Mariah Jacobs

11/28/19

SI 507 Final Project Proposal Revision

MOTIVATION: Readers are often interested in knowing when their favorite books are made into movies, and movie-watchers are sometimes surprised to find that a movie is an adaption of a previously published book. One of the issues both movies and books have is that they generally portray women less often, and in less meaningful roles, than men. The Bechdel test is a measure of representation of women in fiction/media, originating from a comic by Alison Bechdel in order to call attention to the lack of representation of prominent female roles in film. I plan to create an application that displays information about which movies from a Bechdel-test-evaluated data set have accompanying books, and whether the storyline of the media passes or fails the criteria of the Bechdel test.

PROJECT DESCRIPTION: I will use the Google Books API, the Open Movie Database, and a CSV file containing data about movies passing the Bechdel test in order to display information to the user about each title's book and movie, as well as the Bechdel test status.

Back-end implementation: I will use Python to collect and cache the data I retrieve through the Google Books and Open Movie Database APIs. I will also read data from the CSV file in Python. After collecting all of this data, I will create a SQL database to store the Books, Movies, and Bechdel test statistics in their respective tables. I will query the database to connect the different information for each title, and store my analysis of book-plus-movie titles in Python data structures.

Front-end implementation: I will present this analysis by creating a Flask app. I plan to chart the book, movie, and Bechdel test information using HTML/CSS. Users will be able to prioritize whether they want to see books/movies that pass the Bechdel test or fail it, and what year(s) within a specific period they want results for. As a stretch goal, I would like to present the user with a data visualization of the representation of women in media over time by plotting the titles by year that do pass the Bechdel test using Plotly.

Audience: The intended audience of this application is anyone who likes books and/or movies, especially books that have been made into movies and story/plotlines that have increased representation of women.

Data Sources and Challenge Score: Google Books is a Web API I haven't used before (challenge score: 3), the Open Movie Database is a Web API I haven't used before that requires an API key (challenge score: 4), and the CSV file of Bechdel test movie data is greater than 1000 lines (challenge score: 2). All together, I expect my challenge score to total 9.