

PRESIDENTIAL INAUGURAL ADDRESS ANALYSIS

INFO 633 – Project C

Caleb Miller, Luke Chesley, and Harrison Pauwels

PURPOSE

- Explore and visualize results of analysis of presidential inaugural speeches
- These speeches are the subject of analysis because:
 - Speaking on the state of the nation
 - They even go as far to speak on the state of the world
 - Focused on goals for upcoming term
- These speeches reflect America and the American people

DATASET

- Dataset gathered from The American Presidency Project
 - Created and maintained by UC Santa Barbara
 - They house many presidential documents, such as:
 - Eulogies
 - Executive Orders
 - State of the Union Addresses
 - Etc.
- A web scraping program was used to obtain the inaugural addresses.

ACQUISITION AND PREPROCESSING

- This stage of the project was completed in Python
 - Requests and BeautifulSoup modules
- Looped between individual URLs to pull all presidential inaugural addresses
- Some non-text characters (such as new line and tab characters) had to be removed from data
- Final dataset is raw text

ANALYSIS

- Python string manipulation functions assisted with simple analysis of data, producing:
 - Number of words in each speech, characters per word, number of unique words, etc.
- NLTK helped with more complex analysis such as:
 - Average number of words in each sentence, what part of speech each word is, etc.
 - Also used to find and remove stop-words
- Flesch-Kincaid score used throughout research
 - Measures readability of English language passages
 - Allows us to give a score to compare addresses

DASHBOARD CREATION

- We wanted coherent graphs that are:
 - Easily viewable
 - Interactive
- Graphs created with the Python module Plotly
- Dashboard created with the Python module Dash
- Combining the functions of these libraries allows the viewer to filter the graphs by presidents and years

DATA EXPLORATION

- Dashboard: <https://final-633-b74277a80624.herokuapp.com/>

LIMITATIONS

- The dataset is focused on only presidents which may represent trends or relationships that are not representative of the entire historical spectrum
- The small sample size of these speeches may skew the interpretation of the trends that were found
- Inaugural addresses occur every 4 years, so there is a limitation in drawing comprehensive conclusions about long term trends