



Sentiment Analysis  
DH in Practice Winter 2020

## Computational Sentiment Analysis Process (condensed)

- Material is broken into units for analysis (i.e., tokens)
  - (words, lines, sentences, paragraphs, tweets...)
- Material is compared to a sentiment lexicon, as unigrams (1 word at a time), and classified or scored
- The scores for each unigram are totaled for each token to produce the results.

# Interpreting Tweets Like a Computer

- AFINN: rates terms on a scale of -5 to +5 (created 2009-2011)
- NRC: classifies terms according to 8 human emotions & 2 sentiments (updated 2011)
  - anger, fear, anticipation, trust, surprise, sadness, joy, and disgust
  - positive/negative
- Bing: classifies terms as positive or negative (created in 2004)

## Lexicon links

AFINN: [http://corpustext.com/reference/sentiment\\_afinn.html](http://corpustext.com/reference/sentiment_afinn.html)

NRC:

<https://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm>

Bing: <https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html>

# Lexicon links (to full text)

[AFINN](#)

[NRC](#)

[Bing](#)