Sherlock Holmes Over Time

Problem Description

Since his introduction to the public over 125 years ago, Sherlock Holmes has grown to become one of the most popular and enduring characters in literature. With over 250 film and TV portrayals and an untold number of literary adaptations, there is a vast timeline of Sherlock stories available for analysis. Both fans and scholars would be interested in exploring trends over time from the original stories to modern adaptations, such as in word usage, character co-occurrence, and the overall sentiment expressed. These trends may offer insights into changes in literature, film, and television as a whole.

Target Audience

Although the target audience could perhaps be considered to encompass only die-hard Sherlockians, there may be more generally meaningful and interesting trends regarding the changes through time from the original stories to modern adaptations, such as in the overall tone of the stories. These trends may be of interest to literature researchers and scholars. In addition, the general public may find the trends to be of interest.

Description of the Data

Version/Book (Nominal) — Name (e.g. Sherlock, Elementary)
Year Written (Quantitative) — Year (e.g. 1887, 2009)
Total Word Count (Quantitative) — Number of words in the story
Word/Phrase (Nominal) — User specified word or phrase
Character (Nominal) — Holmes, Watson, Hudson, Lestrade, Mycroft, Adler, Moriarty
Sentiment (Nominal) — Positive, Neutral, Negative
Page Number (Quantitative) — Number
Line Number (Quantitative) — Number

Encodings

Both Version/Book and Year Written are simply encoded as text. In both examples, the total word count is encoded as length, one of the most accurate encodings for determining relative size. The Word/Phrase is displayed as text, and is also encoded using color, making use of preattentive processing for quick recognition. The same is true for the Character. Each possible value for Sentiment is encoded as a color (Red, Yellow, and Green). The page number and line number (used to display the location of Words/Phrases, Characters, and Sentiments) are encoded using position for accuracy.

Interactions

Books/versions can be sorted either by the year they were created in or by their length. Words and phrases can be entered in order to see their locations within the texts. When hovering over the location of a word or phrase, the specific line surrounding it is displayed. When hovering over a character’s name, characters with which they have a co-occurrence will glow in the location where the co-occurrence takes place.

Aesthetics

The visualization displays a vast overview of multiple adaptations simultaneously; the colors are chosen sparingly, with the intent that they not clash with one another. Sufficient white space is used in order to separate each individual piece of the visualization.

Specific Questions

- How often does the word/phrase [insert word or phrase] appear in the original text? In newer adaptations?
- What are the differences in word appearance between versions?
- What is the distribution of a specific word/character over time or in the space of the book/adaptation?
- What are the relationships between the characters in the original text? In the new adaptations?
- How often do the characters appear (old vs. new)?
- What is the general sentiment of the original text/new versions?
- Are there any trends over time?

References

http://www.jstor.org/stable/1390665
http://neoformix.com/2013/NovelViews.html