his username “imVkohli” below the view of his tweets.



Step 4: write the extracted file in xlxc formate

Step 5: Perform sentimental analysis on the tweets extracted .

Step 6: Save the required column in the text format.

Step 7:Read the text file and perform the syuzhet, bing, afinn,nrc methods on the text file to determine the setiments.

Step 8: below is the nrc data

anger anticipation disgust fear joy sadness surprise trust

1 0 0 0 0 0 0 0 0

2 0 1 0 1 3 0 1 1

3 0 1 0 0 1 0 1 1

4 1 2 0 1 2 0 0 0

5 0 2 0 0 2 0 1 3

6 0 0 0 0 0 0 0 0

negative positive

1 0 0

2 0 3

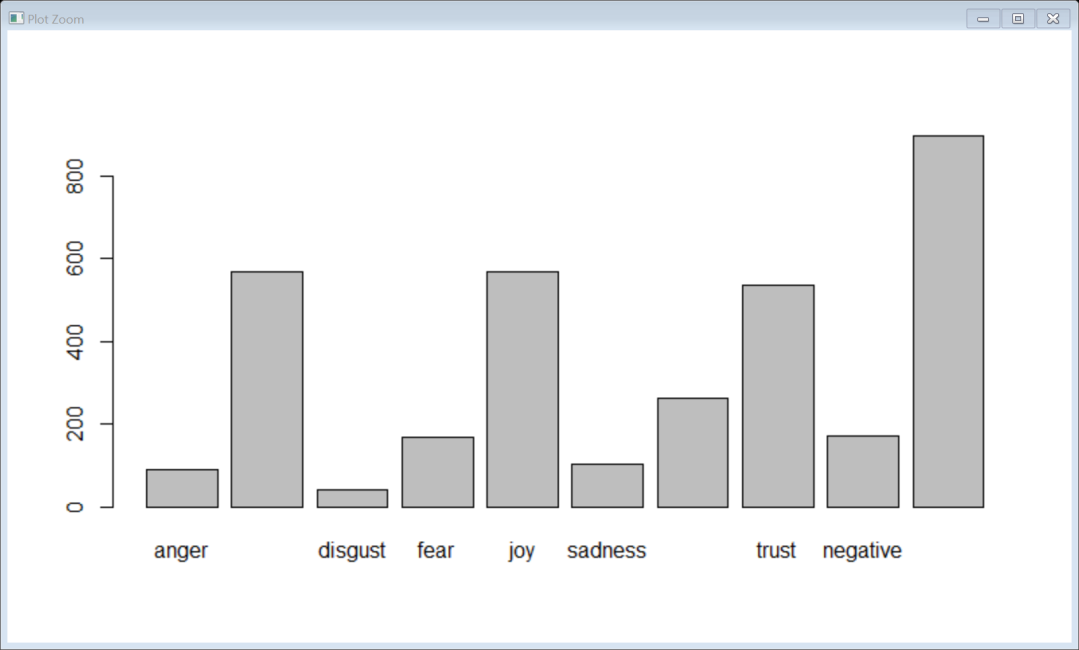
3 0 1

4 1 4

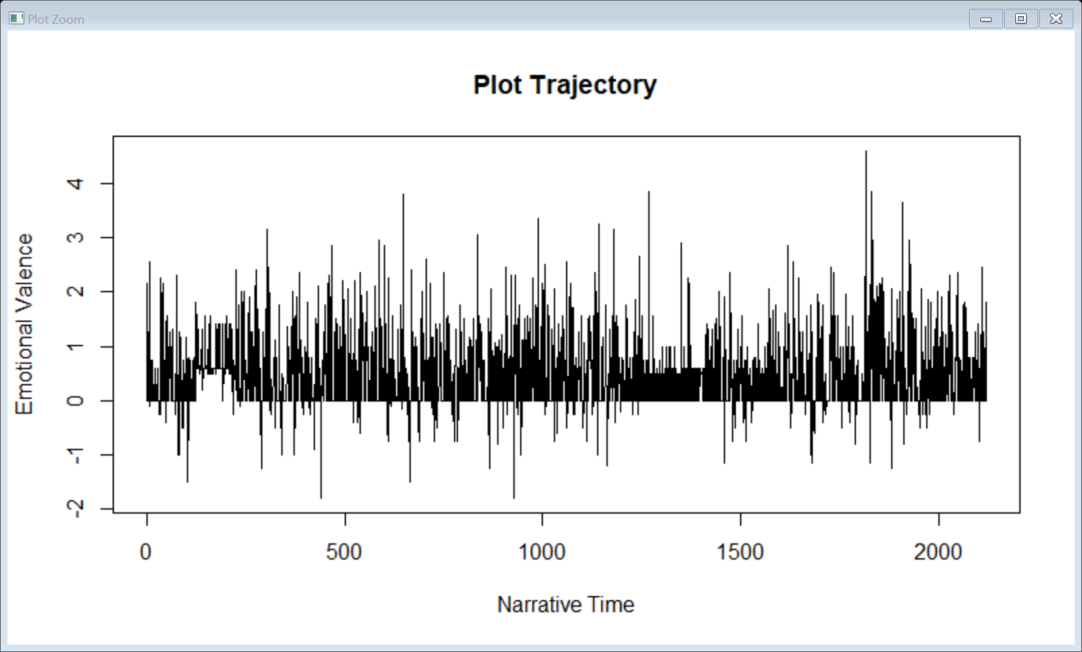
5 0 3

6 0 0

Step 9: below is the barplot of the nrc which shows positive words highest followed by joy,anticipation, trust,.



