“Not” isn’t the only term that provides some context for the following word. We could pick four common words (or more) that negate the subsequent term, and use the same joining and counting approach to examine all of them at once.

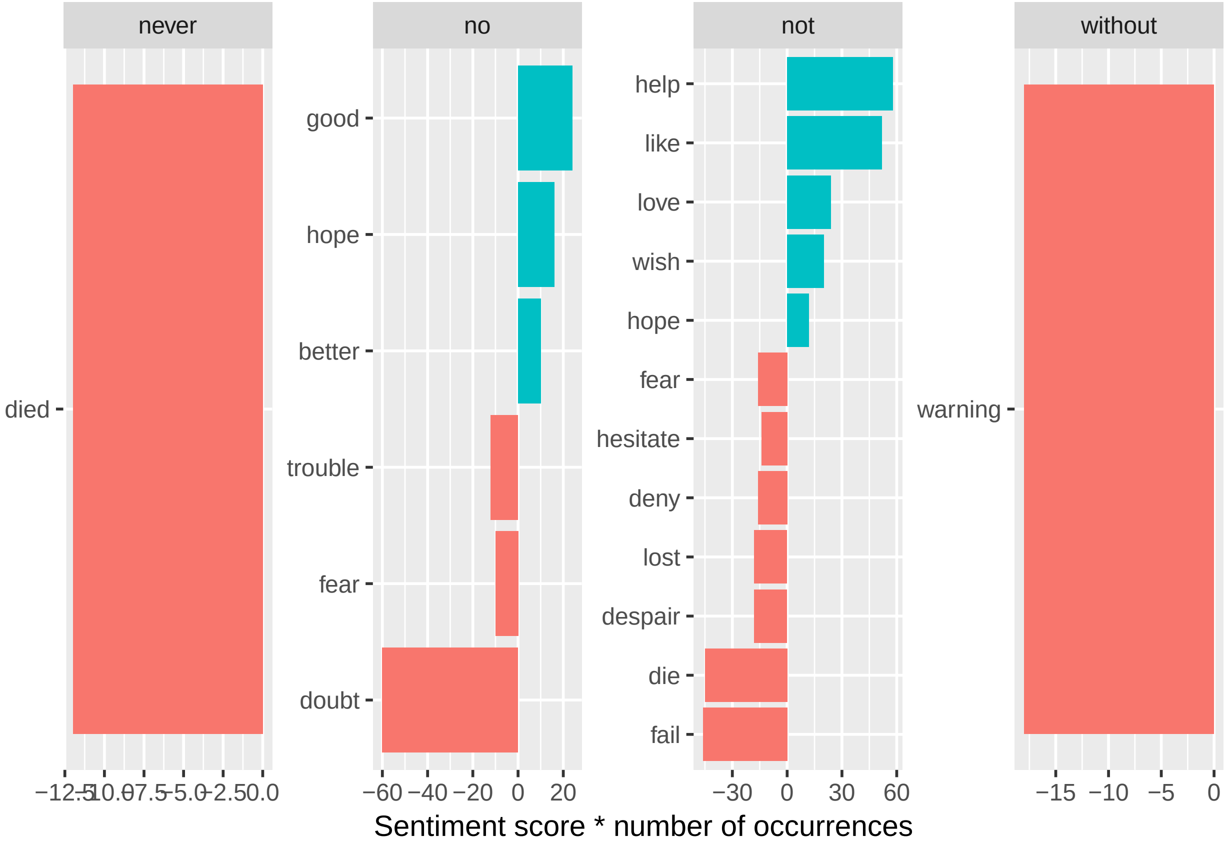
```
negation_words <- c("not", "no", "never", "without")

negated_words<-bigrams_separated %>%
  filter(word1 %in% negation_words) %>%
  inner_join(AFINN, by = c(word2 = "word")) %>%
  count(word1, word2, score, sort = TRUE)
head(negated_words)
```

```
## # A tibble: 6 x 4
##   word1 word2 score     n
##   <chr> <chr> <int> <int>
## 1 no    doubt    -1    60
## 2 not   help      2    29
## 3 not   like      2    26
## 4 not   fail     -2    23
## 5 not   wish      1    20
## 6 not   die      -3    15
```

```
negated_words$word1<-as.factor(negated_words$word1)
unique(negated_words$word1)
```

```
## [1] no      not     never   without
## Levels: never no not without
```

```
negated_words %>%
  mutate(contribution = n * score) %>%
  arrange(desc(abs(contribution))) %>%
  head(20) %>%
  mutate(word2 = reorder(word2, contribution)) %>%
  ggplot(aes(word2, n * score, fill = n * score > 0)) +
  geom_col(show.legend = FALSE) +
  labs(x = NULL, y = "Sentiment score * number of occurrences")+
  facet_wrap(~ word1,ncol =4,scales="free")+
  coord_flip()
```



"not doubt" and "not help" are the two most common examples, we can also see pairings such as "no hope" and "never forget." We could combine this to reverse the AFINN scores of each word that follows a negation.

