Franz Kafka's The Metamorphosis

Chris Kozak

December 15, 2017

Abstract

In this article, we'll do some basic sentiment analysis. We're going to construct a wordcloud for Franz Kafka's The Metamorphosis. This is one of Kafka's most well-known stories.[1].

1 Packages

First, we've already loaded the following packages: dplyr for querying, tidytext and text mining (tm) for working with text, gutenbergr for downloading open-source books, wordcloud for generating visualizations, and knitr for compiling documents from R Markdown, Sweave and various other formats.

```
get_sentiments('nrc')

## # A tibble: 13,901 x 2
##              word sentiment
##             <chr>     <chr>
## 1         abacus     trust
## 2        abandon      fear
## 3        abandon  negative
## 4        abandon   sadness
## 5      abandoned     anger
## 6      abandoned      fear
## 7      abandoned  negative
## 8      abandoned   sadness
## 9    abandonment     anger
## 10   abandonment      fear
## # ... with 13,891 more rows
```

---

[1]The short story was first published in German in 1915. Many translations exist, but we'll be using the 2002 version by David Wyllie

2   Downloading The Metamorphosis

For this exercise, we'll be working with Franz Kafka's The Metamorphosis, and strictly comparing the amount of joy in the book to the amount of fear. In this book, a young man turns into a giant cockroach. His family at first tries to take care of him, then get sick of taking care of him, and sick of the burdens that his metamorphosis has imposed on the family, and let him die.

```
metamorphosis <- gutenberg_download(5200)

metamorphosis_words <- metamorphosis %>%
  unnest_tokens(word, text)
```

Word is the name of the new column we're making with the unnest$_t okens command$

3   Quantifying Fear and Joy

Next, we'll be using the tidytext package to quantify sentiment. The tidytext package comes with several built-in sentiment lexicons, including bind and AFINN, which assign values to words based on their perceived positive or negative values. But for this exercise, we will focus on the NRC emotion lexicon. This assigns words based on the seven primary emotions: joy, sadness, fear, anger, disgust, surprise, and trust. For this exercise, we only want two emotions: joy and fear.

4   Fear

For the next step, we will get only the fear words from metamorphosis. As Kafka's story involves a cockroach, there should be quite a few. The following commands will create a new dataframe that contains only the fear words from the NRC sentiment lexicon, and then join those words with the ones found in The Metamorphosis.

```
nrc_fear <- get_sentiments('nrc') %>%
  filter(sentiment == "fear")

metamorphosis_fear_words<-inner_join(nrc_fear,metamorphosis_words)
```

5   Fear in the Metamorphosis

Next, we'll see how many unique fear words we have, and how what kind of words they are using the dim() and head() commands from R:

```
dim(metamorphosis_fear_words)[1]
```

```
## [1] 271
```

```
head(metamorphosis_fear_words$word,n=20)
```

```
##  [1] "abandoned" "abandoned" "accused"   "afraid"    "afraid"
##  [6] "afraid"    "alarm"     "alarm"     "alarm"     "alarm"
## [11] "alarm"     "alarm"     "anxiety"   "anxious"   "attack"
## [16] "attack"    "austere"   "avoid"     "avoid"     "avoid"
```

This shows that The Metamorphosis has 271 words expressing fear, and lists some examples above.

6   Fear Wordcloud

```
wordcloud(metamorphosis_fear_words$word,
          colors = c("red","black","gray20",
                     "indianred4","tomato4"))
```
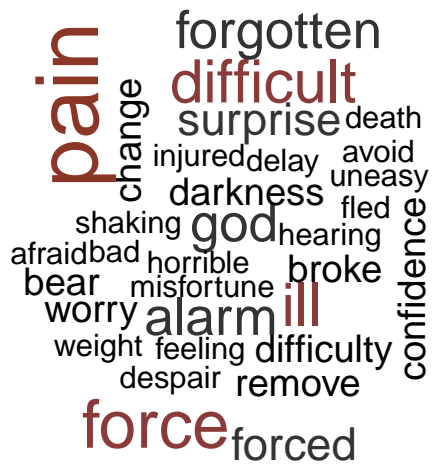
7 Joy

And now, we will do the same thing for the joy words from The Metamorphosis.

```
nrc_joy <- get_sentiments('nrc') %>%
  filter(sentiment == "joy")

metamorphosis_joy_words<-inner_join(nrc_joy,metamorphosis_words)
```

8 Joy in The Metamorphosis

```
dim(metamorphosis_joy_words)[1]

## [1] 344

head(metamorphosis_joy_words$word, n=20)

##  [1] "affection"    "appreciation" "beautiful"    "beautiful"
##  [5] "beer"         "captivate"    "cash"         "cash"
##  [9] "cheer"        "child"        "child"        "child"
## [13] "clean"        "clean"        "clean"        "comfort"
## [17] "confidence"   "confidence"   "confidence"   "confidence"
```

9   Joy Wordcloud

We can see from the dimension - or dim() - command, that there are 344 unique words expressing joy in The Metamorphosis. Beer is one of them. Ironically, it features more unique words expressing joy than fear. Finally, we will generate a wordcloud to display the most common words of joy.

```
wordcloud(metamorphosis_joy_words$word,
          colors =c("black","blue","darkcyan",
                    "darkolivegreen","aquamarine3"))
```

entertaining finally

improvement god music

mother

feeling hope young true

respect clean smile kind save

confidence weight

pay child found white

daughter glad peace

money surprise friendly

food present

good success