Step 10: Below is the plot with narrative time on the x axis and emotional valence on the y axis.



Step 11: To extract the sentence with the most negative emotional valence

Heartbreaking to hear the devastating news.

Step 12: : To extract the sentence with the most positive emotional valence

Happy Birthday Mahi Bai Wish you loads of success and happiness.

**Business Objective** Extract reviews of any product from ecommerce website like snapdeal and amazon

Step 1:Extract the amazon review for MacBook Air.

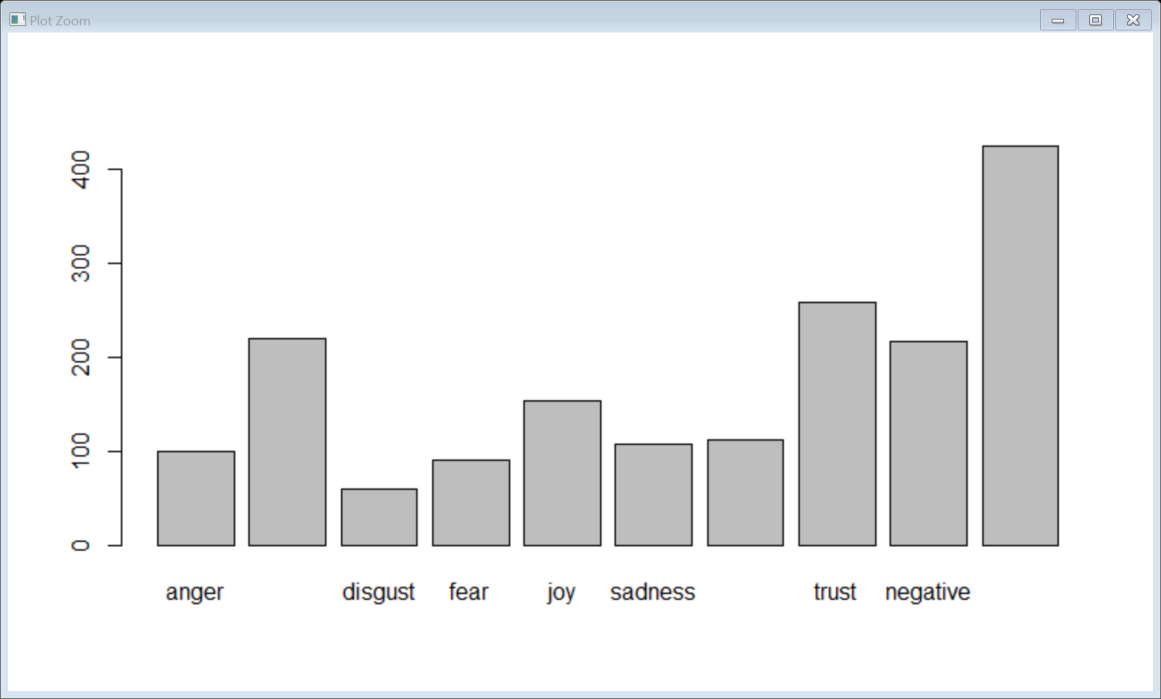
Step 2: Create the for loop and generate the review for 30 pages which has 10 review in each page.

Step 3: Save the review in text format.

