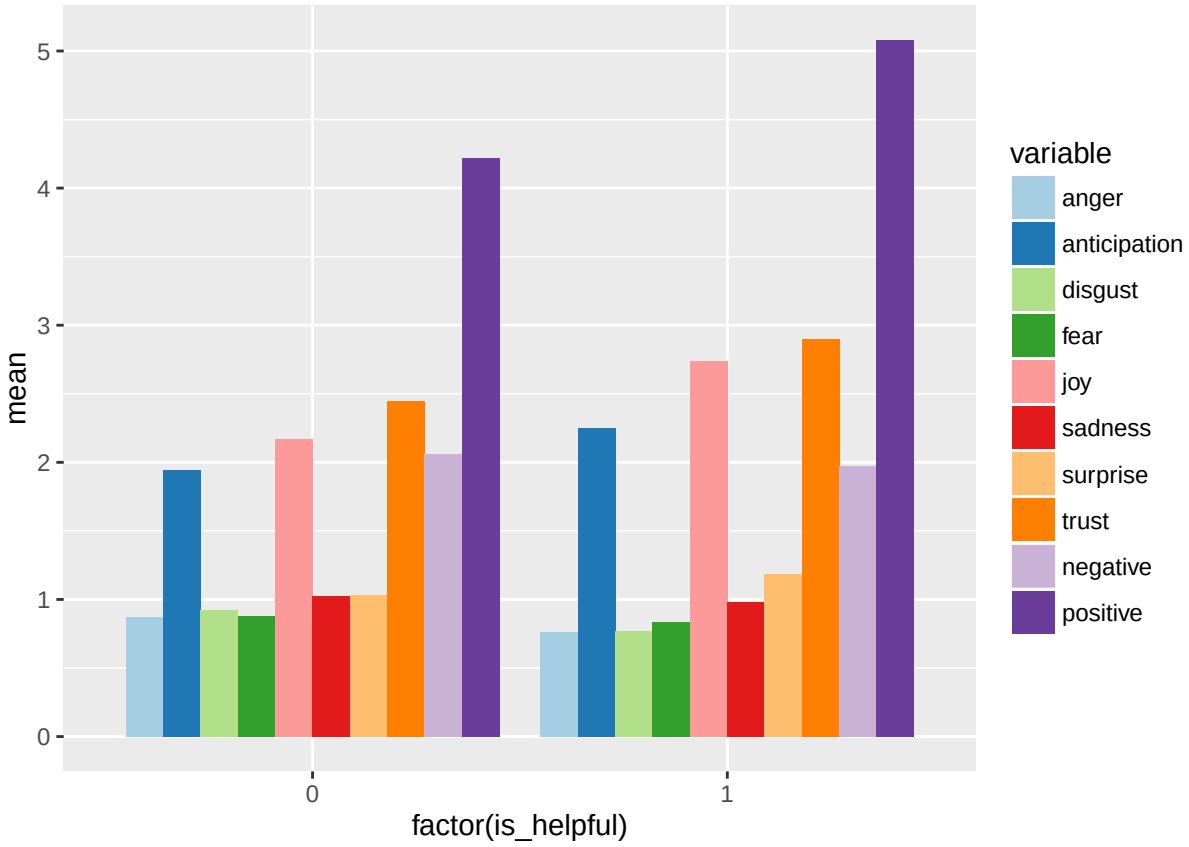
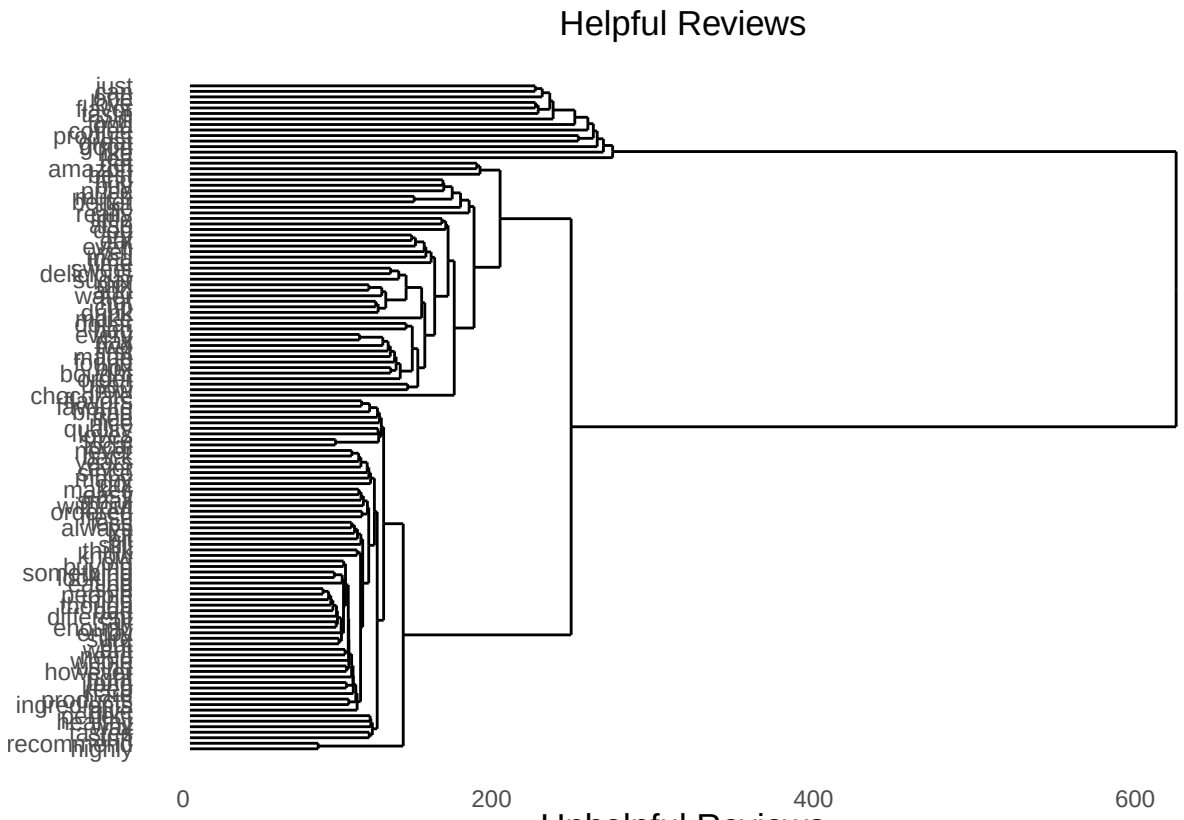
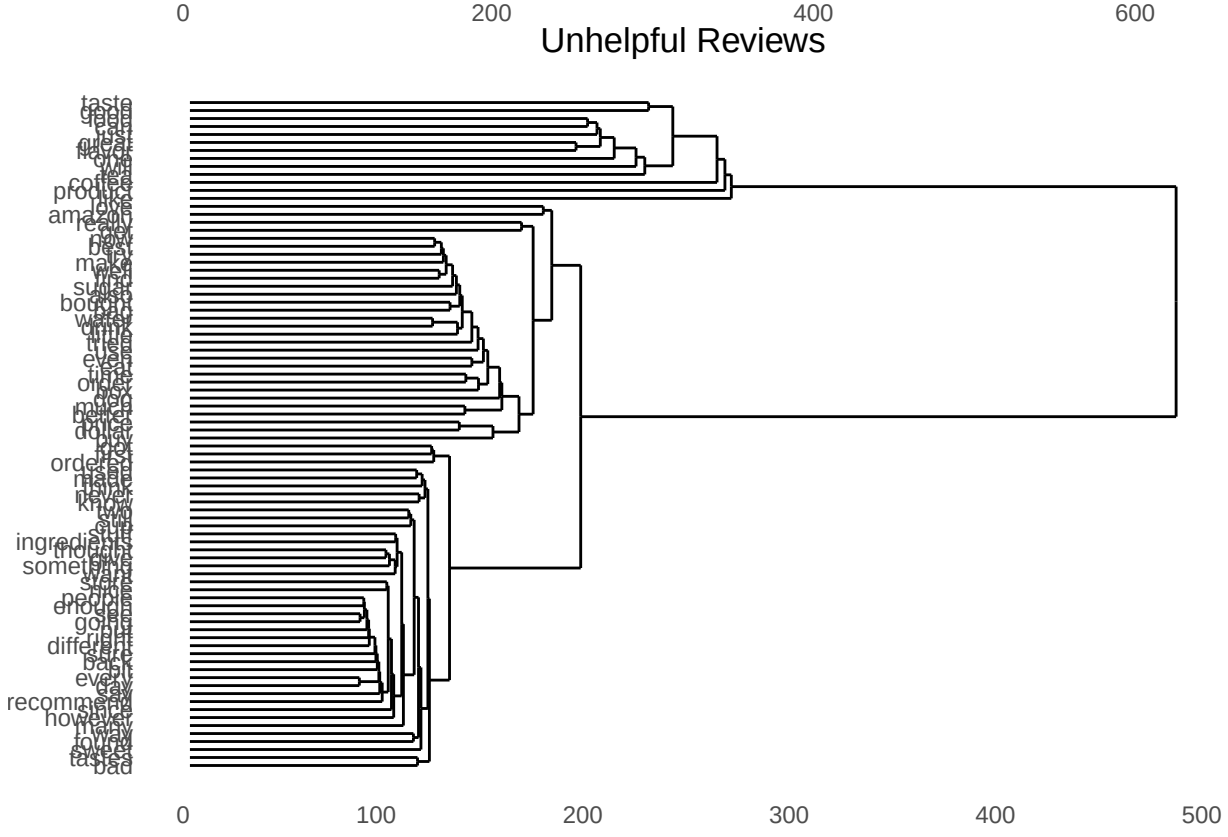
Visualize


mean
5
4
3
2
1
0
0
1
factor(is_helpful)
variable
anger
anticipation
disgust
fear
joy
sadness
surprise
trust
negative
positive






```
## Warning in wordcloud(names(helpful_frequency), helpful_frequency, max.words
## = 25, : order could not be fit on page. It will not be plotted.
```

```
