

Sentiments of World News

[Code ▾](#)[Code](#)[Code](#)

First, I pulled 5000 of the most recent tweets from ten news organizations, both local and abroad, and broke them down into their wording in order to correlate the verbiage with their sentiments.

[Hide](#)

```
nyt_words %>%
  anti_join(get_stopwords()) %>%
  filter(!word == "https") %>%
  filter(!word == "t.co") %>%
  count(word, sort = T)
```

Joining, by = "word"

word	n
<chr>	<int>
coronavirus	1132
new	623
people	395
pandemic	376
u.s	367
said	356
york	291
one	285
now	233
president	220

1-10 of 15,246 rows

Previous **1** 2 3 4 5 6 ... 100 Next

[Hide](#)

```
wsj_words %>%
  anti_join(get_stopwords()) %>%
  filter(!word == "https") %>%
  filter(!word == "t.co") %>%
  count(word, sort = T)
```

Joining, by = "word"

word	n
<chr>	<int>
coronavirus	1031
pandemic	449
new	427
u.s	337
wsjwhatsnow	223
said	188
writes	187
wsjopinion	185
one	175
people	174

1-10 of 12,902 rows Previous **1** 2 3 4 5 6 ... 100 Next

Hide

```
cnn_words %>%  
  anti_join(get_stopwords()) %>%  
  filter(!word == "https") %>%  
  filter(!word == "t.co") %>%  
  count(word, sort = T)
```

Joining, by = "word"

word	n
<chr>	<int>
coronavirus	1360
pandemic	521
new	469
us	447
covid	391
said	385
19	382

president	370
says	347
people	314

1-10 of 12,678 rows Previous **1** 2 3 4 5 6 ... 100 Next

Hide

```
fox_words %>%
  anti_join(get_stopwords()) %>%
  filter(!word == "https") %>%
  filter(!word == "t.co") %>%
  count(word, sort = T)
```

Joining, by = "word"

word <chr>	n <int>
president	260
news	222
potus	219
trump	209
fox	203
democrats	162
midterms2018	154
house	151
border	150
people	149

1-10 of 10,665 rows Previous **1** 2 3 4 5 6 ... 100 Next

Hide

```
wapo_words %>%
  anti_join(get_stopwords()) %>%
  filter(!word == "https") %>%
  filter(!word == "t.co") %>%
  count(word, sort = T)
```

Joining, by = "word"

word	n
<chr>	<int>
coronavirus	906
trump	442
opinion	344
analysis	277
pandemic	228
perspective	227
new	209
u.s	190
covid	187
19	182

1-10 of 10,900 rows

Previous **1** 2 3 4 5 6 ... 100 Next

Hide

```
usa_words %>%  
  anti_join(get_stopwords()) %>%  
  filter(!word == "https") %>%  
  filter(!word == "t.co") %>%  
  count(word, sort = T)
```

Joining, by = "word"

word	n
<chr>	<int>
coronavirus	941
new	273
pandemic	258
said	238
covid	236
19	230
home	210

trump	176
can	174
health	171

1-10 of 12,265 rows Previous **1** 2 3 4 5 6 ... 100 Next

Hide

```
ap_words %>%
  anti_join(get_stopwords()) %>%
  filter(!word == "https") %>%
  filter(!word == "t.co") %>%
  count(word, sort = T)
```

Joining, by = "word"

word	n
<chr>	<int>
coronavirus	1338
new	547
pandemic	482
ap	466
says	421
virus	393
u.s	353
people	270
outbreak	268
home	223

1-10 of 14,556 rows Previous **1** 2 3 4 5 6 ... 100 Next

Hide

```
aje_words %>%
  anti_join(get_stopwords()) %>%
  filter(!word == "https") %>%
  filter(!word == "t.co") %>%
  count(word, sort = T)
```

Joining, by = "word"

word	n
<chr>	<int>
coronavirus	1675
us	427
updates	298
lockdown	290
latest	280
pandemic	262
people	229
covid19	212
new	212
ajopinion	205

1-10 of 11,759 rows

Previous **1** 2 3 4 5 6 ... 100 Next

Hide

```
bbc_words %>%  
  anti_join(get_stopwords()) %>%  
  filter(!word == "https") %>%  
  filter(!word == "t.co") %>%  
  count(word, sort = T)
```

