

RWorksheets#6 (Group 3)

2023-12-21

#Group 3 #Members: #Laurence Aguas #Paula Salvador #Dianah Marie Canonicato #Ann Margareth Camayodo

The Shawhank Redemption #1.

```
library(readr)
library(tm)
```

```
## Loading required package: NLP
```

```
library(NLP)
library(RColorBrewer)
```

```
Shawhank <- read.csv("RevReviews.csv")
ShawhankRatings <- read.csv("RevRatings.csv")
```

```
# Sentimental Analysis Ratings
```

```
ShawhankSentimentAnalysis <- ifelse(ShawhankRatings$Ratings > 5, "Positive",
                                     ifelse(ShawhankRatings$Ratings == 5, "Average",
                                             "Negative"))
```

```
ShawhankSentimentAnalysis
```

```
## [1] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [7] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [13] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [19] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [25] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [31] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [37] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [43] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [49] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [55] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [61] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [67] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [73] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [79] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [85] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [91] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [97] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [103] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [109] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [115] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [121] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [127] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [133] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [139] "Positive" "Negative" "Positive" "Positive" "Positive" "Positive"
## [145] "Positive" "Positive" "Positive" "Negative" "Negative" "Positive"
```

```
## [151] "Negative" "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [157] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [163] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [169] "Positive" "Positive" "Positive" "Positive" "Negative" "Negative" "Negative"
## [175] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [181] "Positive" "Positive" "Positive" "Negative" "Positive" "Positive" "Positive"
## [187] "Positive" "Positive" "Negative" "Positive" "Positive" "Positive" "Positive"
## [193] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [199] "Positive" "Negative" "Positive" "Positive" "Positive" "Positive" "Positive"
## [205] "Positive" "Average" "Positive" "Positive" "Positive" "Positive" "Positive"
## [211] "Positive" "Positive" "Average" "Positive" "Positive" "Positive" "Negative"
## [217] "Negative" "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [223] "Positive" "Positive" "Positive" "Negative" "Positive" "Positive" "Positive"
## [229] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [235] "Positive" "Positive" "Average" "Average" "Positive" "Positive" "Positive"
## [241] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [247] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [253] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [259] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [265] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [271] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [277] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [283] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [289] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [295] "Average" "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [301] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [307] "Positive" "Positive" "Positive" "Negative" "Positive" "Positive" "Positive"
## [313] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
## [319] "Positive" "Positive" "Positive" "Positive" "Positive" "Positive" "Positive"
```

```
# Set encoding to UTF-8
```

```
Shawhank$TextReviews <- iconv(Shawhank$TextReviews, to = "UTF-8", sub = "byte")
```

```
# Corpus for text analysis
```

```
ShawhankCorpus <- Corpus(VectorSource(Shawhank$TextReviews))
```

```
# Corpus for text analysis
```

```
ShawhankCorpus <- Corpus(VectorSource(Shawhank$TextReviews))
```

```
# Remove English stop words
```

```
ShawhankCorpus <- tm_map(ShawhankCorpus, content_transformer(tolower))
```

```
## Warning in tm_map.SimpleCorpus(ShawhankCorpus, content_transformer(tolower)):  
## transformation drops documents
```

```
ShawhankCorpus <- tm_map(ShawhankCorpus, removePunctuation)
```

```
## Warning in tm_map.SimpleCorpus(ShawhankCorpus, removePunctuation):  
## transformation drops documents
```

```
ShawhankCorpus <- tm_map(ShawhankCorpus, removeNumbers)
```

```
## Warning in tm_map.SimpleCorpus(ShawhankCorpus, removeNumbers): transformation  
## drops documents
```



```
## NULL
#Visualization
library(ggplot2)

##
## Attaching package: 'ggplot2'
## The following object is masked from 'package:NLP':
##
##      annotate
# Create a data frame from the sentiment analysis results
sentiment_data <- data.frame(legend = ShawhankSentimentAnalysis)

# Plot sentiment distribution
ggplot(sentiment_data, aes(x = legend, fill = legend)) +
  geom_bar() +
  labs(title = "Visual",
       x = "Ratings",
       y = "Numbers Of Reviewers") +
  scale_fill_manual(values = c("Positive" = "green", "Average" = "yellow", "Negative" = "red"))
```