## Warning in wordcloud(names(helpful_frequency), helpful_frequency, max.words
## = 25, : give could not be fit on page. It will not be plotted.
```

```
## Warning in wordcloud(names(helpful_frequency), helpful_frequency, max.words
## = 25, : cat could not be fit on page. It will not be plotted.
```

Unhelpful Reviews

Helpful Reviews


```
helpful_df <- data.frame(word = names(helpful_frequency),
                         count = helpful_frequency) %>%
  arrange(desc(count)) %>%
  top_n(20)
```

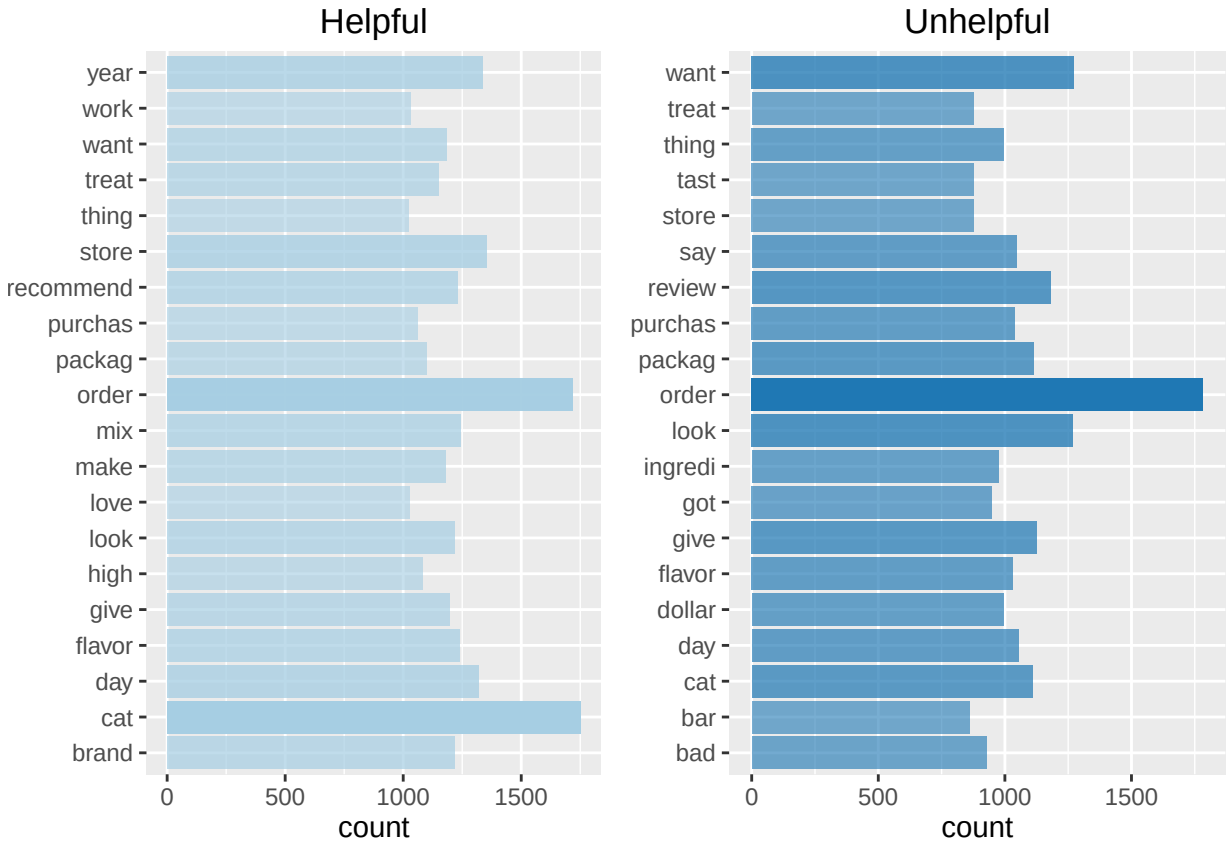
```
## Selecting by count
```

```
unhelpful_df <- data.frame(word = names(unhelpful_frequency),
                           count = unhelpful_frequency) %>%
  arrange(desc(count)) %>%
  top_n(20)
```

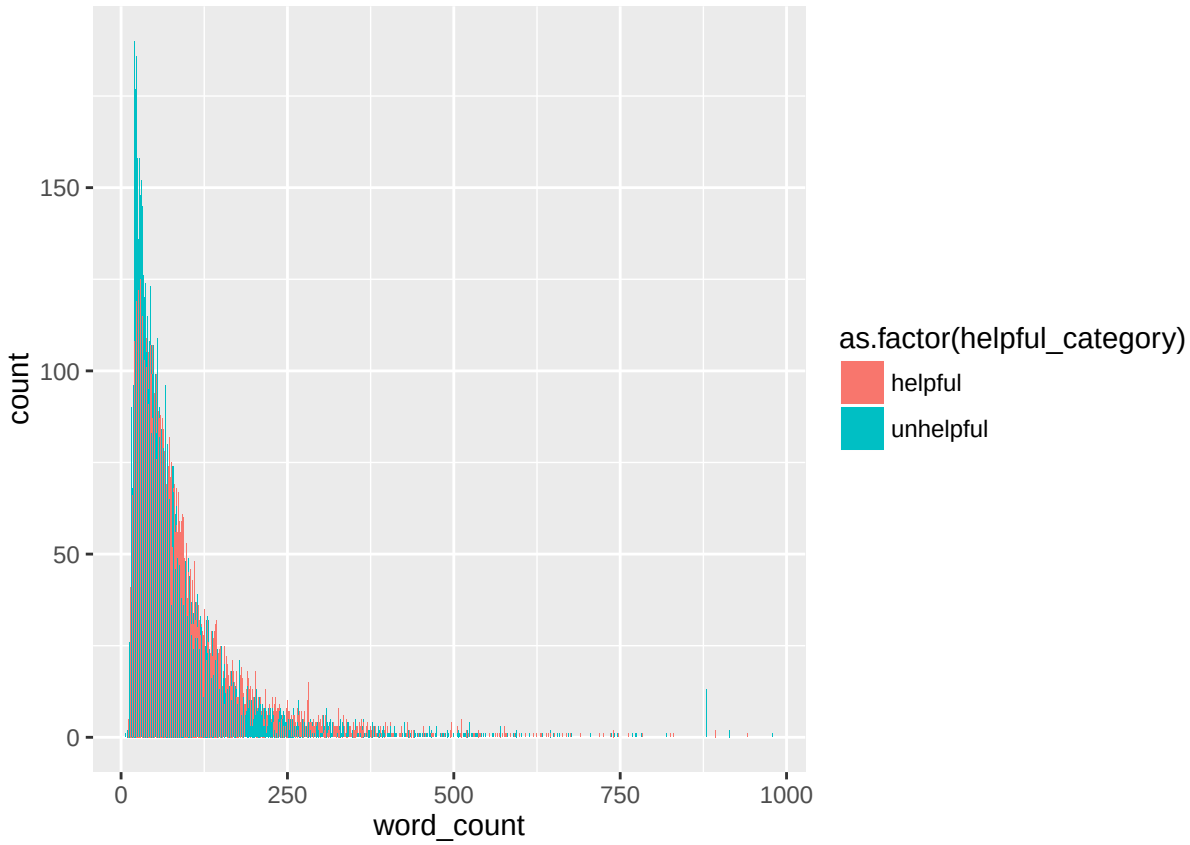
```
## Selecting by count
```

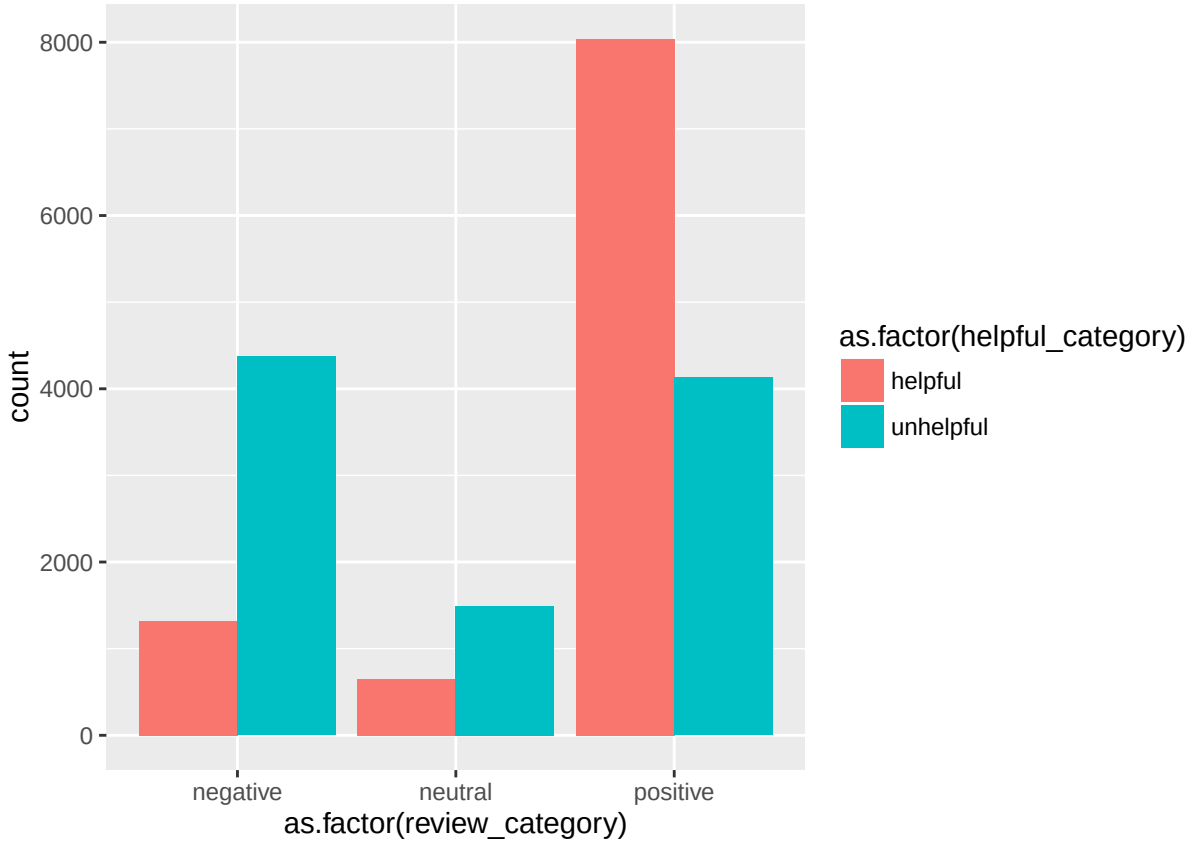
```
helpful_word_plot <- ggplot(helpful_df,
                            aes(x = word, y = count)) +