2: Compare Afinn, Bing with NRC

```
# Keep words that have been classified within the NRC lexicon.
get_sentiments('afinn')
```

```
## # A tibble: 2,476 x 2
##    word       score
##    <chr>      <int>
##  1 abandon       -2
##  2 abandoned     -2
##  3 abandons      -2
##  4 abducted      -2
##  5 abduction     -2
##  6 abductions    -2
##  7 abhor         -3
##  8 abhorred      -3
##  9 abhorrent     -3
## 10 abhors        -3
## # ... with 2,466 more rows
```

```
sentiments_afinn <- inner_join(spooky_wrd, get_sentiments('afinn'), by = "word")
head(sentiments_afinn)
```

```
##        id author      word score
## 1 id26305    EAP        no    -1
## 2 id26305    EAP perfectly     3
## 3 id17569    HPL   mistake    -2
## 4 id11008    EAP   cutting    -1
## 5 id11008    EAP fantastic     4
## 6 id11008    EAP  greatest     3
```

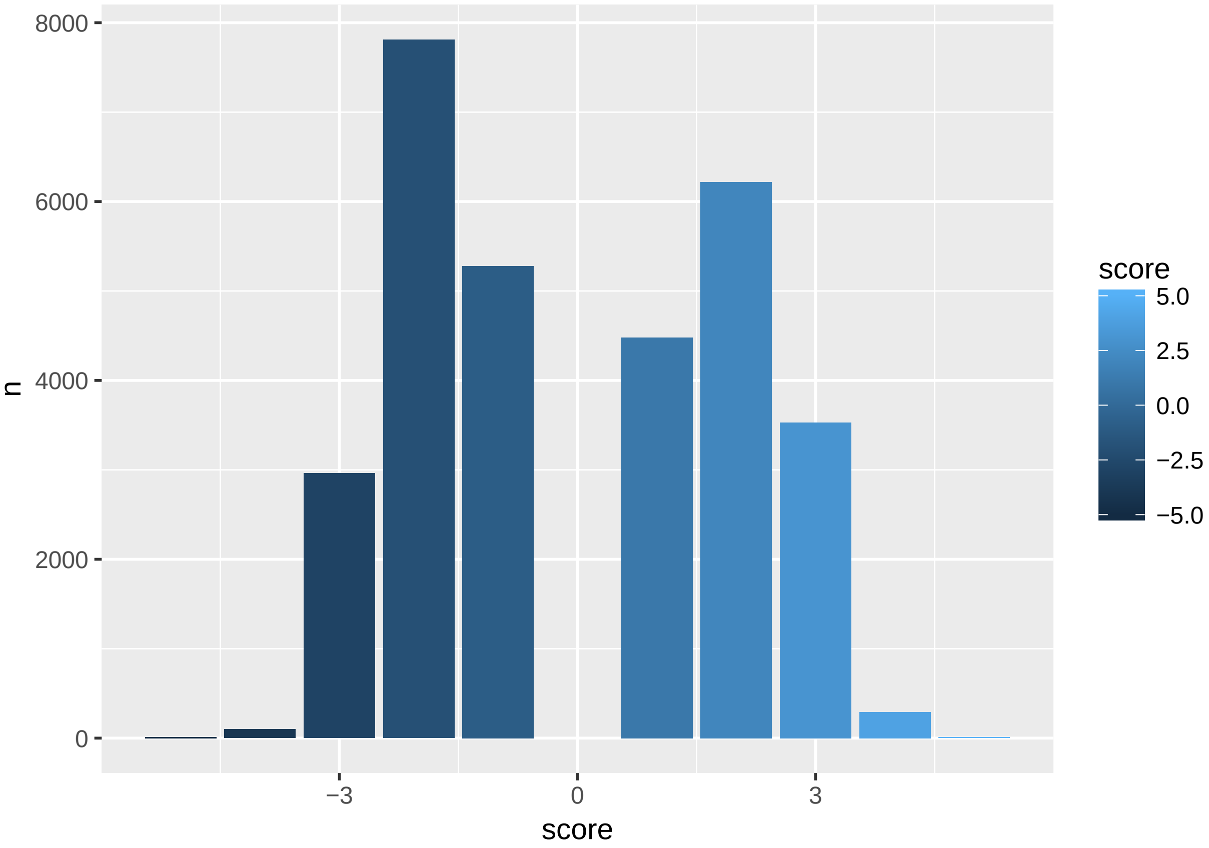
```r
count(sentiments_afinn, score)
```

```
## # A tibble: 10 x 2
##    score     n
##    <int> <int>
##  1    -5    12
##  2    -4    98
##  3    -3  2961
##  4    -2  7810
##  5    -1  5280
##  6     1  4479
##  7     2  6220
##  8     3  3529
##  9     4   291
## 10     5     9
```

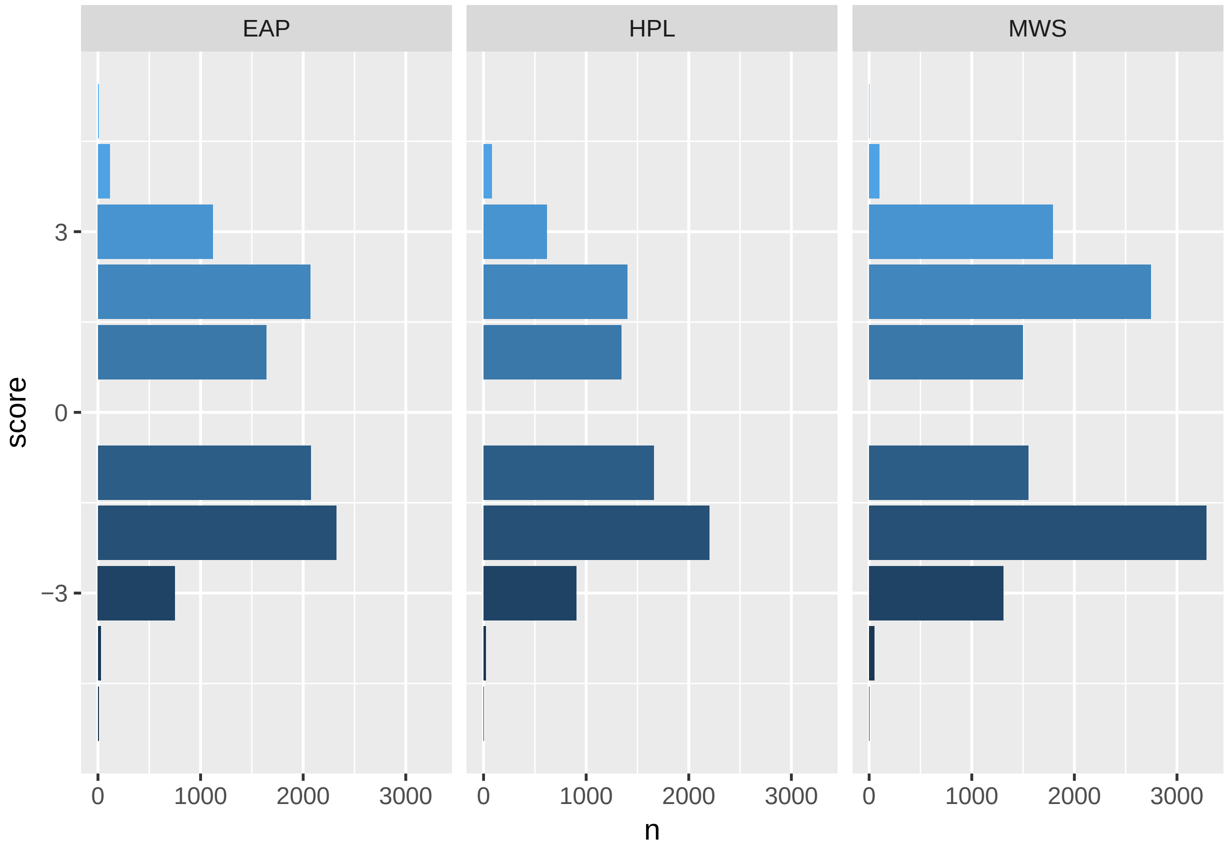
```r
count(sentiments_afinn, author, score)
```

```
## # A tibble: 29 x 3
##    author score     n
##    <chr>  <int> <int>
##  1 EAP       -5     9
##  2 EAP       -4    28
##  3 EAP       -3   751
##  4 EAP       -2  2321
##  5 EAP       -1  2072
##  6 EAP        1  1639
##  7 EAP        2  2071
##  8 EAP        3  1121
##  9 EAP        4   113
## 10 EAP        5     7
## # ... with 19 more rows
```

