Harvard University
Data Science I – Fall 2017
Final Project – Scope of Work

## **Group Members:**

Paul Blankley Nathaniel Stein Ryan Janssen Andrew Lund

1. What is your group #?

4

2. Have you met/communicated with your fellow teammates?

Yes

3. Have you met/communicated with your assigned TF?

Yes, Nick Hoernle, IACS Lounge, 171745Z1117.

- 4. Has your team formulated a well-defined question to address in your project, based on the project description and references?
  - Can we predict the general success of a Spotify playlist using regression?
  - Can we classify a playlist into different popularity classes using the predictors we use for the regression model?
- 5. Briefly describe your team's plans for work to be completed by Nov 28 (milestone three). Please assign specific tasks to team members and deadlines for when these tasks are to be completed.
  - Mostly complete: Nate will put together spotipy functions into project-specific methods.
  - TBD: Paul will investigate using data sources other than the API (Wikipedia, external ranking sites, possibly Million Song Database) to enhance dataset.
  - 20 Nov: Group meeting to discuss finalized data collection and formalize EDA plan.
  - 22 Nov: All members will complete data collection.
  - 22 Nov: All members work to build out a significant and relatively clean dataset given data and metrics we can pull from the Spotify API.
  - 26 Nov: All members complete preliminary EDA.
  - 27 Nov: Group meeting to consolidate EDA, describe data, and summarize findings for milestone 3.
  - 28 Nov: Andrew will write final version of milestone 3 and submit on Canvas.