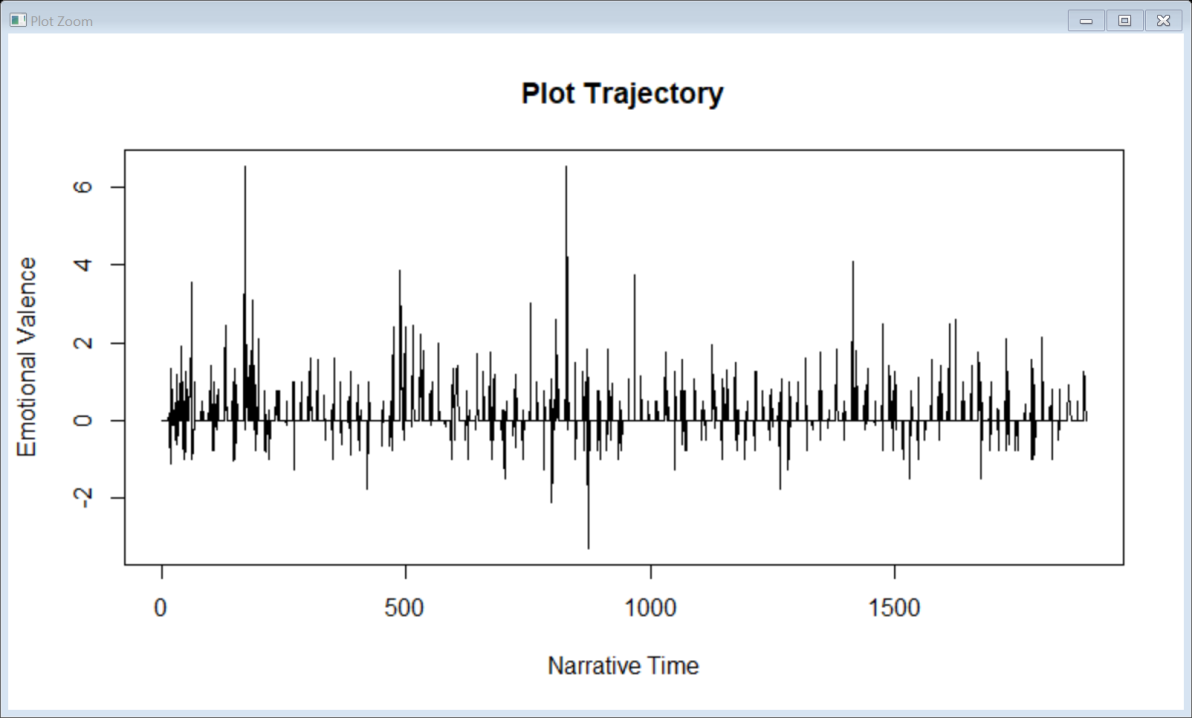
Step 4: Generating the sentimental analysis on the text data extracted from the amazon review.

Step 5: Read the text file and perform the syuzhet, bing, afinn,nrc methods on the text file to determine the sentiments.

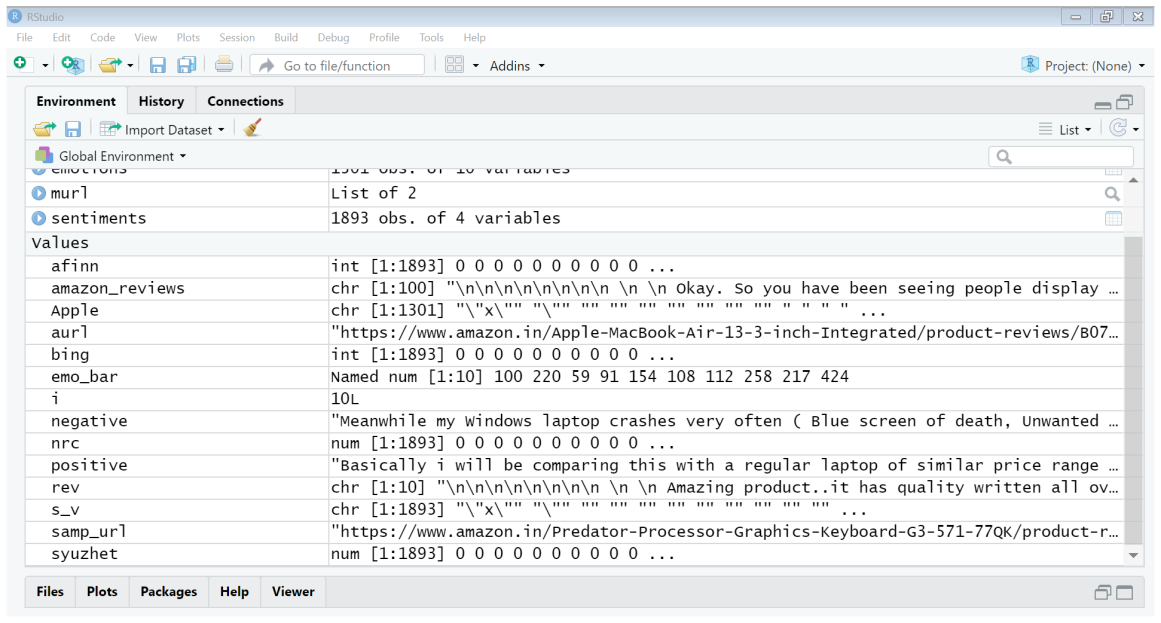
Step 6: Below is the barplot of the nrc which shows positive word highest followed by trust,negative, Anticipation.



Step 7: Below is the plot with narrative time on the x axis and emotional valence on the y axis.



