**Assignment – 1**

Last date: 23rd March 2018 (Friday)

Pull data from Twitter API for the hashtag #datascience and do the following

1. Apply text cleaning process
2. Create word cloud for unigram for top 50 words
3. Create bigram for top 50 tokens
4. Create a word cloud with words starting with # character i.e. word cloud for hashtags only
5. Create DTM matrix for top 50 words and join the same with the original data. Identify the top 5 words which has highest correlation with likes column