Parlor trick or worthwhile?

SENTIMENT ANALYSIS IN R

Ted Kwartler
Data Dude

Interesting visuals

Good Visuals

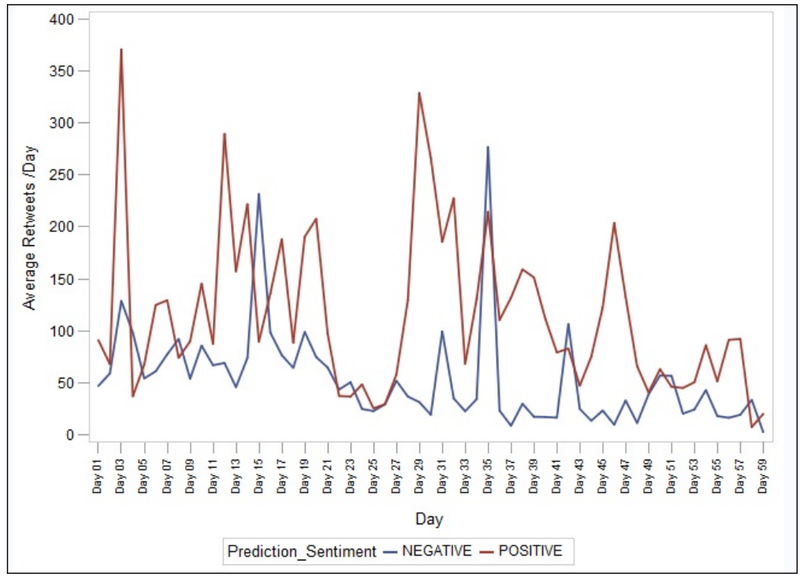
- **Simple** to interpret
- **Confirm** or **elucidate** data aspects
- **Context** for the audience
- **Appropriate** type e.g. line charts for time, bars for amounts

Bonus:

