**Capstone 1-on-1 Prep Worksheet**

Please send me an electronic version of this by resaving it with your name, e.g. 1on1\_prep\_JasonWiener.docx to this folder by 4pm Tuesday, Feb. 8.

*Who is the audience (in the rhetorical sense we’ve discussed in class) for your paper or project?*

A team of analysts at either a music streaming or entertainment management firm

*What is the research question your paper is attempting to answer? (Or if you are engineering something, what is the problem you are engineering a solution to?)*

Can a song’s popularity be predicted from its audio features (energy, acousticness, key, mode, bpm, etc.)?

*What hypotheses will you be testing with your analysis? (Or, if you are engineering something, please describe or sketch out the mechanics of what you are engineering?)*

My hypothesis is that there is a significant relationship between certain audio features and a song’s popularity.

The higher the energy of a song, the more likely it is to be deemed popular

In rap songs, there is a significant relationship between beats per minute and a song’s popularity

In the last five years, there has been a shift in significant features of a song’s popularity from lyrical content to beats per minute and energy of a song.

What specific data will you use to test each hypothesis and what do you think will be the appropriate statistical tests? (Or, if you are engineering something, what specific data will you be loading/transforming/combining/presenting?)

I will be using data from Spotify’s API pertaining to audio features of a song, as well as things like the song’s name, genre, year of release, and most importantly, popularity.

Quickly (don’t think too hard on it) write up an outline for your paper.

Introduction

* Talk about music analytics and its presence in the music industry
* Talk about the variability in the music industry, the potential profitability in using analytics to predict something like song popularity
* Introduce the research question and hypotheses I’ll be addressing

Body

* Talk about the methodology and the approach I took to address the problem.
* Descriptive statistics on the dataset, give an overview of things like range of values and what values mean in context
* Describe the model(s) I made, go through developmental process to give value to the decisions I made and why
* Show regression output, interpret the numbers and results
* Identify the key features to songs that were significant in predicting popularity
* Visualize results

Conclusion

* End with a brief overview of everything discussed
* Discuss the use cases for such a model and how it can impact the entertainment indistry

Quickly (again, don’t think too hard on it) describe a few of the visualizations you expect to build for your paper and/or poster.

I’ll aim to visualize trends among features and how they range across various genres, or time periods. For example, visualizing how lyrical uniqueness differs in rap songs from pop songs.

If there is the data for it, showing how a song’s popularity has changed over time (All I want for Christmas is you changing popularity over the seasons per year)

Maybe for keys of a song, since it’s a categorical variable, I can show how average popularity differs among categories

Maybe show how the popularity of a song changes as energy increases