Employee Sentiment Analysis

Elroy Huynh

2023-01-29

Introduction

This project is a Sentiment Analysis of comments/feedback collected from employees during an employee engagement/safety survey. The survey results were collected in Microsoft Forms, and the comments data were extracted and transformed in Excel.

This project is shared with permission from the employer.

Setting Up My Environment

Notes: Setting up my R environment by loading the following packages:

```
library(tidyverse)
```

```
## ── Attaching packages ──────────────────────────────── tidyverse 1.3.2 ──
## ✔ ggplot2 3.4.0      ✔ purrr   1.0.0
## ✔ tibble  3.1.8      ✔ dplyr   1.0.10
## ✔ tidyr   1.2.1      ✔ stringr 1.5.0
## ✔ readr   2.1.3      ✔ forcats 0.5.2
## ── Conflicts ───────────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
```

```
library(tidytext)
library(syuzhet)
library(dplyr)
library(tibble)
library(stringr)
```

Importing The Survey Data

Importing the survey data that was previously cleaned and transformed in Excel

```
df <-read.csv("/Users/elroyhuynh/Employee Survey Sentiment Analysis/Survey_Data.csv", header = TRUE)
```

Convert Text From "Comments" Column To Lower-Case

Converting all the strings from the "Comments" column to lower-case:

```
text.df <- tibble(text = str_to_lower(df$Comments))
```

Use 'syuzhet' Package

Using the syuzhet package to analyze sentiments based on NRC sentiment dictionary, and display the sum of all emotion counts.

```
emotions <- get_nrc_sentiment(text.df$text)
```

```
## Warning: `spread_()` was deprecated in tidyr 1.2.0.
## i Please use `spread()` instead.
## i The deprecated feature was likely used in the syuzhet package.
##   Please report the issue to the authors.
```

```
emo_bar <- colSums(emotions)
emo_sum <- data.frame(count=emo_bar, emotions=names(emo_bar))
emo_sum
```

```
##                  count      emotions
## anger               35         anger
## anticipation       105  anticipation
## disgust             21       disgust
## fear                44          fear
## joy                 58           joy
## sadness             54       sadness
## surprise            21      surprise
## trust              159         trust
## negative           100      negative
## positive           251      positive
```

Visualizing The Data:

Load ggplot2 library to visualize 'emo_sum' results, create a bar chart showing the counts of all eight different emotions and positive/negative ratings.

```
library(ggplot2)

SAplot <- ggplot(emo_sum, aes(x = reorder(emotions, -count), y = count, fill = emotion
s)) + geom_bar(stat = 'identity') +theme_minimal() +theme(legend.position="none")
```

Employee Sentiment Analysis 2022 Plot

Add chart title for and axis labels for Sentiment Analysis bar chart

Employee Sentiment Analysis 2022
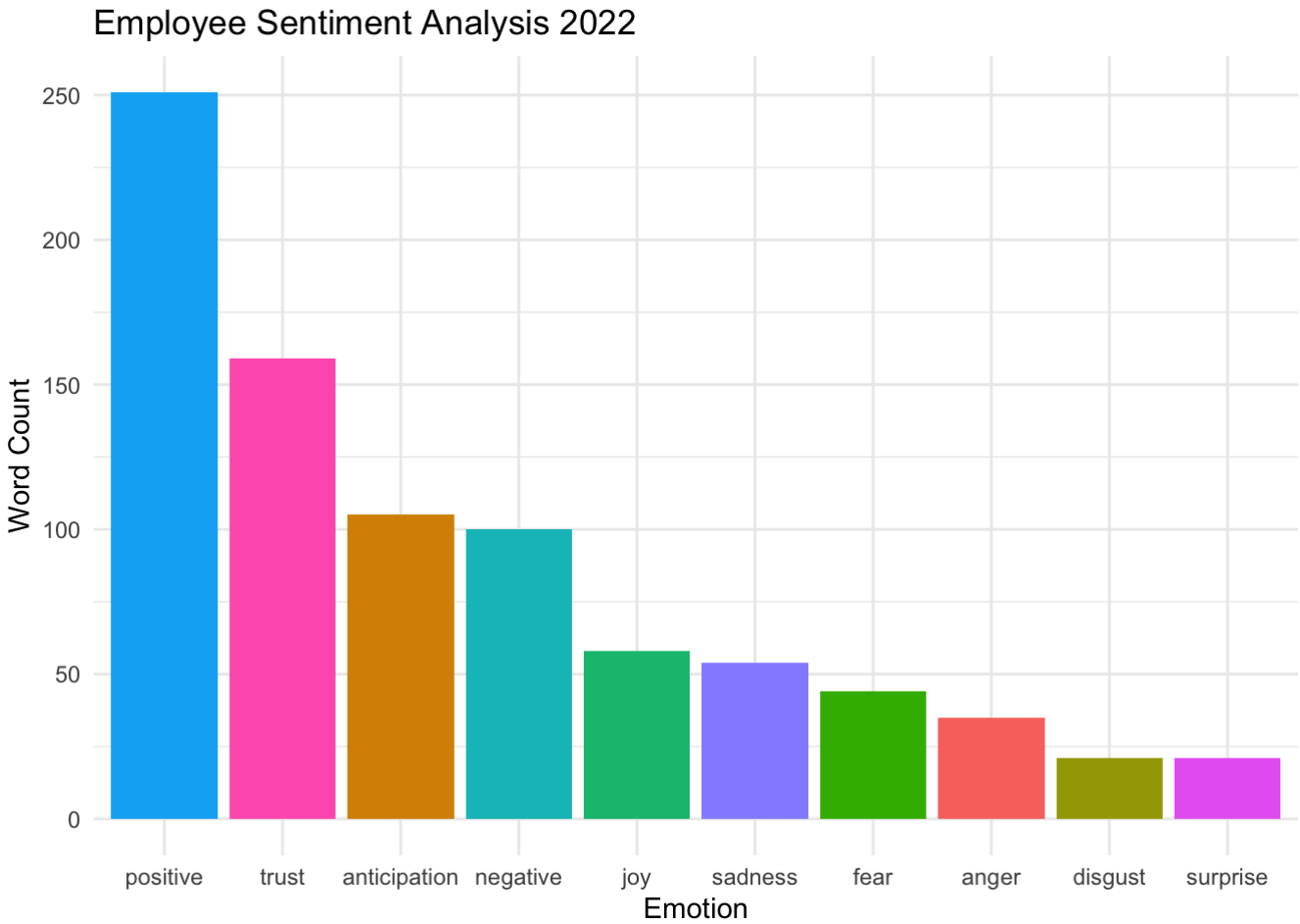
Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.