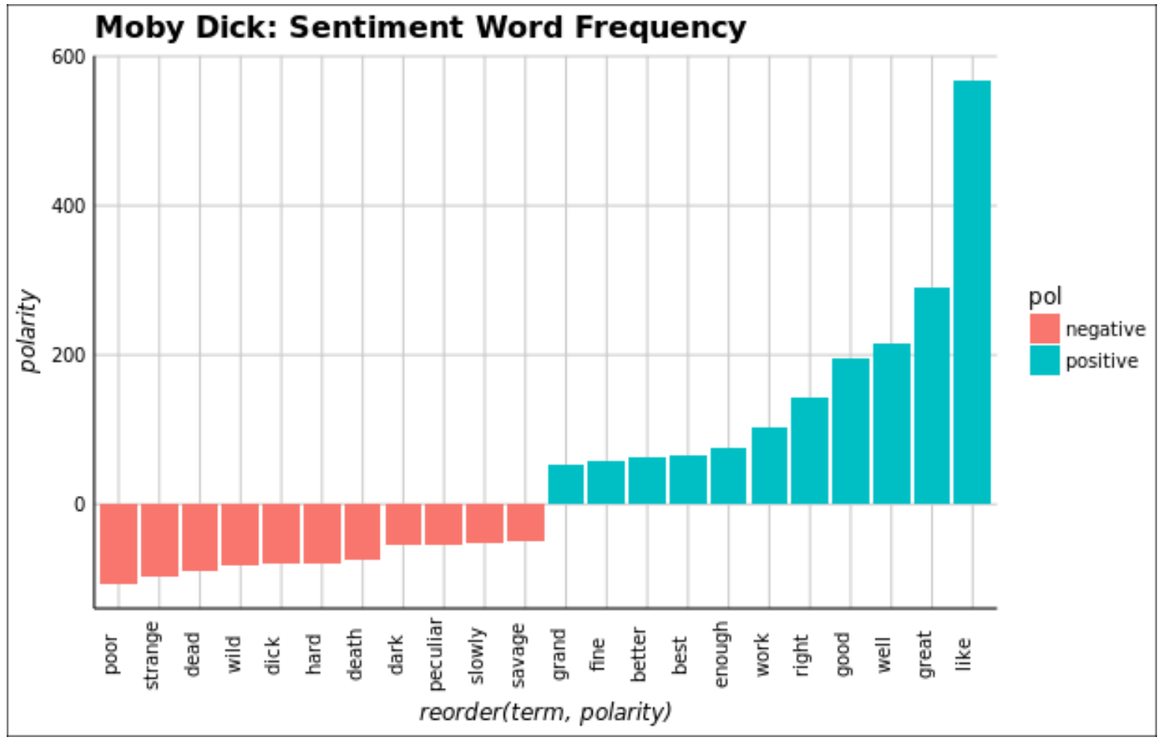
- Avoid word clouds

Tracking sentiment over time

Sentiment timeline - a way of displaying sentiment values in chronological order. It is typically a graphic design showing time periods, such as months, as the X axis and the sentiment values as Y axis values either as a line or series of bars.



Simple frequency analysis



`ggplot2` is a popular library based on the "grammar of graphics" for constructing visuals in R.

Let's practice!

