Spotify Popular Song Attributes

Created in d3.js

Aria Wang, Ellie Tan, Jue Wang, Shirley Liu, Zach Zhu

Background

According to the most recent spotify data, the average amount of time spent listening to music on Spotify by the platform's active user was 25 hours per month.

We aim to build a visualization tool for music listeners and Spotify to understand music, trend, and app usage better by providing insights from a variety of dimensions including audio features, time, title words, and monthly change of stream counts across the world.

In our visualization system, we used the data "Top Spotify songs from 2010-2019 - BY YEAR" from Kaggle. The dataset contains the information about the song attributes, stream counts, popularity, etc. for 130,663 most popular songs worldwide on Spotify from 2010 to 2019. (https://www.kaggle.com/leonardopena/top-spotify-songs-from-20102019-by-year)

System Introduction

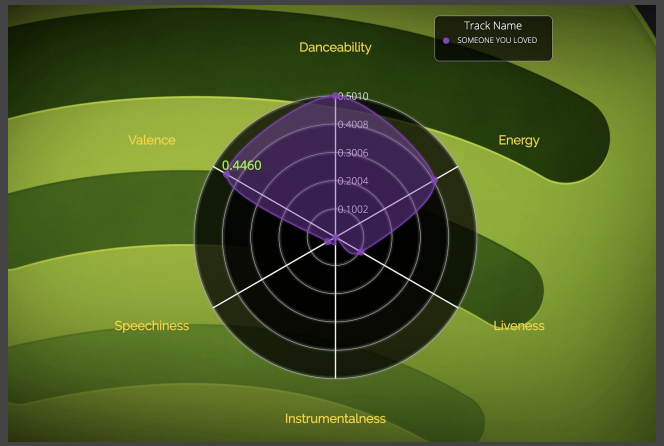
The system is composed of a homepage and four interactive visualization elements:

- Homepage: navigate to any visualizations
- Radar Chart: audio features of a song or comparing audio features of two songs
- Bubble Chart: popularity, danceability and valence of the songs by a specific artist
- Bubble Map: monthly change of stream counts around the world in 2017
- Word Cloud: popular words in the 130,663 top songs' titles from 2010 to 2019

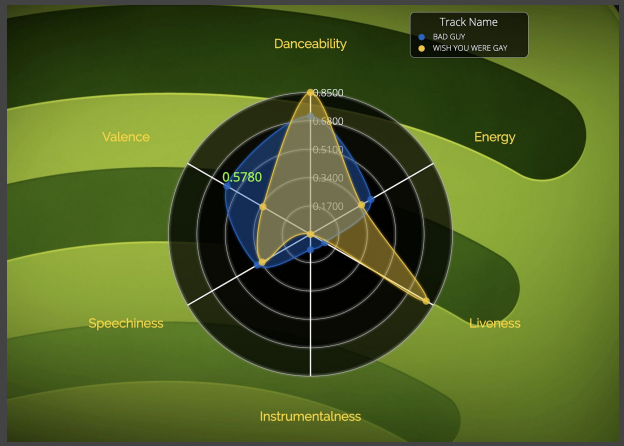
* Users can go back to the home page anytime

Approach 1: Radar

• For each song that the user inputs, our system creates a radar chart showing the scores of six audio features: **Energy, Danceability, Valence, Speechiness, Liveness and Instrumentalness** (exact values will show up when mouse hovering).



• Users can also compare two songs' attributes. Our system will output two radar charts layered on top of each other with the same-scale coordinates (the corresponding song will be highlighted when mouse hovering).



