

Alice Bogdan (ucn3qn@virginia.edu)
DS 5001
03 May 2022
Manifest: Gilmore Girls

Provenance: This project focused on the analysis of tv scripts for the show *Gilmore Girls*. We acquired the TV scripts from the website “Forever Dreaming Transcripts” through the following link: [1]

<https://transcripts.foreverdreaming.org/viewforum.php?f=22>

It included scripts for all seven seasons of *Gilmore Girls* from 2000 to 2007 and the four-part mini-series *Gilmore Girls: A Year in the Life* that aired in 2016. For the purposes of this analysis, we limited our analysis to the original tv show and excluded the mini-series for various reasons including a break in timeline and different organizational structuring. We utilized web scraping to acquire the tv scripts for each episode.

Location: The source files for this project be found in the following location:

<https://virginia.box.com/s/chr80kl77gdv9s9556qhaf34yt0i33q9>

Source files included, but are not limited to, episode scripts (dump.zip), DATA, CORPUS, VOCAB, and BOW tables, Jupyter notebooks, etc.

Description: *Gilmore Girls* is an American comedy-drama television series that ran for seven seasons from 2000 to 2007. It follows the life of two main characters Lorelai and Rory Gilmore, (mom and daughter duo) starting from Rory’s sophomore year of high school and ending with her college graduation. The events revolve around their life living in a small town and their personal, educational, and professional ambitions and growth ranging from pursuits of attending prestigious schools and opening their own business to developments in familial, friendly, and romantic relationships.

Gilmore Girls is well-known for having longer scripts due to the fast dialogue of both Lorelai and Rory; sources state that the shows scripts averaged around 80 pages long while the average length of hour-long tv shows are 40-50 pages. [2] In addition to being one of my favorite tv shows, the opportunity to analyze a tv show with a potentially larger corpus was a driving factor for my choice in text.

Format: We utilized web scraping to acquire the tv scripts for each episode and stored them as HTML. These files can be accessed via the “dump.zip” file located in the Virginia Box link provided above in the **Location** section. The HTML files were merged into one raw data file and stored as csv file (file name: gilmore-girls.csv) with an internal structure of ['title_raw', 'time_raw', 'line_num', 'doc_raw'].

References

[1] *Gilmore girls transcripts - forever dreaming*. Index - Forever Dreaming. (n.d.). Retrieved March 22, 2022, from <https://transcripts.foreverdreaming.org/viewforum.php?f=22>

[2] Grady, C. (2016, November 22). *Why Everyone on Gilmore Girls Talks a Mile a Minute*. Vox. Retrieved March 16, 2022, from <https://www.vox.com/culture/2016/11/22/13554566/gilmore-girls-fast-talking-explained>