Joining, by = "word"

word	n
<chr>	<int>
coronavirus	1452
us	392
people	253
says	247
lockdown	231
world	192
president	165

new	160
trump	137
outbreak	133
1-10 of 12,577 rows	
Previous	1 2 3 4 5 6 ... 100 Next

Hide

```
reuters_words %>%
  anti_join(get_stopwords()) %>%
  filter(!word == "https") %>%
  filter(!word == "t.co") %>%
  count(word, sort = T)
```

Joining, by = "word"

word	n
<chr>	<int>
coronavirus	819
u.s	385
new	269
covid	241
lockdown	226
19	225
pandemic	198
says	187
first	151
said	148
1-10 of 12,735 rows	
Previous	1 2 3 4 5 6 ... 100 Next

These tables show the words used by each news station, organized by count.

Hide

```
news_words <- cnn_words %>%
  full_join(fox_words) %>%
  full_join(nyt_words) %>%
  full_join(usa_words) %>%
  full_join(wapo_words) %>%
  full_join(wsj_words) %>%
  full_join(aje_words) %>%
  full_join(ap_words) %>%
  full_join(bbc_words) %>%
  full_join(reuters_words)
```

```
Joining, by = c("screen_name", "word", "created_at")
Joining, by = c("screen_name", "word", "created_at")
Joining, by = c("screen_name", "word", "created_at")
Joining, by = c("screen_name", "word", "created_at")
Joining, by = c("screen_name", "word", "created_at")
Joining, by = c("screen_name", "word", "created_at")
Joining, by = c("screen_name", "word", "created_at")
Joining, by = c("screen_name", "word", "created_at")
Joining, by = c("screen_name", "word", "created_at")
```

I combined all of the words utilized by the stations into one overarching group in order to analyze them as a whole.

Hide

```
news_words %>%
  anti_join(get_stopwords()) %>%
  filter(!word == "https") %>%
  filter(!word == "t.co") %>%
  count(word, sort = T)
```

```
Joining, by = "word"
```

word <chr>	n <int>
coronavirus	10711
new	3295
pandemic	2908
people	2129
trump	1965
u.s	1868
says	1836

president	1727
us	1725
covid	1695
1-10 of 68,835 rows	Previous 1 2 3 4 5 6 ... 100 Next

This table shows the most frequently used words, ranked by count. Coronavirus is (unsurprisingly) at the top.

Hide

```
news_words %>%
  anti_join(get_stopwords()) %>%
  filter(!word == "https") %>%
  filter(!word == "t.co") %>%
  count(word, sort = T)%>%
  top_n(1000) %>%
  wordcloud2(size = .6)
```

```
Joining, by = "word"
Selecting by n
```