SENTIMENT ANALYSIS IN R

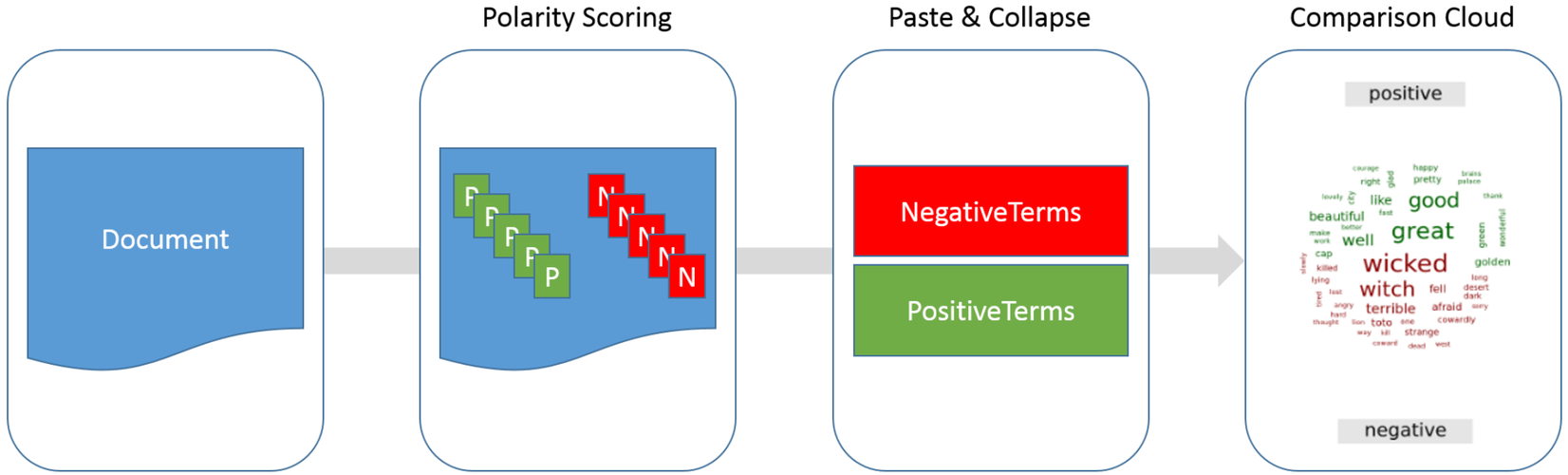
Introspection using sentiment analysis

SENTIMENT ANALYSIS IN R

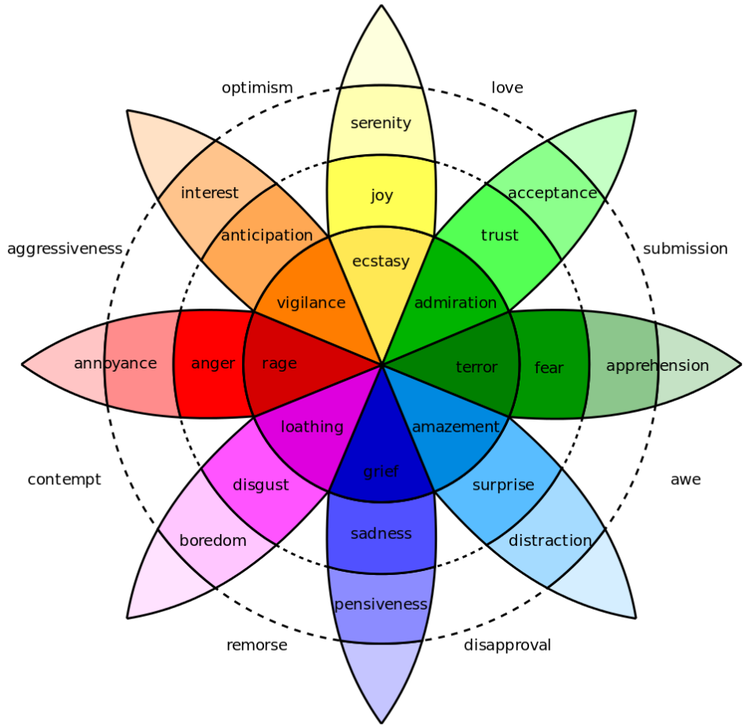
Ted Kwartler
Data Dude

qdap's polarity for subsetting corpora

```
library(qdap)
polarity(text.var, grouping.var = NULL)
```



Comparing frequent words in Plutchik's Framework



Where's Waldo? Where isn't Waldo?

```
x <- c("Nicole", "Nick", "Waldo")
grep("Waldo", x)
```

```
[1] 3
```

```
grepl("Waldo", x)
```

```
[1] FALSE FALSE  TRUE
```

```
!grepl("Waldo", x)
```

```
[1] TRUE  TRUE FALSE
```