  geom_bar(aes(alpha = count),
           stat="identity",
           fill = color_palette[[1]]) +
```

```
  coord_flip() +
  labs(x = NULL, y = "count", title = "Helpful") +
  scale_alpha_continuous(range = c(0.6, 1), guide = FALSE)

unhelpful_word_plot <- ggplot(unhelpful_df,
                              aes(x = word, y = count)) +
  geom_bar(aes(alpha = count),
           stat="identity",
           fill = color_palette[[2]]) +
  coord_flip() +
  labs(x = NULL, y = "count", title = "Unhelpful") +
  scale_alpha_continuous(range = c(0.6, 1), guide = FALSE)

grid.arrange(helpful_word_plot, unhelpful_word_plot, ncol = 2)
```

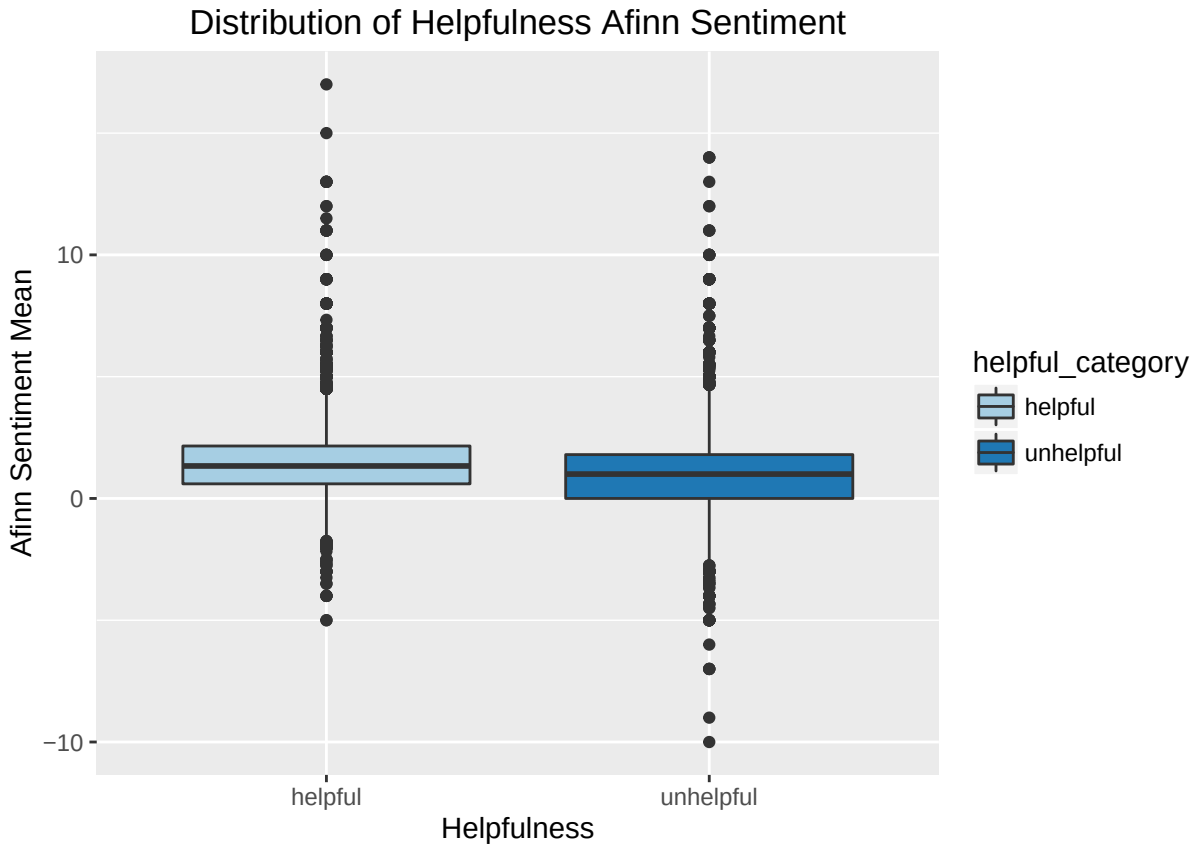



```
ggplot(reviews %>% filter(word_count < 1000),
       aes(word_count, fill = as.factor(helpful_category))) +
  geom_bar(position = "dodge")
```

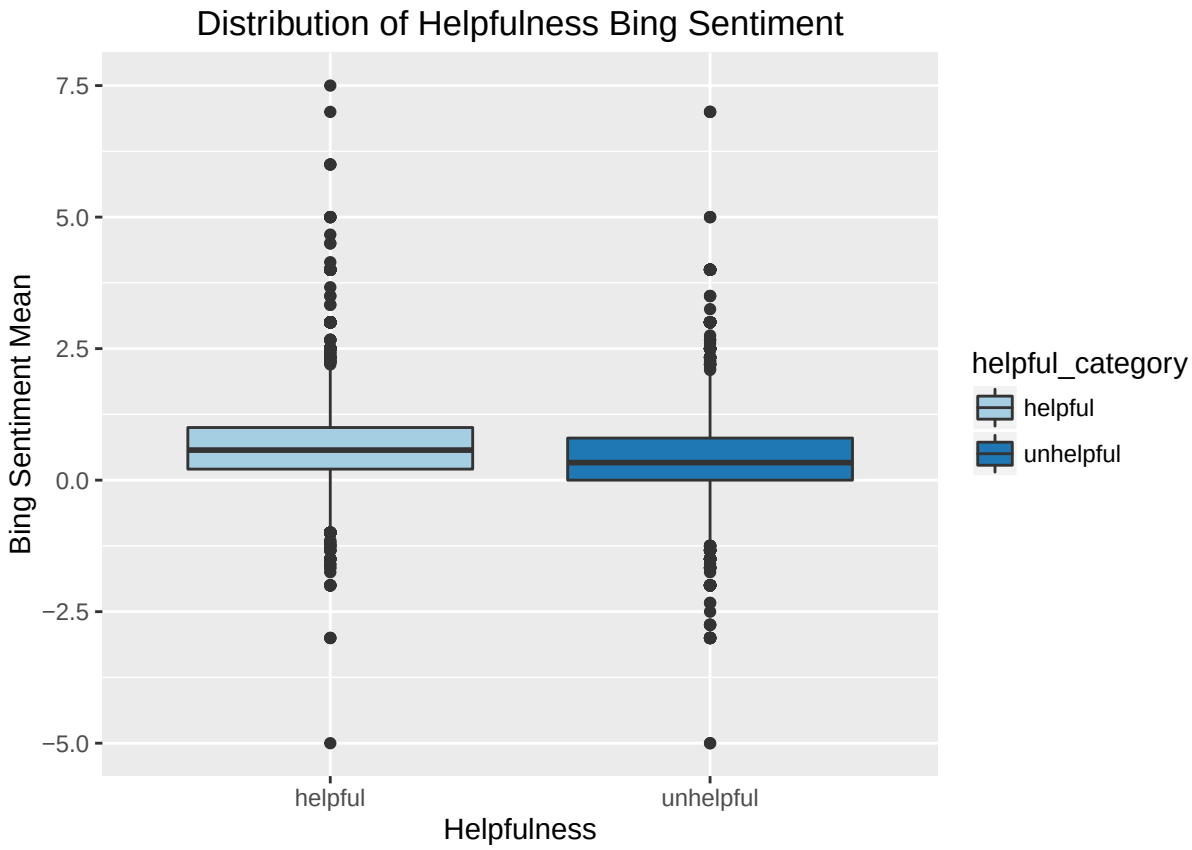

count
150
100
