ReadMe file

=============

1. groupedTweets are the clustered file
   1. group is the cluster numbers; in other words all tweets with the same group id form the cluster
   2. tweetID is the unique tweet identifier
   3. you have also username, text, qScore and nScore
2. tweets.csv is the tweets collected with
   1. \_id is the unique tweet identifier, username, text, qScore and nScore

topic modelling is done on the text field

qScore is the quality score discussed in the class

nScore is the newsworthy score discussed in the class

if you don’t find any use for these then please ignore

It is up to you which part of the information you may use