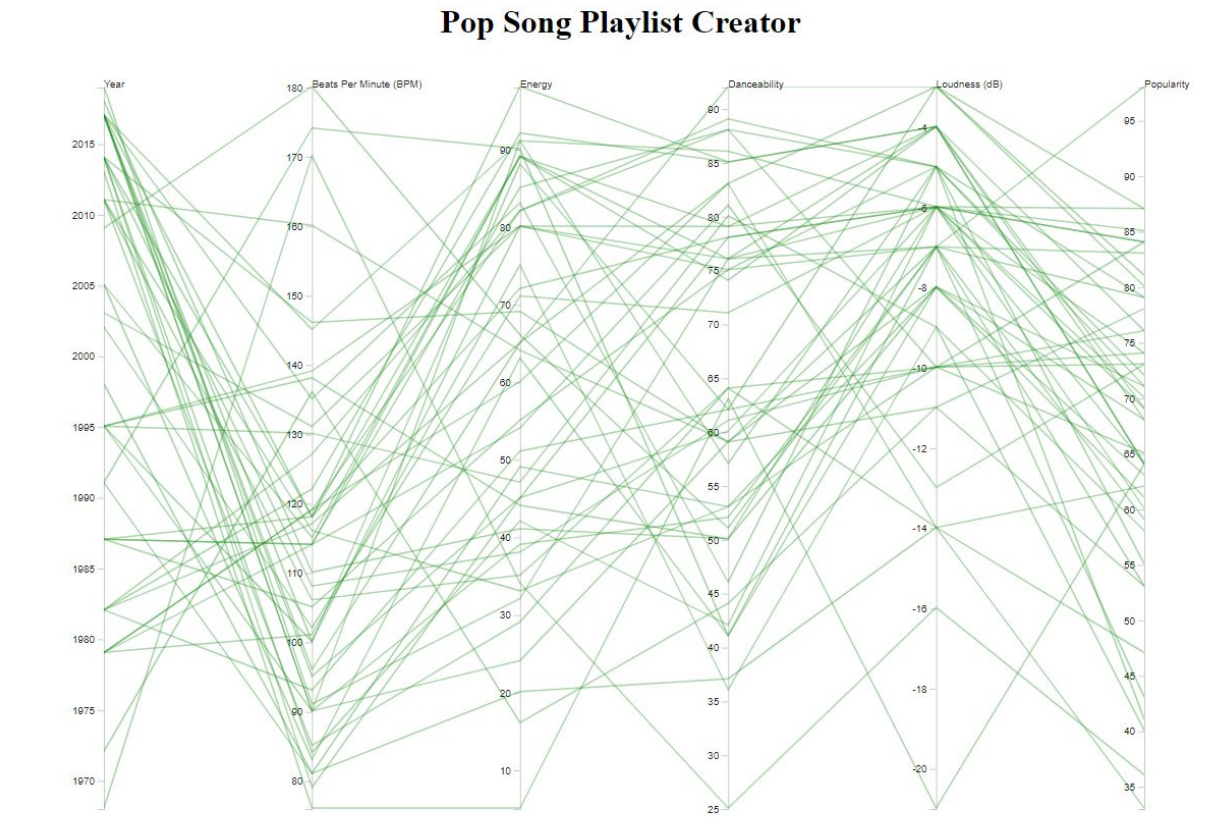
Pop Song Playlist Creator
Visualization Final Project

Jared Butler
Lars Christensen
Kylie McArthur
Owen McCarty
Hayden Smith



Title	Artist	Year	Beats Per Minute (BPM)	Energy	Danceability	Loudness (dB)	Popularity
She Will Be Loved	Maroon 5	2002	102	71	71	-6	74
Dirty Diana	Michael Jackson	2003	131	80	75	-5	55
Fireflies	Owl City	2009	180	66	51	-7	79
Will You Be There - Single Version	Michael Jackson	2005	83	63	41	-7	53
Black or White - Single Version	Michael Jackson	2005	115	89	74	-4	64
The A Team	Ed Sheeran	2011	85	29	64	-10	76

Demo Video Link:
https://youtu.be/Dzwk4SuGwyY

The aim of this visualization is to help a user sort through a long list of possible songs using their attributes to create a playlist tailored to specific desired characteristics. This was done through the use of a main interactive parallel coordinate plot and a supporting table.

The song attributes dataset, as provided in class, was used. Since this project had a focus on pop songs, the main dataset was filtered down to its pop-only songs. This way, the smaller dataset

also could be displayed more neatly on the chart. A parallel coordinate plot was selected because the data was multivariate and this chart type facilitates the visualization of such data—especially when it is beneficial for the different variables to have their own axis. Other options, like a stacked bar chart, were not chosen because, for example, the variable "Loudness" is negative, which would make those chart types less usable. A parallel coordinate plot is also able to provide the user with the ability to easily filter on the variable(s) they are interested in, which was one of the main goals of our design. Additionally, since the data is represented in the form of a line it is relatively easy to see trends in the data.

Six total variables were chosen: the release year, beats per minute, energy, danceability, loudness, and popularity. These attributes were chosen because they give a good representation of the song characteristics which the user may be interested in. Among the other available attributes that were not used, speechiness and liveness were not included since these variables did not vary much between all of the pop songs in the dataset. The song length was also not used since the user is presumably interested in the musical characteristics of the songs, not the play time.

Other important data included the song name and artist. These variables were not visualized with the parallel coordinate plot since that would result in less user-friendly visual noise—the only options to show these on the plot would be to either have two extra axes or through a tooltip-type method where the user is required to hover over every single song to find its name. In addition to being difficult to view if there were axes for the song name and artist, there would be no meaningful way to order the values. Instead, a supporting table was included below the chart, which lists all of the relevant data, including song name and artist, that the user has selected. When the user brushes the plot axes, the table updates accordingly. This way, a clean and concise method of showing the selected data is offered. Additionally, any song names longer than 35 characters in length were trimmed to keep the data in the table neat. All songs in the data set were uniquely distinguishable even after this post-processing.

Our group worked together to create this design primarily based around the parallel coordinate plot. The most involved challenge was getting the multi-axis brushing to not only work but also correctly filter the data and link to the table. A few methods were used to solve this problem, including natively in D3 with SVG objects and also with line drawing on a canvas. Additionally, options for the display table, including a traditional HTML table and a simple tab-space table, were explored. Team members worked on these different solutions to find the best solution for a final configuration.

Ultimately, this project offered the team an opportunity to not only create a useful visualization using good practices as learned in class, but also exercise the use of D3 and other web tools. The visualization reflects this with a clean and visually appealing, interactive design which allows the user to appropriately filter data to see songs that share various attributes.