


Help save net neutrality! A free, open internet is once again at stake—and we need your help.

Learn more

 [bendazz](#) / [gradprac1](#)

Branch: master ▾ [gradprac1](#) / [index.Rmd](#)

Find file Copy path

 **bendazz** adding solution 0fb9b52 4 days ago

1 contributor

44 lines (30 sloc) 1.32 KB

title	author	date	output
Practice Problem 1		December 2, 2017	html_document

1. Download this file: [Mercyhurst Catalog](#).
2. You'll need tidytext,dplyr, and ggplot2.
3. Create a plot illustrating the net positive and negative sentiments over line-blocks of size 80, progressing from the beginning of the document to the end. Use the afinn lexicon.
4. Save your work in a single R script file. Use your last name to name the file. For instance, if my name is Smith, I would name the file smith.R.
5. Upload the file using this link: [upload link](#)

###Solution###

The plot should look like this:

```
library(dplyr)
library(tidytext)
library(ggplot2)
catalog_scan<-scan('mercyhurst.txt',what=character(),sep='\n')
catalog_lines<-data_frame(line=1:24066,text=catalog_scan)
catalog_words<-unnest_tokens(catalog_lines,word,text)

affin<-get_sentiments('afinn')
catalog_words<-inner_join(catalog_words,affin)

catalog_words<-catalog_words%>%
  mutate(group=line%/%80)

catalog_groups<-catalog_words%>%
  group_by(group)%>%
  summarize(sentiment=sum(score))

ggplot()+
  geom_col(data=catalog_groups,aes(x=group,y=sentiment))
```