Visualizing Text – Harding vs. Coolidge

Crystal Mosley

Northwestern University

Visualizing Text – Harding vs. Coolidge

Summary

President Harding's address's in 1921 and 1922 consisted of quite of a few of the same topics, such as: "American", "will", "public", "world", and "government". His 1921 address seems much lighter in subject, and less firm, in comparison of his 1922 address. In his 1922 address, he spoke more about "railway", "responsibility", "transportation", and "war"—all signs of necessary change.

President Coolidge's address's in 1924 and 1925 were only the same when it came to the topic of national debt. His address in 1924 was much larger than 1925. "Country", "government", "congress", "economic", and "policy" were the center of his agenda and topics in his 1924 address.

Overall, both Presidents had a lot of the same topics in speech, but both were focused on different agenda's.

Research design, measurement, statistical methods, and predictive models employed—

The address's of Harding and Coolidge were picked due to the large size of their speeches, as well as the United States turn of events in the early 1920's. Their speeches posed different topics in the center of their agenda, which can be seen via the *wordcloud*.

The *wordcloud* was chosen because it's easy to visually see key words from the data being used. The different sizing of the words shown in the *wordcloud*, as well as the different font edits that can be utilized for the text, help the viewers get an overall idea of the full text being displayed. It's a quick technique, and it's visually interesting and engaging, all at the same time. The viewers do not have to be technical do get a quick understanding of the data representation being shown.

Overview of programming work—

The programming work that was utilized came from the "Jump Start Code" for the *wordcloud* that was provided via *455_wordcloud_jump_start_v001.zip* file.

The Operating System used was Windows. The package libraries utilized were "wordcloud" to bring in the *wordcloud* framework, as well as "tm" to bring in the text mining capability.

To get the code to work, two addresses per each President chosen were put into separate work directories, which is where each separate session was pointed to during execution. The data text files were read into the RStudio, and then the *wordcloud* code was built. The minimum frequency required for words being pulled into the *wordcloud* was set to 5, and the maximum words to be displayed was set to 150 so that the *wordcloud* doesn't look too convoluted. Harding's *wordcloud's* were set to "black" so that all text shown was black; consequently, it didn't look as appealing as it could have been; therefore, Coolidge's *wordcloud's* were set with the color code *colors=brewer.pal(6, "Dark2")* so that the viewers could see a difference in the aesthetics of each cloud.