Comparison Barplots

Charles Redmond

November 17, 2017

Outline

Install and Load Libraries

- ```
  library(dplyr)
  ```
  ```

Install and Load Libraries

- ```
  library(dplyr)
  ```
- ```
  library(tidytext)
  ```

Install and Load Libraries

- `library(dplyr)`
- `library(tidytext)`
- `library(gutenbergr)`

Install and Load Libraries

- `library(dplyr)`
- `library(tidytext)`
- `library(gutenbergr)`
- `library(ggplot2)`

Install and Load Libraries

- `library(dplyr)`
- `library(tidytext)`
- `library(gutenbergr)`
- `library(ggplot2)`
- `library(stringr)`

Access Project Gutenberg

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

## [1]   345 10150

df$title

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

Download Dracula

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

## [1] "gutenberg_id" "text"

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

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

Unpack the Words

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

## [1] "gutenberg_id" "word"

dracula_words[498:500,]

## # A tibble: 3 x 2
##   gutenberg_id  word
##          <int> <chr>
## 1          345  fail
## 2          345    to
## 3          345  have
```

The Bing Lexicon

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

## [1] "word"      "sentiment"

  bing[498:500,]

## # A tibble: 3 x 2
##          word sentiment
##         <chr>     <chr>
## 1     bereave  negative
## 2 bereavement  negative
## 3      bereft  negative
```

The Inner Join

```
  dracula_words<-inner_join(dracula_words,bing)

## Joining, by = "word"

  dracula_words$gutenberg_id<-NULL
  dracula_words[498:500,]

## # A tibble: 3 x 2
##       word sentiment
##      <chr>     <chr>
## 1   great  positive
## 2    love  positive
## 3 crowded  negative
```