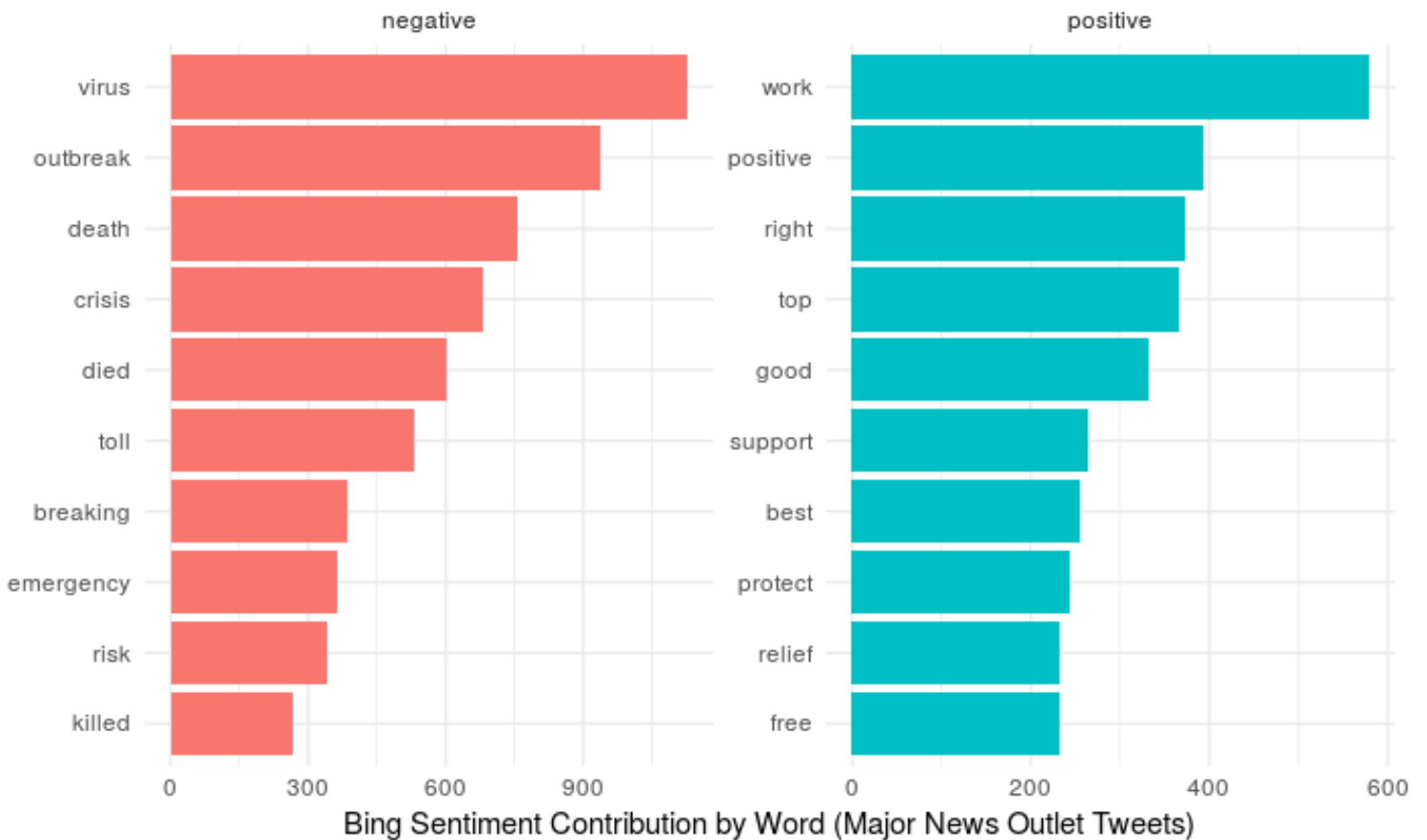


```

news_words %>%
  inner_join(bing_words) %>%
  filter(!word == "trump") %>%
  filter(!word == "like") %>%
  count(word, sentiment, sort = TRUE) %>%
  group_by(sentiment) %>%
  top_n(10) %>%
  ungroup() %>%
  mutate(word = reorder(word, n)) %>%
  ggplot(aes(word, n, fill = sentiment)) +
  geom_col(show.legend = FALSE) +
  facet_wrap(vars(sentiment), scales = "free") +
  labs(y = "Bing Sentiment Contribution by Word (Major News Outlet Tweets)",
       x = NULL) +
  coord_flip() +
  theme_minimal()

```

Joining, by = "word"
 Selecting by n



Here are the top words organized by Bing Sentiment. The pandemic is terrifying. Here is proof.

Hide

```
nrc_words <- get_sentiments("nrc")
```

[Hide](#)

```
nrc_sentiments <- news_words %>%
  inner_join(nrc_words) %>%
  filter(!word == "trump") %>%
  count(word, sentiment, screen_name, sort = TRUE) %>%
  group_by(sentiment) %>%
  top_n(1000) %>%
  ungroup() %>%
  mutate(word = reorder(word, n))
```

Joining, by = "word"
Selecting by n

[Hide](#)

```
sentiments_labeled <- nrc_sentiments
number_of_bar <- nrow(sentiments_labeled)
angle <- 90 - 360 * (nrc_sentiments$n-0.5) / number_of_bar
sentiments_labeled$hjust <- ifelse( angle < -90, 1, 0)
sentiments_labeled$angle <- ifelse(angle < -90, angle+180, angle)
```

