Sentimientos en informes de estabilidad financiera

Luis Ortiz-Cevallos

2022-07-07

Contents

1 INTRODUCCIÓN

El análisis de sentimiento te ayuda extraer de un texto el estado de animo del editore de ese texto.

Por ejemplo, a partir de un texto que contenga la conversación entre dos personas podemos usar la función (`qdap's polarity()`) la cual califica el texto en dos escalas: **positivas y negativas**.

Noten que para aplicar esa función los input son: + simples carácteres o dataframe + agrupamiento de palabras

En tanto el output es un objeto de la clase polaridad

1.1 El text mining con bolsas de palabras

Corpus es un conjunto de texto a partir del cual se pueden aplicar varios procesos.

La función (`VectorSource()`) pasa un vector de carácteres a una fuente de texto. A continuación un ejemplo:

La función (`VCorpus()`) pasa una fuente de texto a un Corpus.

Luego es importante limpiar tu Corpus. con funciones como: (`removePunctuation() y stripWhitespace()`) del paquete (`tm`) y la función (`replace_abbreviation()`) del paquete (`qdap`)

```
library(qdap)
library(tidyverse)
library(pdftools)
library(stringr)
library(tm)
#pdf.text <- pdf_text("muestra.pdf")
#write.table(pdf.text, "muestra.txt", sep=";")
pdf.text2 <- pdf_text("muestra.pdf") %>% str_split("\n")
pdf.text2<-unlist(pdf.text2)
write.table(pdf.text2, "muestra.txt", sep=";")
tm_define<-read.csv("muestra.txt")
```

```
tm_define<-tm_define[1,1]
tm_define
```

```
## [1] "1;Text mining is the process of distilling actionable insights from text."
```

```
tm_vector <- VectorSource(tm_define)
tm_corpus <- VCorpus(tm_vector)
content(tm_corpus[[1]])
```

```
## [1] "1;Text mining is the process of distilling actionable insights from text."
```

```
clean_corpus <- function(corpus){
  corpus <- tm_map(corpus, content_transformer(replace_abbreviation))
  corpus <- tm_map(corpus, removePunctuation)
  corpus <- tm_map(corpus, removeNumbers)
  corpus <- tm_map(corpus, removeWords, c(stopwords("en"), "coffee"))
  corpus <- tm_map(corpus, content_transformer(tolower))
  corpus <- tm_map(corpus, stripWhitespace)
  return(corpus)
}
tm_clean <- clean_corpus(tm_corpus)
content(tm_clean[[1]])
```

```
## [1] "text mining process distilling actionable insights text"
```

An unnumbered section

2 Cross-references

Cross-references make it easier for your readers to find and link to elements in your book.

2.1 Chapters and sub-chapters

There are two steps to cross-reference any heading:

1. Label the heading: `# Hello world {#nice-label}`.
   - Leave the label off if you like the automated heading generated based on your heading title: for example, `# Hello world` = `# Hello world {#hello-world}`.
   - To label an un-numbered heading, use: `# Hello world {-#nice-label}` or `{# Hello world .unnumbered}`.
2. Next, reference the labeled heading anywhere in the text using `\@ref(nice-label)`; for example, please see Chapter 2.
   - If you prefer text as the link instead of a numbered reference use: any text you want can go here.

2.2 Captioned figures and tables

Figures and tables *with captions* can also be cross-referenced from elsewhere in your book using `\@ref(fig:chunk-label)` and `\@ref(tab:chunk-label)`, respectively.

See Figure 1.

```
par(mar = c(4, 4, .1, .1))
plot(pressure, type = 'b', pch = 19)
```

Don't miss Table 1.

```
knitr::kable(
  head(pressure, 10), caption = 'Here is a nice table!',
```

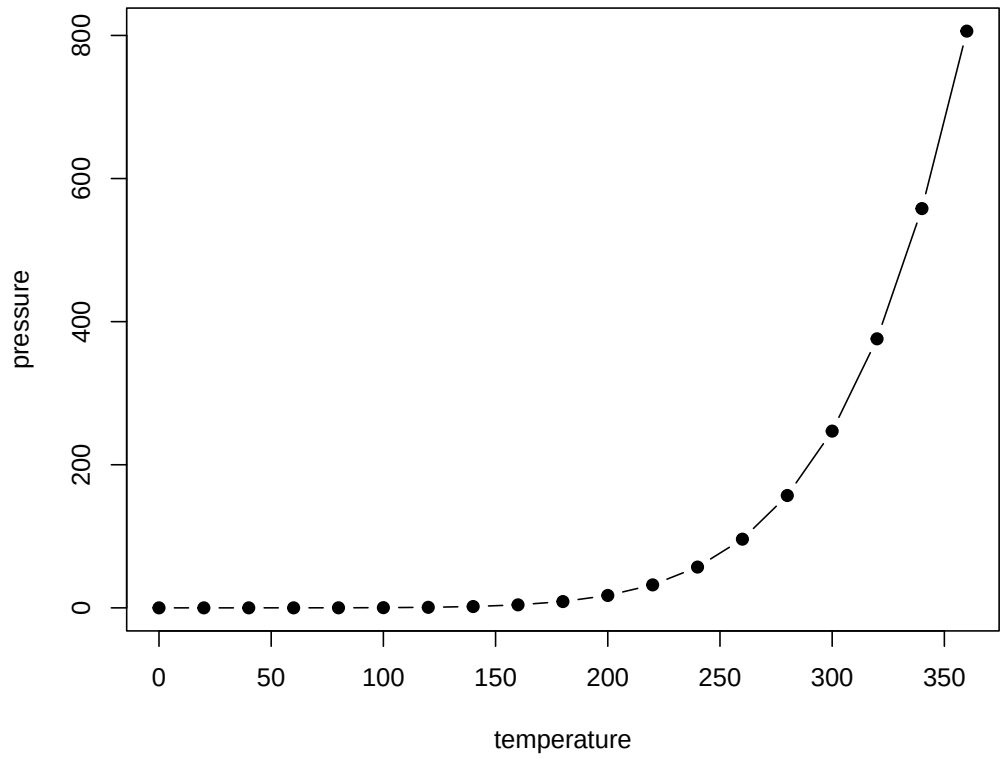

Figure 1: Here is a nice figure!

```
  booktabs = TRUE
)
```

3 Parts

You can add parts to organize one or more book chapters together. Parts can be inserted at the top of an .Rmd file, before the first-level chapter heading in that same file.

Add a numbered part: `# (PART) Act one {-}` (followed by `# A chapter`)

Add an unnumbered part: `# (PART\*) Act one {-}` (followed by `# A chapter`)

Add an appendix as a special kind of un-numbered part: `# (APPENDIX) Other stuff {-}` (followed by `# A chapter`). Chapters in an appendix are prepended with letters instead of numbers.

Table 1: Here is a nice table!

temperature	pressure
0	0.0002
20	0.0012
40	0.0060
60	0.0300
80	0.0900
100	0.2700
120	0.7500
140	1.8500
160	4.2000
180	8.8000