Adding an "or" operator

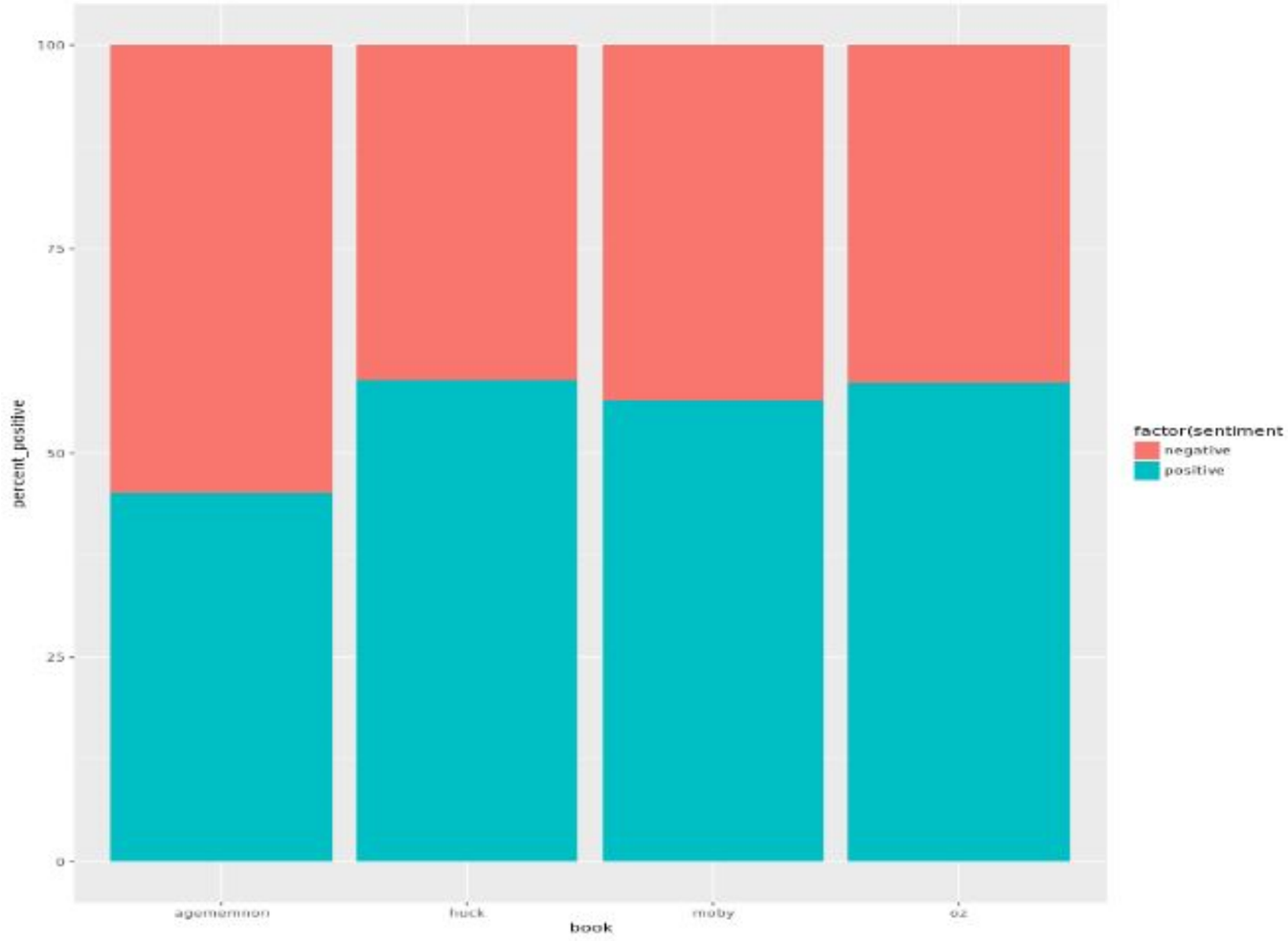
```
x <- c("Nicole", "Nick", "Waldo")
grepl("Waldo|Nicole", x)
```

```
[1] TRUE FALSE  TRUE
```

```
!grepl("Waldo|Nicole", x)
```

```
[1] FALSE  TRUE FALSE
```

Stacked comparisons for polarity mixture



Let's practice!

SENTIMENT ANALYSIS IN R

Interpreting a kernel density, box plots & radar charts

SENTIMENT ANALYSIS IN R

Ted Kwartler
Data Dude

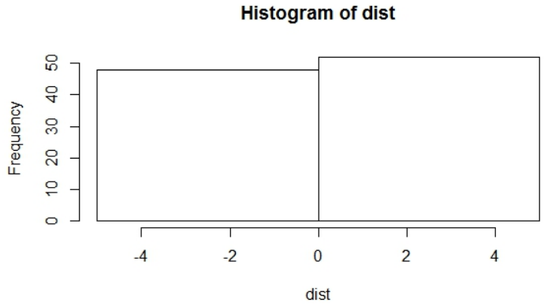
More visualizations

- Kernel density plot
- Box plot
- Radar chart
- Treemap

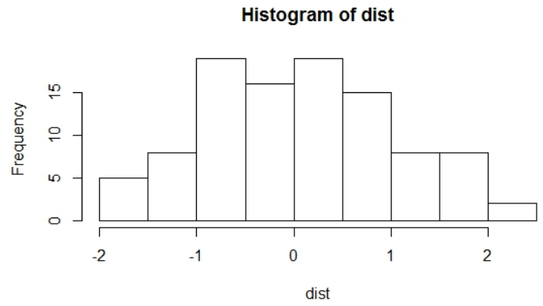
Kernel density plots vs histogram

```
dist <- rnorm(100,
              mean = 0,
              sd = 1)
```