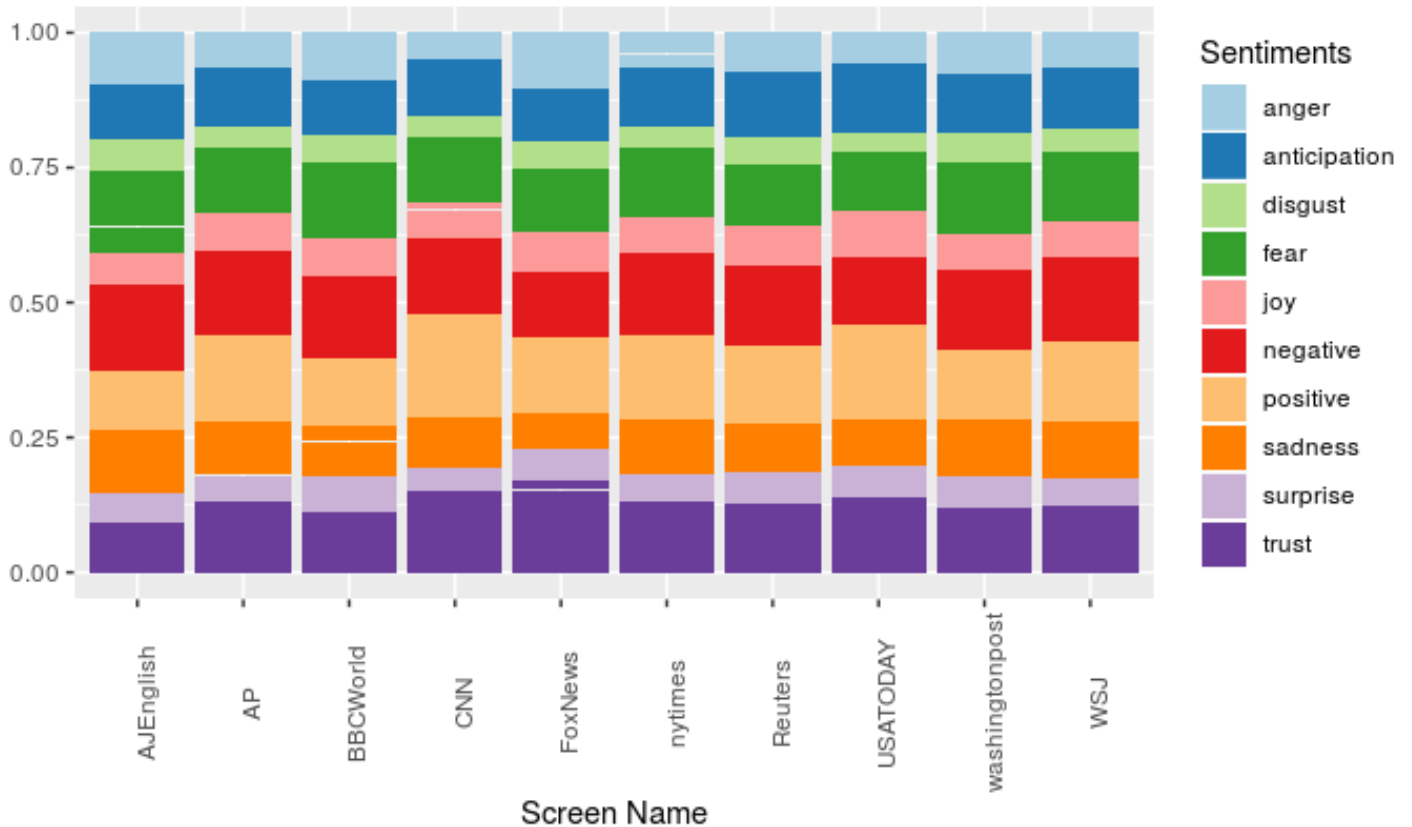
[Hide](#)

```
nrc_sentiments %>%
  ggplot(aes(x = screen_name, y = n , fill = sentiment)) +
  geom_bar(position="fill", stat = "identity") +

  labs(y = "",
       x = "Screen Name",
       title = "NRC Sentiment Contribution by Word",
       fill = "Sentiments") +

  theme_gray() +
  scale_fill_brewer(palette="Paired") +
  theme(axis.text.x = element_text(angle = 90))
```