Step 8: Below is Screenshot of the sentences with the most negative and positive emotional valence



**Business Objective**: Extract movie reviews for any movie from IMDB and perform sentimental analysis

Step 1: Extract movie reviews for any movie from IMDB , Venom movie review.

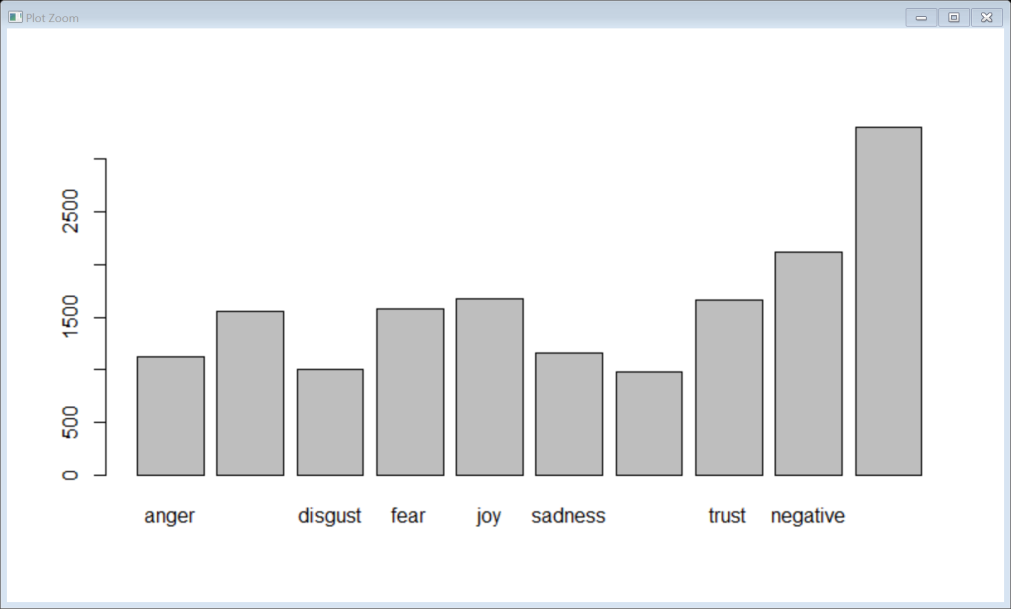
Step 2: Copy the URL and generate the for loop and in IMDB all the review is copied for specific movies as page loads the entire reviews.

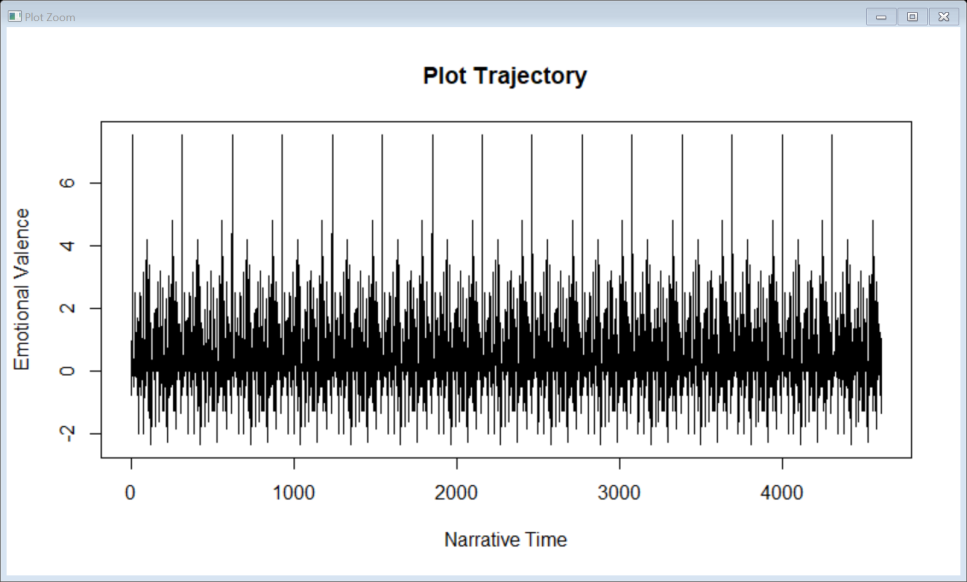
Step 3 : write it into the text file.

Step 4: Generating the sentimental analysis on the text data extracted from the IMDB review

Step 5: Read the text file and perform the syuzhet, bing, afinn,nrc methods on the text file to determine the sentiments.

Step 6: Below is the barplot of the nrc which shows positive word highest followed by trust,negative, Anticipation, joy.





Step 7: Above is the plot with narrative time on the x axis and emotional valence on the y axis.

Step 8: Below is Screenshot of the sentences with the most negative and positive emotional valence

