

Digital Text Analysis Project

Group 4

RESEARCH QUESTION

Lin-Manuel Miranda's hit Broadway show, *Hamilton*, is a biographical musical that aims to tell the life story of Alexander Hamilton. Using text analysis software, we are looking to reveal the main themes and topics presented in the opening show tune 'Alexander Hamilton,' and how these ideas contribute to the show's view of Hamilton and to our overall understanding of Hamilton's legacy. Based on the known historical context and biographical nature of the musical, we predict that the text analysis will reveal themes of ambition, legacy, and conflict that took place during the American Revolution.

RESULTS

Text Analysis Using Voyant

The themes of ambition and legacy are certainly reflected in the prominence of words like "man," "Alexander," and "Hamilton," which suggest a focus on the individual and the historical figure himself. The term "New York" relates to the setting of much of Hamilton's work and the political landscape during the American Revolution, while words like "father," "struggled," and "dead" might indicate conflict, which aligns with the theme of conflict that was predicted. From this analysis, the song places a strong emphasis on Alexander Hamilton's personal origin story. We can see that in the latter half of the song, the word "new" peaks in frequencies (higher than "York") which could indicate some sort of shift or change in Hamilton's life.

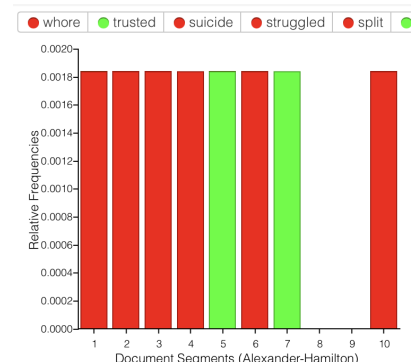
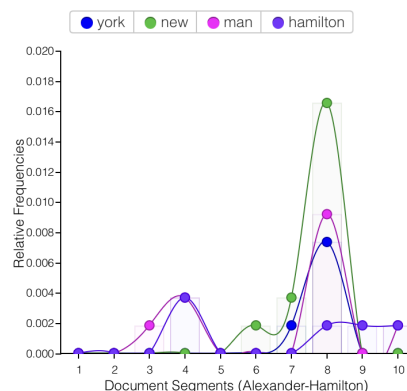


Figure 1: Word Cloud From Voyant Figure 2: Word Frequencies From Voyant

Figure 3: Frequency of possessive and negative terms

Text Analysis Using R

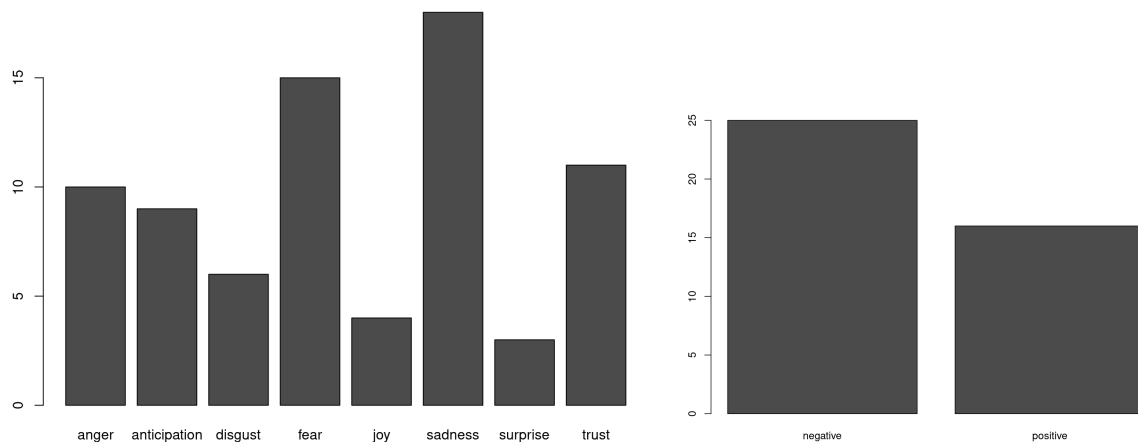
In R, we are also able to create word clouds; however, we are also able to use a g-score to determine the emotion expressed in different parts of the song. Because of this, we are able to gain a better general understanding of the specific emotions presented in the text (anger, fear, joy, etc.).



g_speech_sen	sentiment_vector
How does a bastard, orphan, son of a whore And a Sc...	-1.25
The ten-dollar founding father without a father Got a ...	-5.70
Alexander Hamilton My name is Alexander Hamilton ...	-0.85
Will they know you rewrote your game?	0.00
The world will never be the same, oh The ship is in th...	-2.90
Alexander Hamilton	0.00

Figure 5. Sentiment scoring of each sentence, where lower values indicate negative sentiment and higher values indicate positive sentiment

Figure 4: Word Cloud From R



g_sentiment	1 obs. of 8 variables
\$ anger	: num 10
\$ anticipation	: num 9
\$ disgust	: num 6
\$ fear	: num 15
\$ joy	: num 4
\$ sadness	: num 18
\$ surprise	: num 3
\$ trust	: num 11

Figure 6. Sentiment analysis shows positive and negative emotions, as well as specific emotions

DISCUSSION

Comparison of Text Analysis

Both Voyant and R were able to provide text based on word frequency and some degree of sentiment analysis. Both were able to create a word cloud based on word frequency, with the option to exclude certain words. The output results of each word cloud were similar, as they both highlighted frequent words such as "man," "Alexander," and "Hamilton." As for the sentiment analysis, R was able to use a `g_score` to assign specific sentiments to each line of the text, whereas Voyant was only able to distinguish between positive and negative sentiments. Despite R providing a more detailed sentiment analysis, both softwares produced similar outcomes of the song being composed of primarily negative emotions. The results of the word cloud are consistent with our hypothesis that the song would be centered around Hamilton, however the sentiment analysis results do not highlight the themes of ambition and legacy, as we initially predicted. More detailed text analysis and a close reading of the song as a whole is necessary to reach a conclusion about the true depth and understanding of the song.

Comparison of Software

Comparing Voyant and R for text analysis reveals distinct advantages and challenges. **Voyant** is user-friendly, with an intuitive interface that lowers the barrier to entry for those new to text analysis, making it accessible without extensive programming knowledge. It offers a range of predefined tools that may be difficult to code in R, such as making word trees, identifying collocates, and mapping locations of specific places mentioned within the text. However, it might limit customization and depth of analysis due to its predefined tools. **R**, on the other hand, offers comprehensive control over the analysis process, allowing for customized and complex analyses. Additionally, since the user is the one controlling the methods in which the text is analyzed, the confidence of the output results is clearer. This flexibility comes with a steeper learning curve, requiring proficiency in R programming; although pre-written R code is publicly available, tailoring it to a specific need is more tedious and requires more extensive knowledge about coding and R. Ultimately, the choice depends on the user's technical skills and specific analysis needs.

Voyant is designed for accessibility, providing a straightforward, web-based interface for quick analyses. R, being a programming language with various packages for text analysis, offers more customization and is aimed at users with programming skills seeking detailed, complex analyses.

For this project, we preferred R because we were able to clearly see the sentiment of each line and how it was assigned. High level coding experience was not necessary for the goals of this project so we preferred the analysis software where we could be more confident about our results and directly see how they were obtained.