```r
ggplot(count(sentiments_afinn, score)) +
  geom_col(aes(score, n, fill = score))
```



```
ggplot(count(sentiments_afinn, author, score)) +
  geom_col(aes(score, n, fill = score)) +
  facet_wrap(~ author) +
  coord_flip() +
  theme(legend.position = "none")
```



```
get_sentiments('bing')
```

```
## # A tibble: 6,788 x 2
##    word        sentiment
##    <chr>       <chr>
##  1 2-faced     negative
##  2 2-faces     negative
##  3 a+          positive
##  4 abnormal    negative
##  5 abolish     negative
##  6 abominable  negative
##  7 abominably  negative
##  8 abominate   negative
##  9 abomination negative
## 10 abort       negative
## # ... with 6,778 more rows
```

```
sentiments_bing<- inner_join(spooky_wrd, get_sentiments('bing'), by = "word")
head(sentiments_bing)
```

```
##        id author        word sentiment
## 1 id26305    EAP     dungeon  negative
## 2 id26305    EAP   perfectly  positive
## 3 id17569    HPL     mistake  negative
## 4 id11008    EAP        gold  positive
## 5 id11008    EAP   fantastic  positive
## 6 id11008    EAP incessantly  negative
```

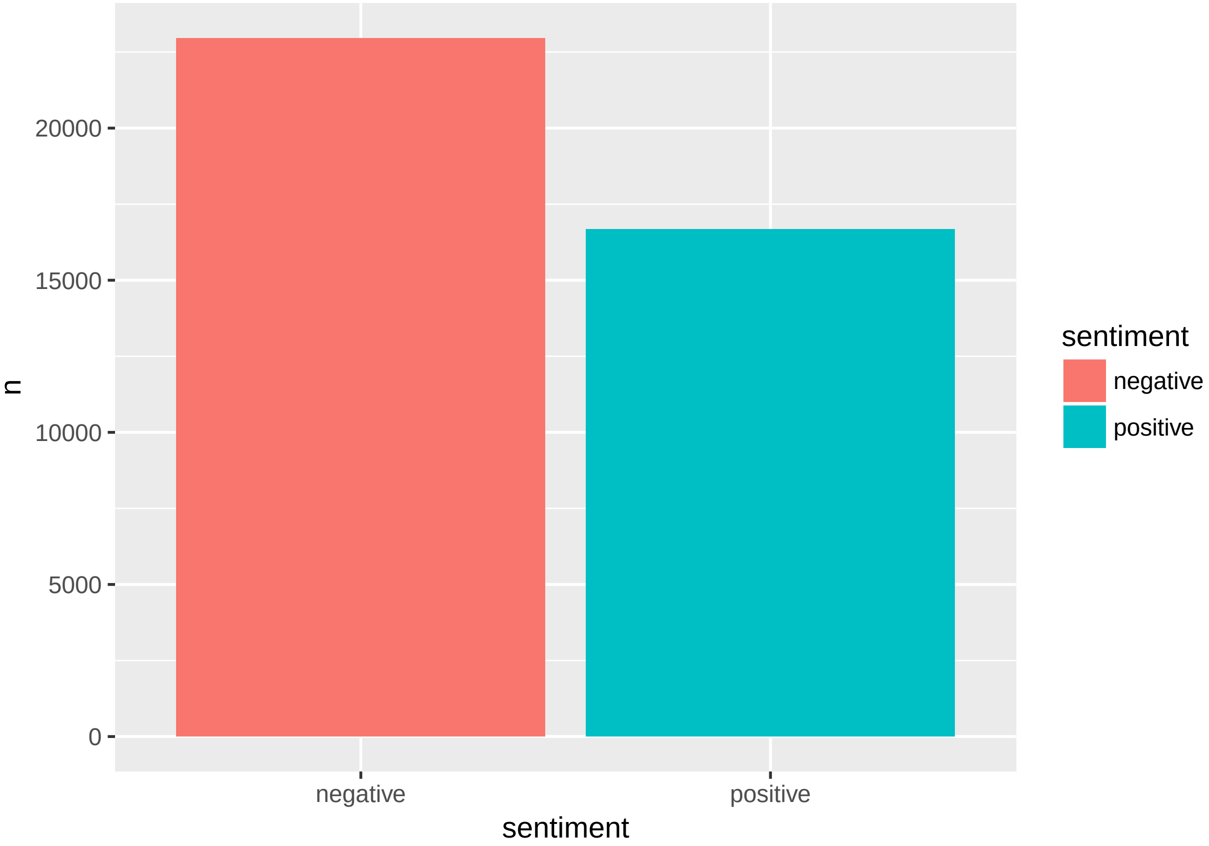
```r
count(sentiments_bing,sentiment)
```

```
## # A tibble: 2 x 2
##   sentiment     n
##   <chr>     <int>
## 1 negative  22958
## 2 positive  16674
```

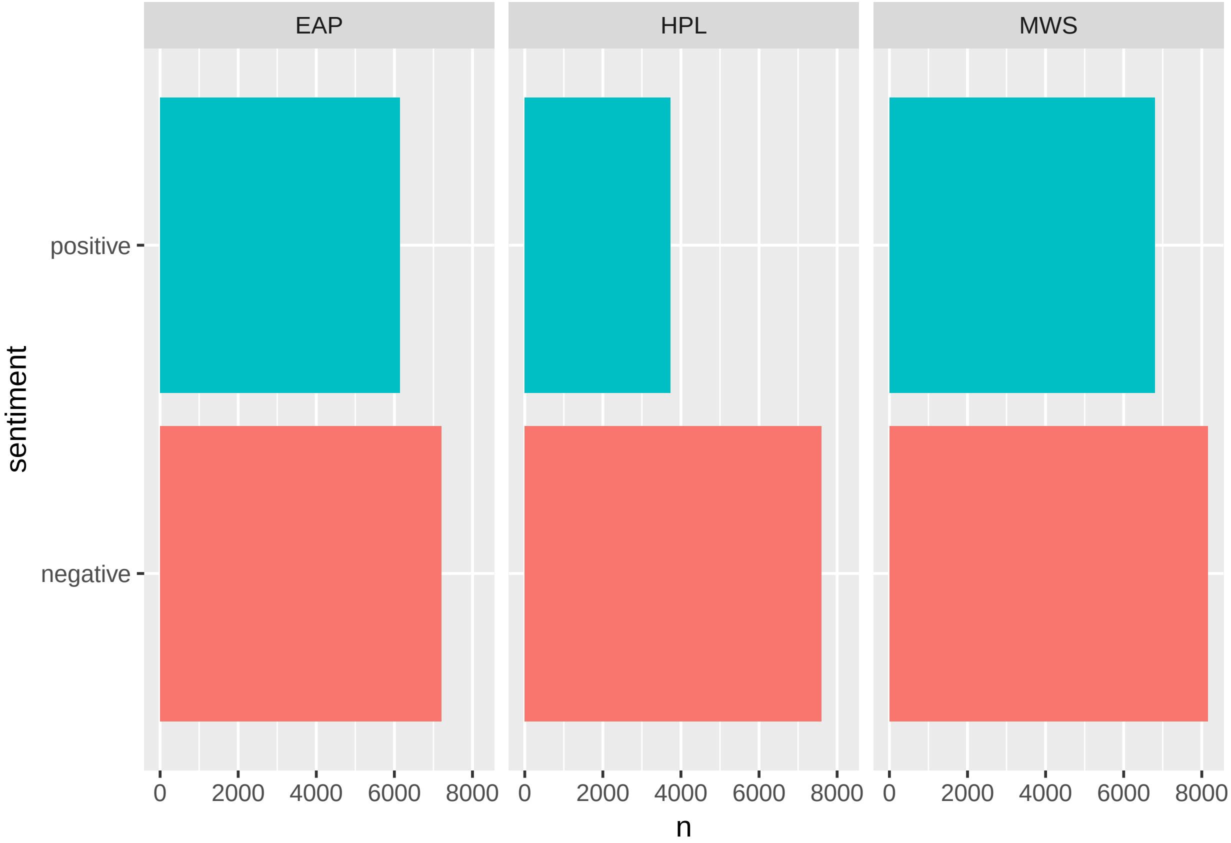
```r
count(sentiments_bing,author,sentiment)
```

```