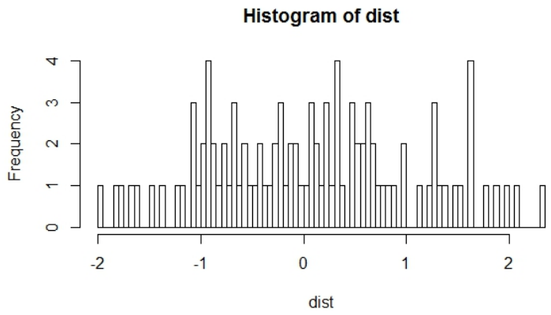
```
hist(dist, breaks = 1)
```




```
hist(dist, breaks = 10)
```

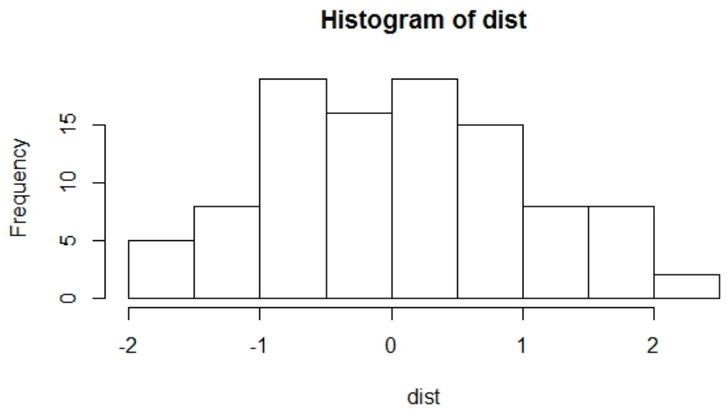



```
hist(dist, breaks = 100)
```