NRC Sentiment Contribution by Word



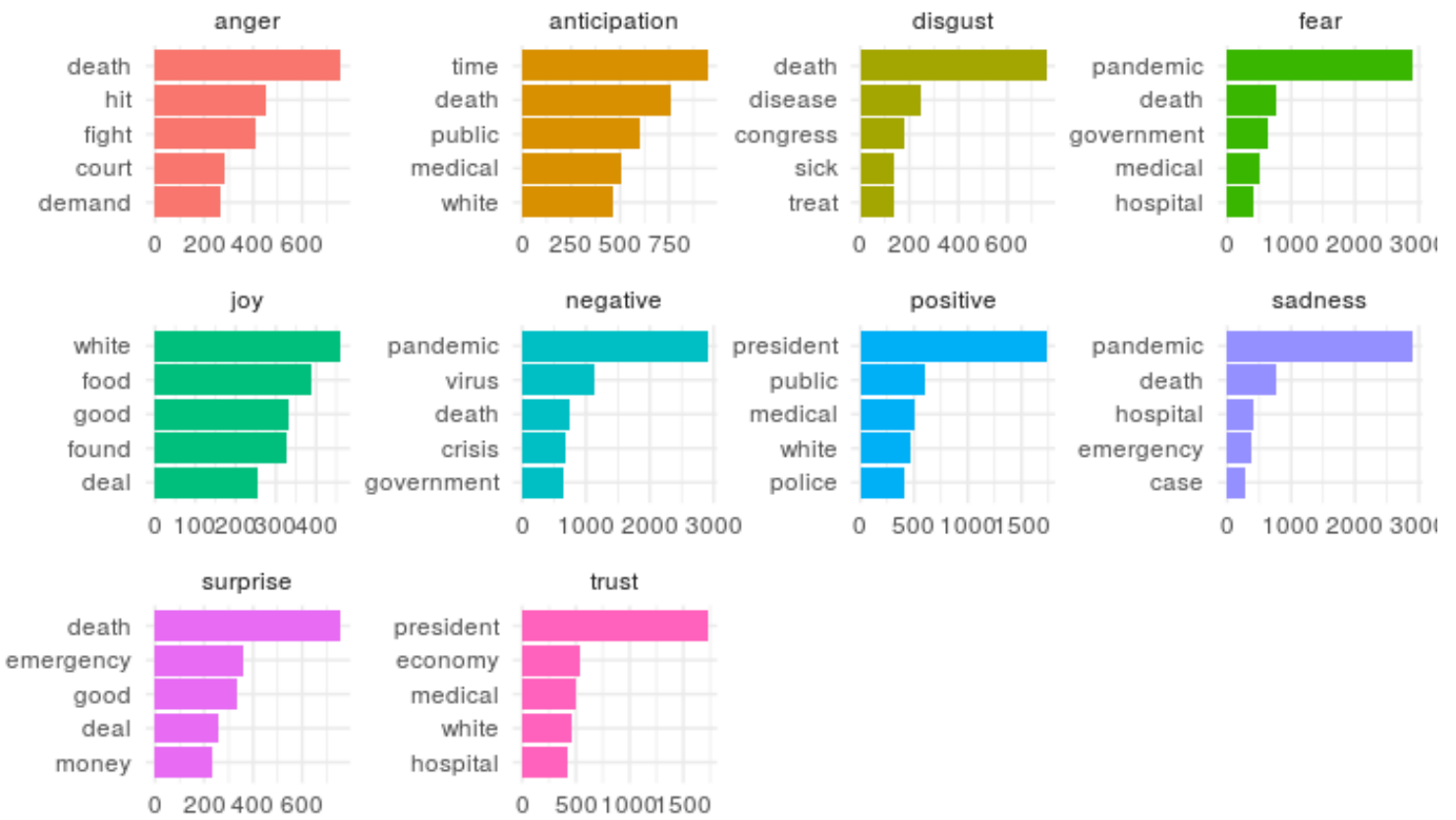
This is a visual representation of the words used and the sentiments they portray across all ten news stations. The most angry and also the most trusting news station is Fox. “Most Watched, Most Trusted”

CNN beats out USA by a close margin for the least angry.

Hide

```
news_words %>%
inner_join(nrc_words) %>%
  filter(!word == "trump") %>%
  count(word, sentiment, sort = TRUE) %>%
  group_by(sentiment) %>%
  top_n(5) %>%
  ungroup() %>%
  mutate(word = reorder(word, n)) %>%
  ggplot(aes(word, n, fill = sentiment)) +
  geom_col(show.legend = FALSE) +
  facet_wrap(vars(sentiment), scales = "free") +
  labs(y = "NRC Sentiment Contribution by Word (Major News Network Tweets)",
       x = NULL) +
  coord_flip() +
  theme_minimal()
```

Joining, by = "word"
Selecting by n



NRC Sentiment Contribution by Word (Major News Network Tweets)

This is a more specific representation of what words are being used across news media.