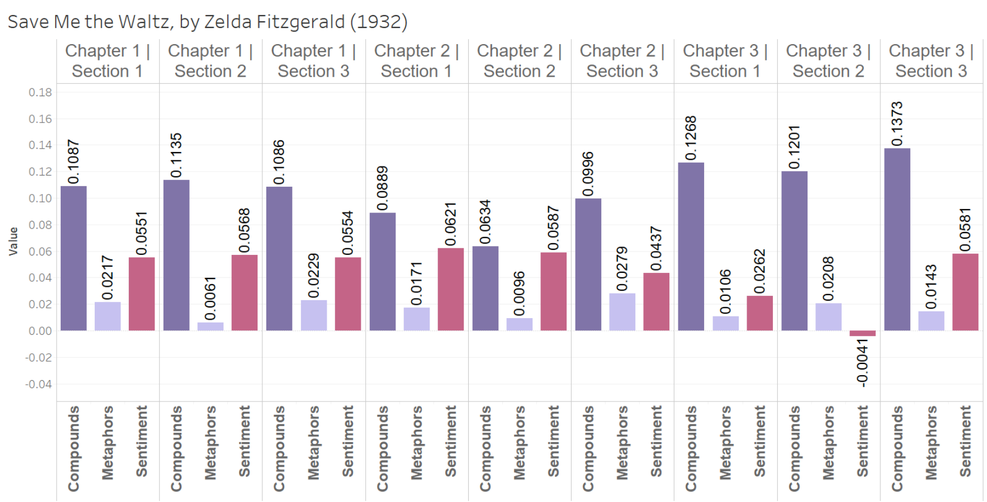
Another Hammer in the Humanities Librarians' Toolkit: **Text Analysis/Mining**

Working with faculty members on text analysis/mining projects has **enhanced** the status of humanities librarians by giving them **entrée** to more collaborative projects

Josh Been – Eileen Bentsen – Bill Hair

BAYLOR University Libraries

Use of Language and Sentiment in *Save Me the Waltz*



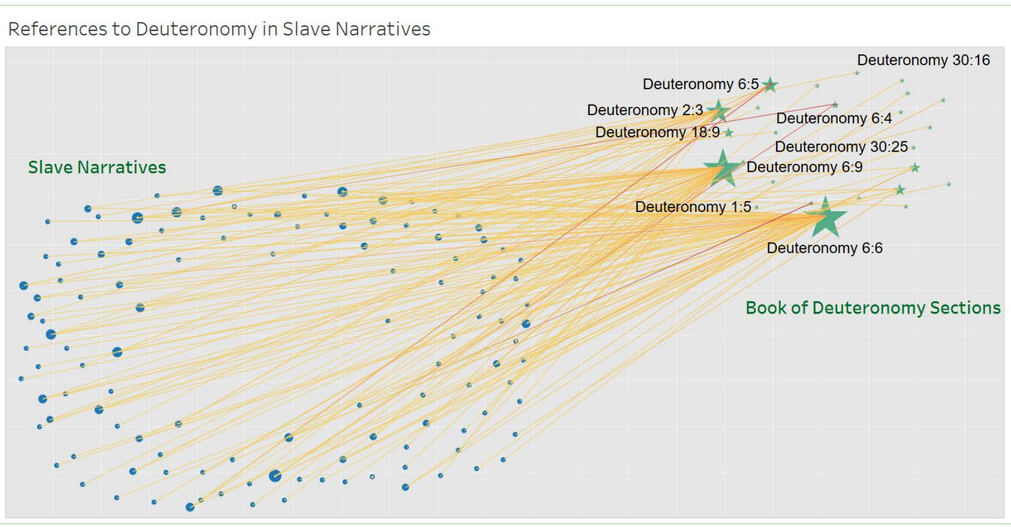
Identifying compound sentences, metaphor, and sentiment for each sentence

First text analysis project we undertook! - Introduced English faculty to the library's digital humanities services and gave humanities librarian the confidence to promote these services.

Methods:
(1) Compound: FANBOYS
(2) Metaphor: Two noun definition similarity
(3) Sentiment: VADERSentiment

Source: HTML Online Website

Fugitive and Former Slave Narrative References to Deuteronomy



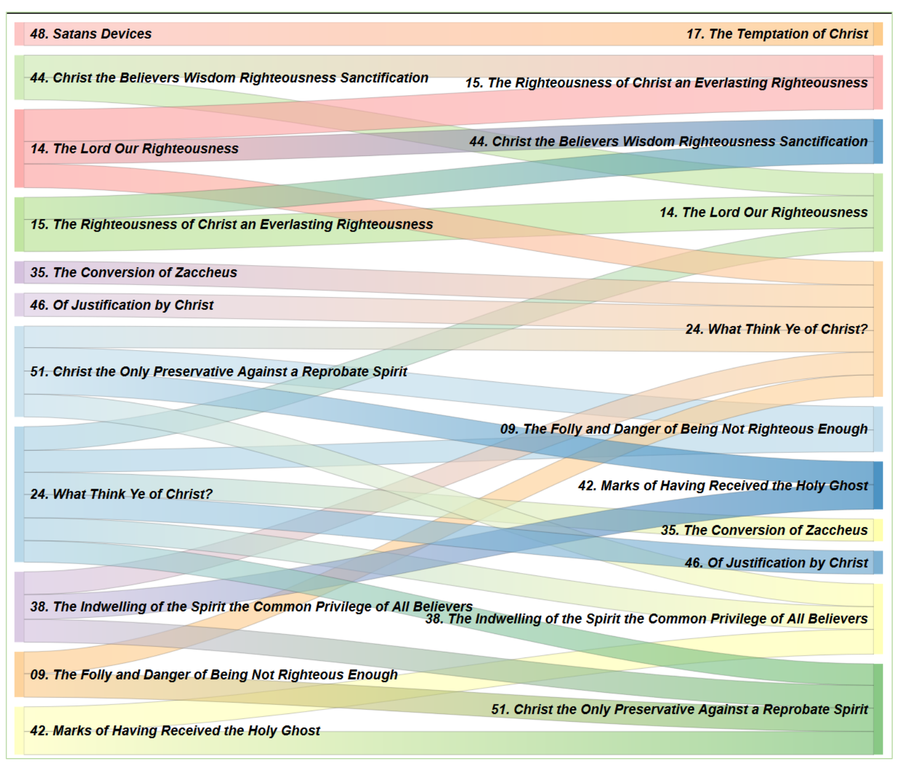
Identifying biblical references from transcribed fugitive and former slave narratives

Opened doors to more collaborative projects with Seminary faculty, including a department-wide digital humanities demonstration and additional faculty research partnerships around text analysis.

Methods: Circular network created using NodeXL and visualized using Tableau. Calculated using fuzzy n-gram similarities.

Source: UNC: Documenting the American South – North American Slave Narratives

Similarities Between George Whitfield's Sermons



Measuring the similarities between 57 of George Whitfield's sermons with each other

This project was done in partnership with a music graduate student, and this is opening doors to topic modeling, similarity analysis, and TF-IDF keyword generation across music lyrics.

Methods: Sankey Chart using Google Charts JavaScript Classes. All similarity scores .35 or higher (35% similar) displayed. Term frequency-inverse document frequency (TF-IDF) cosine difference used to calculate similarity scores.

Source: Center for Reformed Theology and Apologetics