

Compraison Barplots

Justin Minsk

November 17, 2017

Load and Install Libraries

► `library(dplyr)`

Load and Install Libraries

- ▶ `library(dplyr)`

- ▶ `library(tidytext)`

Load and Install Libraries

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

Load and Install Libraries

► `library(dplyr)`

► `library(tidytext)`

► `library(gutenbergr)`

► `library(ggplot2)`

Load and Install 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 <- gutenberglownload(345)
colnames(dracula)

## [1] "gutenberg_id" "text"

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

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


Unpacking the Words

```
dracula_words <- dracula%>%  
  unnest_tokens(word, text)  
colnames(dracula_words)  
  
## [1] "guttenberg_id" "word"  
  
dracula_words[498:500,]  
  
## # A tibble: 3 x 2  
##   guttenberg_id  word  
##           <int> <chr>  
## 1           345  fail  
## 2           345   to  
## 3           345  have
```

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
```

Interjoin

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

```
## Joining, by = "word"
```

```
dracula_words$gutenberg_id <- NULL  
colnames(dracula_words)
```

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

```
dracula_words[498:500,]
```

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