Positive and Negative Words in Dracula

Ron Richardson

November 15, 2017

Outline

Install and Load Libraries

- `library(dplyr)`

Install and Load Libraries

- `library(dplyr)`

- `library(gutenbergr)`

Install and Load Libraries

- `library(dplyr)`

- `library(gutenbergr)`

- `library(tidytext)`

Install and Load Libraries

- `library(dplyr)`

- `library(gutenbergr)`

- `library(tidytext)`

- `library(stringr)`

Install and Load Libraries

- `library(dplyr)`

- `library(gutenbergr)`

- `library(tidytext)`

- `library(stringr)`

- `library(ggplot2)`

Searching for a Novel

```
df<-gutenberg_works(str_detect(title, 'Dracula'))
df$gutenberg_id
```

```
## [1]   345 10150
```

```
df$title
```

```
## [1] "Dracula"          "Dracula's Guest"
```

Download Dracula from Project Gutenberg

```r
dracula<-gutenberg_download(345)
colnames(dracula)

## [1] "gutenberg_id" "text"

substr(dracula$text[500], 1, 21)

## [1] "my own disappointment"
```

Unpack the Words

```
words<-dracula%>%
  unnest_tokens(word, text)
colnames(words)

## [1] "gutenberg_id" "word"

words[500,]

## # A tibble: 1 x 2
##    gutenberg_id  word
##           <int> <chr>
## 1           345  have
```

The Bing Lexicon

```
bing<-get_sentiments('bing')
colnames(bing)

## [1] "word"      "sentiment"

bing[500,]

## # A tibble: 1 x 2
##     word sentiment
##    <chr>     <chr>
## 1 bereft  negative
```

Joining Dracula with Bing

```
words<-inner_join(words, bing)
words$gutenberg_id<-NULL
colnames(words)

## [1] "word"      "sentiment"
```