Dear Project Sponsors,

Hope you have had another great week, we have been having some pretty good time working on the Spotify's project and have made some good progress since the last report. To give you an idea on **our current status**:

1. We tested Doc2Vec, Word2Vec and LDA and TFIDF models on textual information from Genius, Youtube and Lyrics data. We have yet to identify the best, but we will do so soon once we have some baseline metrics.
2. Made a pilot pipeline for playlist generation using a similarity matching approach.
3. Implement some evaluation metrics as described in the recsys challenge 2018.
4. We have brainstormed about testing evaluation metrics along with user experience. We have yet to device the experimental hypothesis and variables, but have an initial mockup sketch for playlist comparison: http://spandan-project-manager.herokuapp.com/
5. Have further refined our data collection processes, currently we are working with 5k playlist..but will be able to scale up to the one million playlist dataset.

Our plan for the next weeks:

1. Further refine our generative models, we will explore using a hierarchical bayesian model and a deep generative model (RNN, GAN, VAE) for playlist generation and continuation.
2. Improve our evaluation metrics to cover more desired properties in playlists.
3. Further refine our idea for measuring user experience.

Please let us know if you have any suggestion or concerns regarding our project and we will look forward to keeping in contact with you along the term!

Sincerely,

Benjamin Sanchez-Lengeling, Timothy Lee,
Mehul Smriti Raje, Spandan Madan
Team Spotify for APCOMP 297r