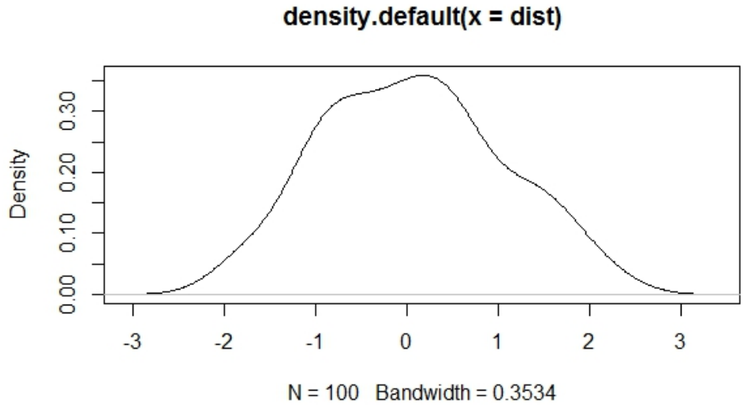


Kernel density plots vs histogram

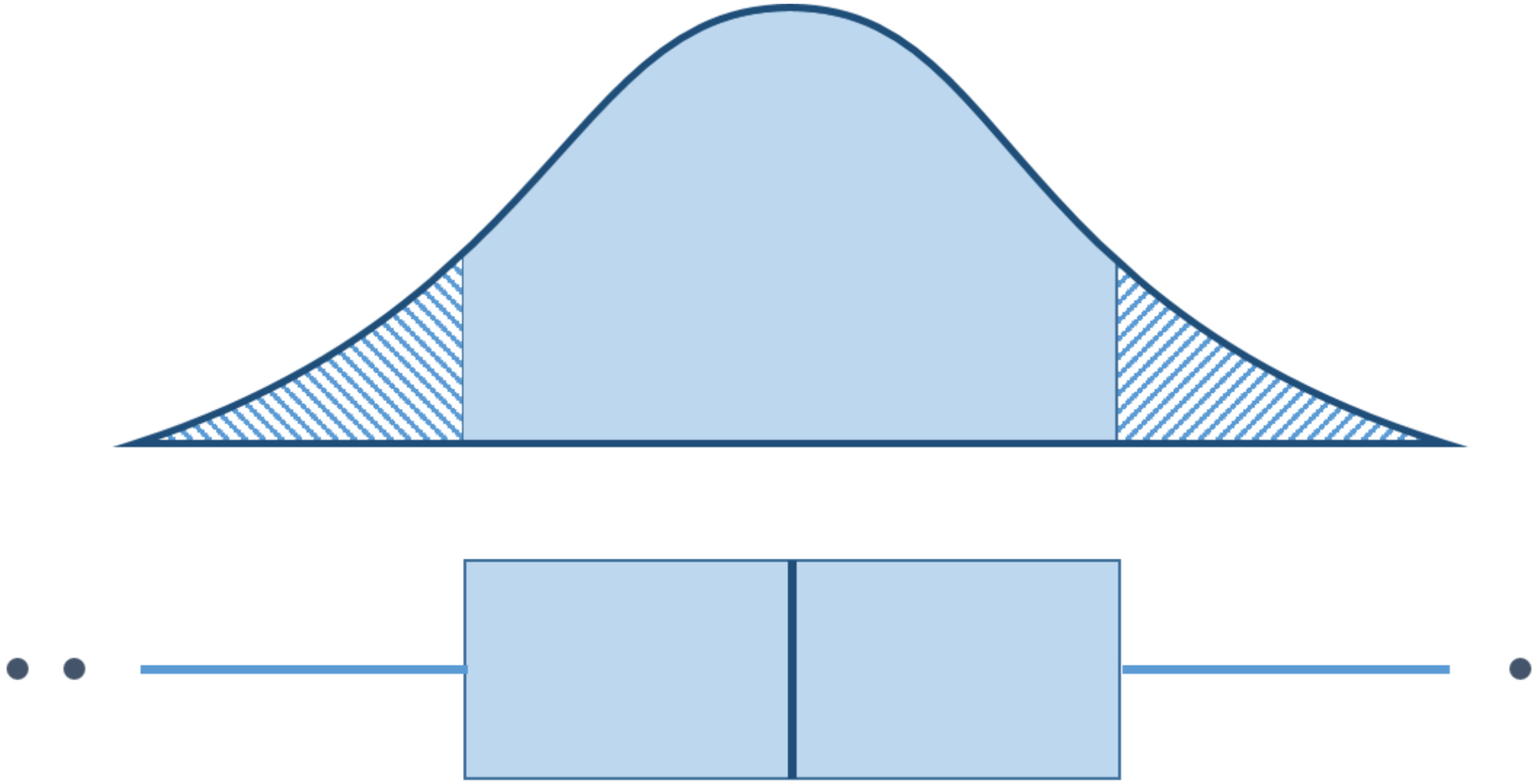
```
dist <- rnorm(100,
              mean = 0,
              sd = 1)
hist(dist, breaks = 10)
```