50
0
0
250
500
750
1000
word_count
as.factor(helpful_category)
helpful
unhelpful


count
8000
6000
4000
2000
0
negative
neutral
positive
as.factor(review_category)
as.factor(helpful_category)
helpful
unhelpful

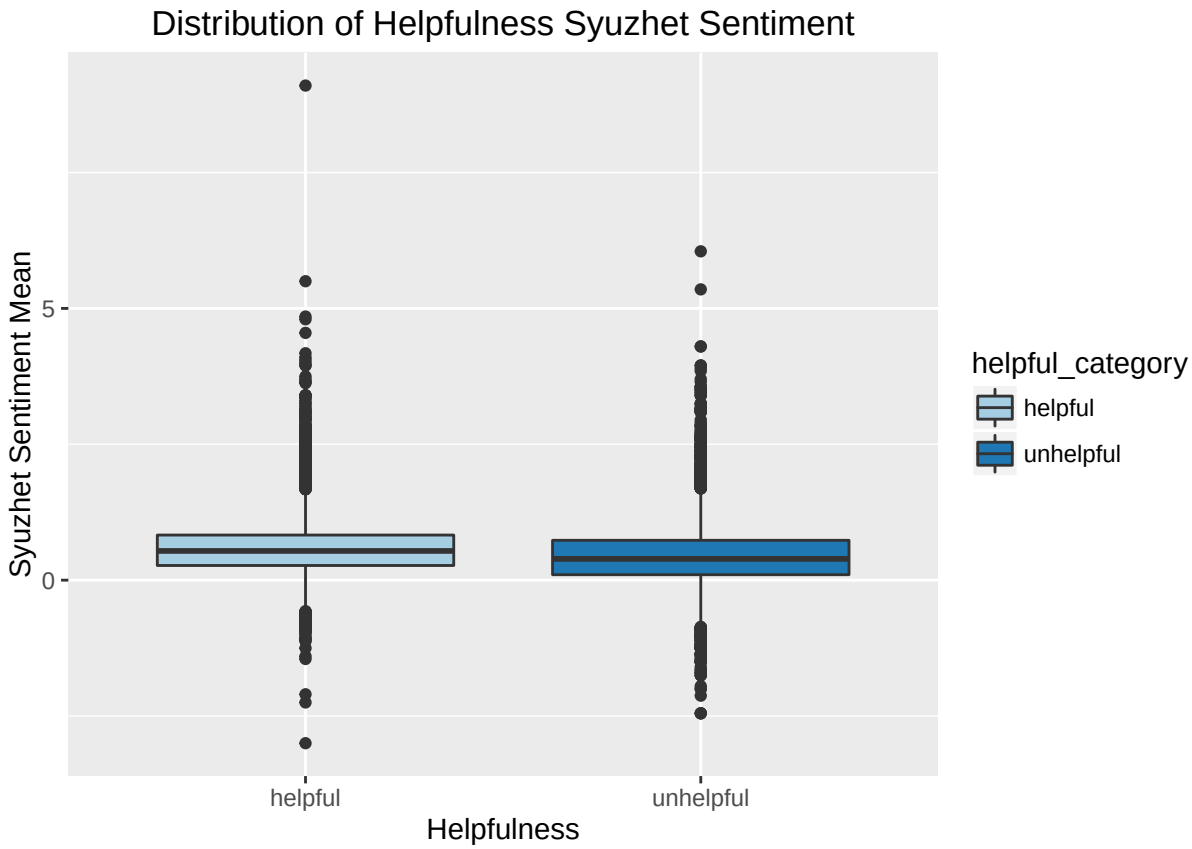
```
ggplot(reviews,
    aes(x = helpful_category,
        y = afinn_sentiment_mean,
        fill = helpful_category)) +
    geom_boxplot() +
    labs(x = "Helpfulness",
         y = "Afinn Sentiment Mean",
         title = "Distribution of Helpfulness Afinn Sentiment") +
    scale_fill_manual(values = color_palette)
```




```
ggplot(reviews,
    aes(x = helpful_category,
        y = bing_sentiment_mean,
        fill = helpful_category)) +
    geom_boxplot() +
    labs(x = "Helpfulness",
         y = "Bing Sentiment Mean",
         title = "Distribution of Helpfulness Bing Sentiment") +
    scale_fill_manual(values = color_palette)
```




```
ggplot(reviews,
    aes(x = helpful_category,
        y = sentiment_mean,
        fill = helpful_category)) +
    geom_boxplot() +
    labs(x = "Helpfulness",
         y = "Syuzhet Sentiment Mean",
         title = "Distribution of Helpfulness Syuzhet Sentiment") +
    scale_fill_manual(values = color_palette)
```


Distribution of Helpfulness Syuzhet Sentiment
Syuzhet Sentiment Mean
5
0
helpful
unhelpful
Helpfulness
helpful_category
helpful
unhelpful