Eric Randall
4/20/2018
MSDS 455 - SEC 55
Individual Assignment 2: Text Analysis

A record company is looking for feedback from three potential reissues by The Beatles. Using a dataset scraped from RateYourMusic.com, analysis is done to determine which of the records are viewed most positively.
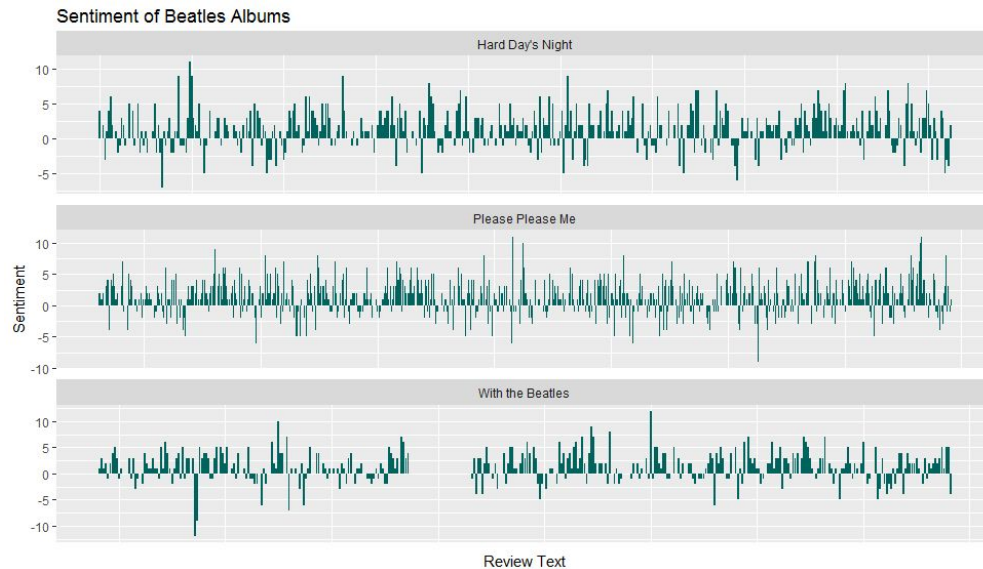
- Individual reviews have been combined into a separate file for each album and then processed.
- On the left, a traditional word cloud shows some of the most talked about attributes of the albums.
- We can see clustering occasional bad reviews, though overall reviews are positive. Please Please Me has the highest sentiment.

- In a comparison of relative frequency of terms between albums, we see that soundtrack, chord, accompanying, iconic and ballads are all distinct words to describe Hard Day's Night. Likewise, Covers and Debut are terms commonly and distinctly used to describe Please Please Me.
- Both albums could be marketed with a focus on the love songs as shown to be mentioned quite frequently in both albums.

Relative Frequency of Keywords in Beatles Reviews