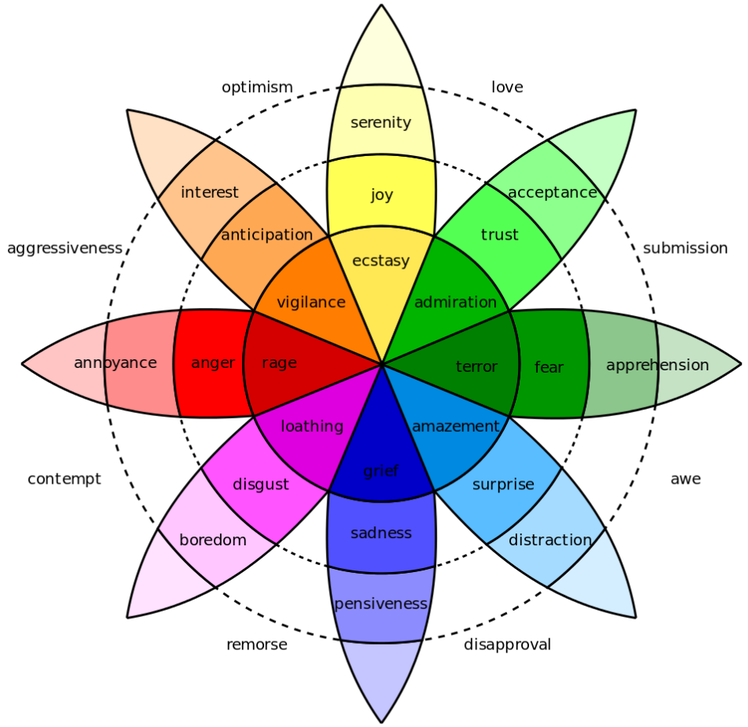

```
d_curve <- density(dist)
plot(d_curve)
```

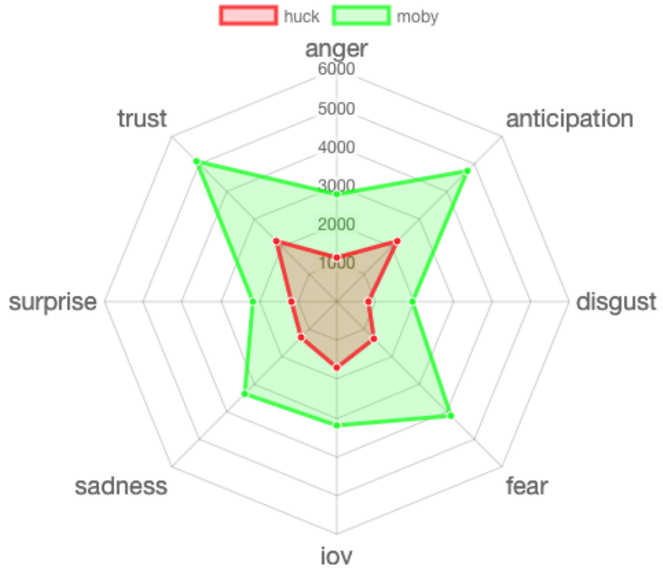


Box plot

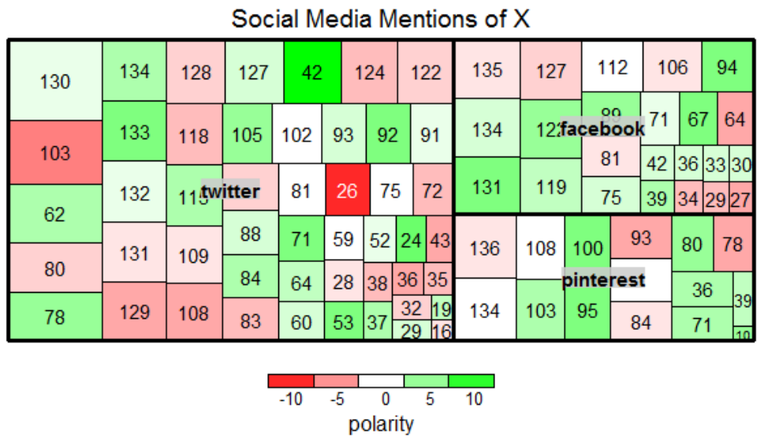


Radar Wheel of Emotion





Treemaps



- Each block represents a data point like a row
- Each block's size is dictated by another data dimension
- Each block is colored according to another data dimension
- Blocks are arranged into like groups using another data dimension

Let's practice!

SENTIMENT ANALYSIS IN R