Hello,

I have used the small chunk of my actual database to do all the testing as my machine will take very long to respond if worked with entire file. Once al set, I will used my actual data set to do the final testing.

My project is in the intermediate stage and so the scrip will be added and modified in the final submission

I have obtained my database from below link. It is publically available.

<https://ai.stanford.edu/~amaas/data/sentiment/>

I have extracted the files using un-zip tools and then as all the reviews are separate txt file, I used excel add-ins to convert them to one CSV.

I then merged the reviews under trial category and then used the dataset called Traintrial.csv

Packages I have used so far are

tm

stopwords

textstem

wordcloud

RColorBrewer

Syuzhet

ggplot2

sentimentr