Sentiment Analysis

```r
{r} dfq=read.csv('Sentiment.csv',header = TRUE)
```

```r
{r} library(tm) corpus <- iconv(dfq$Opinion) corpus <- Corpus(VectorSource(corpus))
```

```r
{r} corpus <- tm_map(corpus, tolower) inspect(corpus[1:5])    {r} corpus <- tm_map(corpus,
removePunctuation) inspect(corpus[1:5])
```

```r
{r} cleanset <- tm_map(corpus, removeWords, stopwords('english')) inspect(cleanset[1:5])
```

```r
{r} #cleanset <- tm_map(cleanset, stemDocument) #cleanset <- tm_map(cleanset, stripWhitespace)
inspect(cleanset[1:5])      {r} tdm <- TermDocumentMatrix(cleanset) tdm <- as.matrix(tdm)
tdm[1:314, 1:24]
```

```r
{r} w <- rowSums(tdm) w <- subset(w, w>=4) barplot(w,         las = 2,         col =
rainbow(50)) {r} library(wordcloud2) w <- data.frame(names(w), w) colnames(w) <- c('word',
'freq') wordcloud2(w,            size = 0.7,            shape = 'circle',          rotateRatio
= 0.5,          minSize = 3) {r} library(wordcloud) w <- sort(rowSums(tdm), decreasing =
TRUE) set.seed(222) wordcloud(words = names(w),          freq = w,          max.words =
150,          random.order = F,          min.freq = 2,          colors = brewer.pal(8,
'Dark2'),          scale = c(5, 0.3),          rot.per = 0.4)
```

Word Association

"'{r} # Find associations findAssocs(tdm, terms = c("health","women","tribal"), corlimit = 0.35)

```r
## sentiment Analysis
```

```r
```{r}
library(syuzhet)
library(lubridate)
library(ggplot2)
library(scales)
library(reshape2)
library(dplyr)
```

"'{r} Open_response <- iconv(dfq$Opinion)

```r
```{r}
s <- get_nrc_sentiment(Open_response)
head(s)
```

"'{r} barplot(colSums(s), las = 2, col = rainbow(10), ylab = 'Count', main = 'Sentiment Scores of Open Response')

```r
```{r}
#transpose
td<-data.frame(t(s))
#The function rowSums computes column sums across rows for each level of a grouping variable.
td_new <- data.frame(rowSums(td))
#Transformation and cleaning
names(td_new)[1] <- "count"
td_new <- cbind("sentiment" = rownames(td_new), td_new)
rownames(td_new) <- NULL
```

```
td_new2<-td_new[1:10,]
#Plot One - count of words associated with each sentiment
quickplot(sentiment, data=td_new2, weight=count, geom="bar", fill=sentiment, ylab="count")+ggtitle("Sur
```

"'{r} hist(D)

```{r}
#Plot two - count of words associated with each sentiment, expressed as a percentage
barplot(
  sort(colSums(prop.table(s[, 1:10]))*100),
  horiz = TRUE,
  cex.names = 0.7,
  las = 1,
  main = "Emotions in Open-ended Responses", xlab="Percentage"
)
```

This bar plot allows for a quick and easy comparison of the proportion of words associated with each emotion in the text. The emotion "trust" has the longest bar and shows that words associated with this positive emotion constitute just over 35% of all the meaningful words in this text. On the other hand, the emotion of "disgust" has the shortest bar and shows that words associated with this negative emotion constitute less than 2% of all the meaningful words in this text. Overall, words associated with the positive emotions of "trust" and "joy" account for almost 60% of the meaningful words in the text, which can be interpreted as a good sign of team health.