Web App Project Proposal

Group: Tore Banta and Charlie Sarano

Working Title: Rap-up

https://bitbucket.org/bantat/cs257_webapp

User Population: This web app is directed toward an audience of users who have a wide range of experience with software and computers in general. A potential user could have little to no experience using a graphical interface based utility, or they could be very comfortable with using software and user interfaces. We are targeting a subset of users who are interested in, or enjoy hip-hop music and its related sub-genres. Since musical preference isn't usually related to tech literacy, we aren't making any assumptions on the ability of our average user to use software.

Abstract: This web app will aim to build a useful resource that will serve as an informational tool for the history of the rap genre. Features will include a graphical timeline, with different points in time that include the release of major and influential albums, as well as additional information on the artists that release those albums. This will serve to help rap aficionados browse the history of rap, as well as see use additional features to examine what common phrases or words were used by that artist, as well as comparing relative vocabulary complexity. By building this tool, fans of the genre will be able to move across the time spectrum, and identify the highs and lows of the genre based on a number of different variables, such as major releases or artist vocabulary.

Data: We are planning to primarily use the Million Song Dataset (http://labrosa.ee.columbia.edu/millionsong/) and the MusixMatch database (https://developer.musixmatch.com/) to power our web app. Additionally, we tentatively plan to use content from Wikipedia for additional description/information, although we aren't planning to utilize any major subset of Wikipedia or use the site's API to develop our dataset. Instead we will develop a smaller database of info we need for our web application. Wikipedia is under the Creative Commons License, so as students using the content for academic purposes, we fall under fair use. The Million Song Dataset and MusixMatch are both designed for public use, with some citation instructions for proper acknowledgement when we use the resources. The Million Song Dataset and MusixMatch both have thorough documentation, with data downloads, APIs,

and Python wrappers for using the datasets, although we expect a relatively steep learning curve in using the datasets.

Features:

- Graphical timeline of significant releases, events in rap history
- Database of artist information, including…
- Most frequently used words
- Artist vocabulary
- Search tool for looking up artists, albums, or dates
- Basic biographical information
- Album release dates

We are very open to suggestions and/or recommendations as to how we might proceed with a project of this nature and scale, so please feel free to give us criticisms or feedback.