Project Proposal:

Kevin Zhou kjzhou2, Beth Fineberg erf3, Cale Wolf caleqw2, Evan Matthews evanmm3

Our project will be dealing with artists on Spotify and their connections with each other. We will be using the Spotify API, and we will use HTTP GET requests to retrieve the data. We understand that we will basically be building our own database, different from the prepared datasets of OpenFlights or the SNAP libraries. In this project, we will be able to input an artist able to be found on spotify, and give information about the artists that they are connected to. In each album that a main artist produces, it also lists the co-artists. This is how we will be building our graph of artist connections. There will not be weights for any of the connections. To access the API, we will be using an HTTP Protocol using cURL. This may change depending on whether we are able to access the package but we would use either the cpr or curlpp packages to send HTTP requests.The data that Spotify will be returning to us will be a JSON object, so we will have to either build our own JSON parser or use nlohmann::json from github to parse the data. With the ability to use GET requests and parse JSON files, we would build a graph of Spotify artists linked to each other. With this data, we will be able to create a playlist that connects any two artists with songs that artists have collaborated on. This would be similar to the shortest path algorithm for the OpenFlights data. Limitations in this project is the speed at which we are able to access the API. With multiple requests, the Spotify API is rate-limiting, which means that we would have to retry the requests after some time if we send too many. This can really slow down our search, especially considering the number of artists that are on the platform. If this does happen, we may be able to remedy this by only selecting the first 20 artists that the main artist has collaborated with. With less connections, there are less calls to the API and could possibly speed up our program.