Ideas:

* [song2vec](https://algoritmaonline.com/song2vec-music-recommender/)
  + Personal data
  + 4 Tasks Under Same Overarching Idea:
    - Spotify enhance (suggest new songs to add to existing playlist)
    - Spotify blend (music similarities based on multiple people’s music)
    - Create entirely new playlist from input attributes (this is Jess’ Music final project)
    - Spotify radio stream (Spotify DJ)
  + [Spotipy API](https://spotipy.readthedocs.io/en/2.22.1/)
  + Would need to request data soon… takes “up to 30 days” but took 3 for me
  + Raw audio files (how would we get this)
* [~~AskOski~~](https://askoski.berkeley.edu/)
  + ~~Recommend courses based on your major~~
* ChatGPT
  + Make data
  + Ask questions & analyze how often it’s answered correctly (ie. make a dataset)
* ~~Discord bot from personal data ???~~
  + ~~Personal data~~
* ~~Ed bot to answer posts like a GSI would~~
  + ~~Scrape data from Ed (can download CSV of staff metrics, JSON of threads)~~
    - ~~Consent of students?~~
  + ~~Ethics of pulling all the data ?~~

**Your group members' names followed by the number of group mates in parentheses**

Jessica Golden, Jonathan Ferrari, Ciara Acosta, Rahul Shah, Arman Kazmi, Abigail Yu, Sean Yang (7)

**A title that generally summates your project (need not be the final name)**

Personal Music Exploration

**A paragraph that explains your project, identifies key stakeholders, and specifies why ML is an ideal approach to addressing your project**

We are interested in providing personalized music recommendations given a user’s Spotify data. There are multiple routes that we are interested in pursuing and that would be of interest to the user. Some examples include recommending songs based on a playlist, combining multiple users’ song preferences into a shared playlist, creating playlists based on input sentiments, or generating playlists based on other user patterns (eg. seasonal patterns). Machine learning would be an ideal approach because it would extract complexity from the data that would not be intuitively derived. Machine learning is also applicable here since we are working with a large dataset and this amount of data is largely unstructured – the connections between songs that our model will learn are hard for a person to draw due to the vast genres and personalities a person could like, which can defer from one's own personal biases. The methods that we will likely utilize are clustering, word2vec, dimensionality reduction, ensemble learning, and RNNs. Some key stakeholders are music listeners as our model would improve user experience by allowing for additional exploration. Artists and other Spotify collaborators would also benefit, as their products would be introduced to a new audience that they may not have previously reached. Spotify, as a company, would also be a key stakeholder as the increase in user retention not only directly benefits Spotify explicitly with additional users with matches their company mission of bringing music to people but also implicity as *their* financial stakeholders will now also benefit from the respective expected increase in their stock price, which comes from more users.