## # A tibble: 6 x 3
##   author sentiment     n
##   <chr>  <chr>     <int>
## 1 EAP    negative   7203
## 2 EAP    positive   6144
## 3 HPL    negative   7605
## 4 HPL    positive   3731
## 5 MWS    negative   8150
## 6 MWS    positive   6799
```

```r
ggplot(count(sentiments_bing,sentiment)) +
  geom_col(aes(sentiment, n, fill = sentiment))
```



```r
ggplot(count(sentiments_bing, author, sentiment)) +
  geom_col(aes(sentiment, n, fill = sentiment)) +
  facet_wrap(~ author) +
  coord_flip() +
  theme(legend.position = "none")
```



```r
get_sentiments('nrc')
```

```
## # A tibble: 13,901 x 2
##    word        sentiment
##    <chr>       <chr>
##  1 abacus      trust
##  2 abandon     fear
##  3 abandon     negative
##  4 abandon     sadness
##  5 abandoned   anger
##  6 abandoned   fear
##  7 abandoned   negative
##  8 abandoned   sadness
##  9 abandonment anger
## 10 abandonment fear
## # ... with 13,891 more rows
```

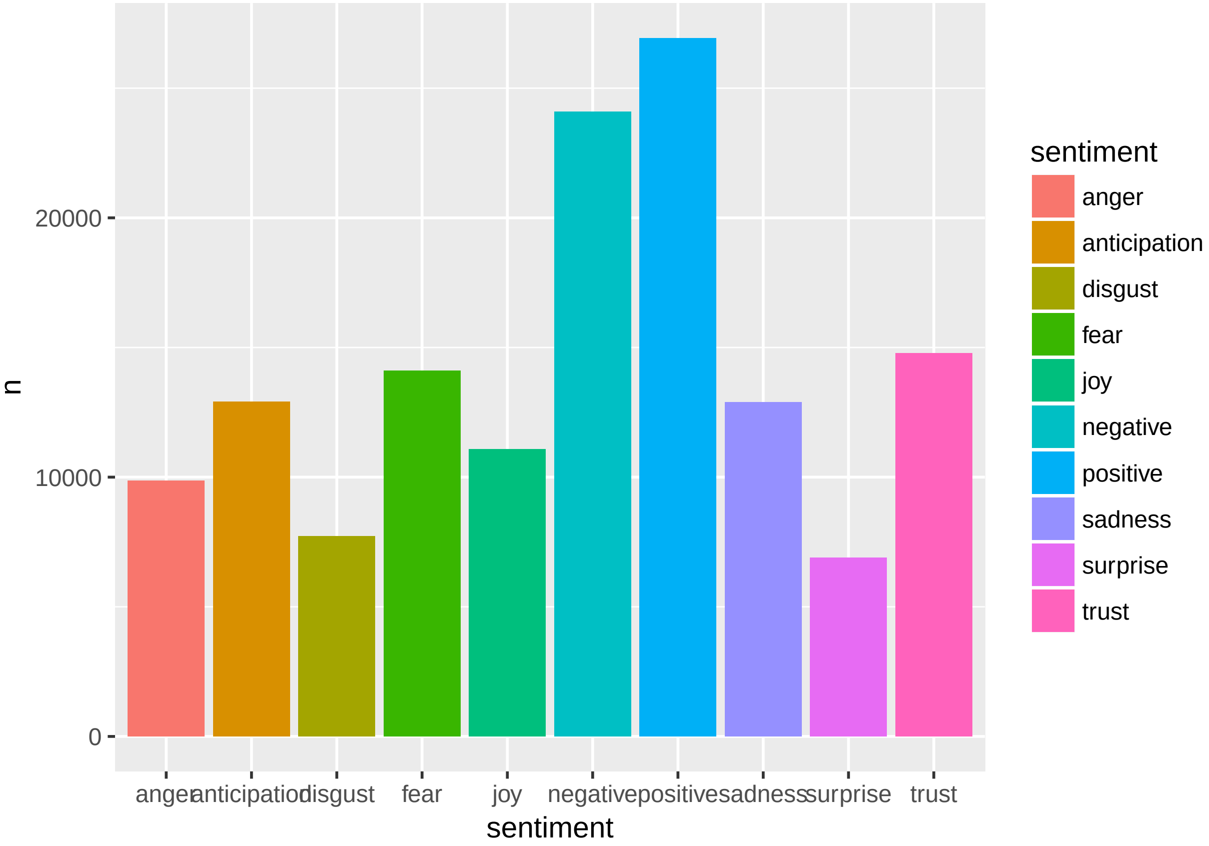
```r
sentiments <- inner_join(spooky_wrd, get_sentiments('nrc'), by = "word")

count(sentiments, sentiment)
```

```
## # A tibble: 10 x 2
##    sentiment        n
##    <chr>        <int>
##  1 anger         9869
##  2 anticipation 12912
##  3 disgust       7731
##  4 fear         14096
##  5 joy          11077
##  6 negative     24084
##  7 positive     26934
##  8 sadness      12896
##  9 surprise      6903
## 10 trust        14777
```

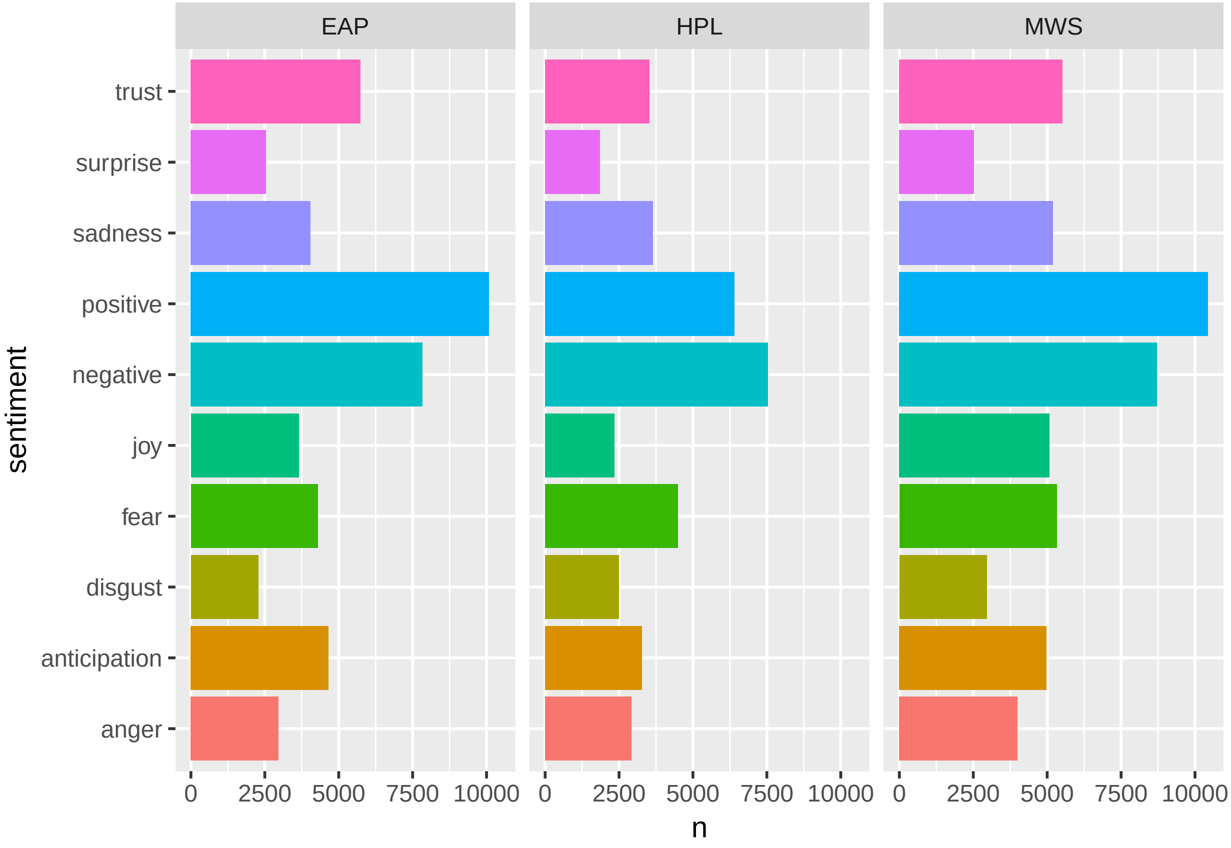
```
count(sentiments, author, sentiment)
```

```
## # A tibble: 30 x 3
##    author sentiment        n
##    <chr>  <chr>        <int>
##  1 EAP    anger         2962
##  2 EAP    anticipation  4656
##  3 EAP    disgust       2273
##  4 EAP    fear          4287
##  5 EAP    joy           3652
##  6 EAP    negative      7833
##  7 EAP    positive     10083
##  8 EAP    sadness       4045
##  9 EAP    surprise      2538
## 10 EAP    trust         5739
## # ... with 20 more rows
```

```
ggplot(count(sentiments, sentiment)) +
  geom_col(aes(sentiment, n, fill = sentiment))
```



```
ggplot(count(sentiments, author, sentiment)) +
  geom_col(aes(sentiment, n, fill = sentiment)) +
  facet_wrap(~ author) +
  coord_flip() +
  theme(legend.position = "none")
```


EAP
HPL
MWS
sentiment
trust
surprise
sadness
positive
negative
joy
fear
disgust
anticipation
anger
n