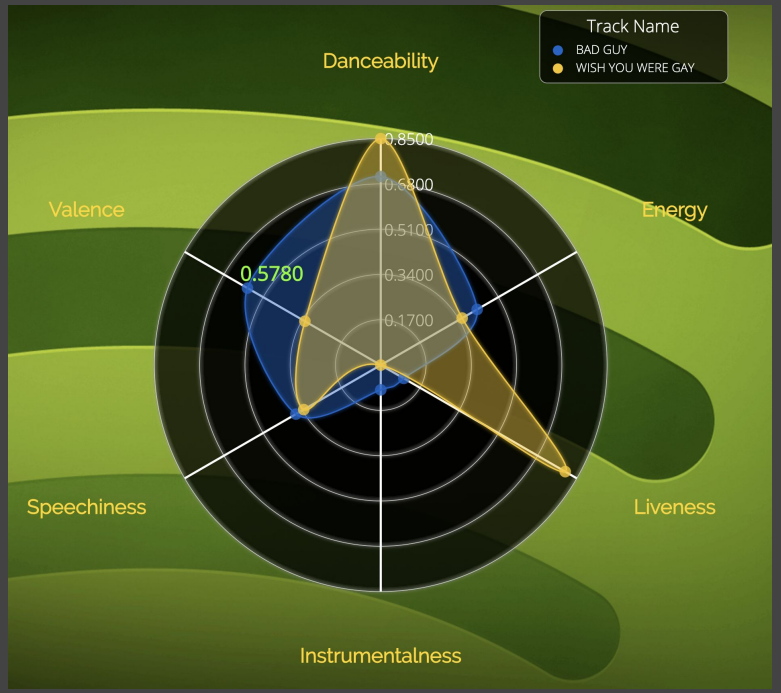
Design Choices

Marks

- **Point** - represent the corresponding attributes value
- **Area** - represent the song's overall performance in six dimensions

Channels

- **Size** - the larger area indicate higher values in given attributes, doesn't indicate goodness of a song
- **Color** - differentiate two songs by two different colors on the same radar plot for straightforward comparison



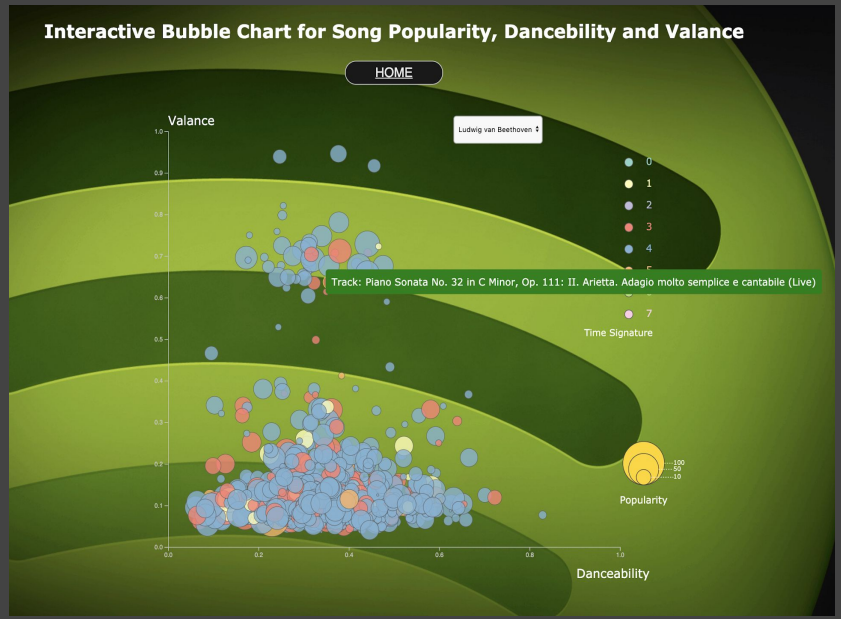
Strengths and Limitations

- **Strengths**
  - Displays and compares different attributes for one song. Users can learn about the strengths and weaknesses of the songs;
  - Compares multiple songs in terms of different attributes on the same scale (layered radar charts).
- **Limitations**
  - Different orders of attributes may lead to significant difference in size / sight interpretation;
  - It may be hard to compare the attributes of more than 2 songs;
  - The differences in attributes may be exaggerated by scaling
- **Alternatives**
  - barplot and lollipop plots

Approach 2: Bubble

Users can choose an artist with at least 10 songs listed from the drop down menu, and our system will output a interactive bubble chart visualizing all the songs by this artist.

Each bubble represents a song's popularity, danceability and valance. The colors represent time signatures. Track name is shown when mouse hovered to the corresponding bubble



Marks

- Point - represents a song. Corresponding axes are the values of danceability and valance

Channels

- Size - indicates the song's popularity
- Color - displays the time signature of the track

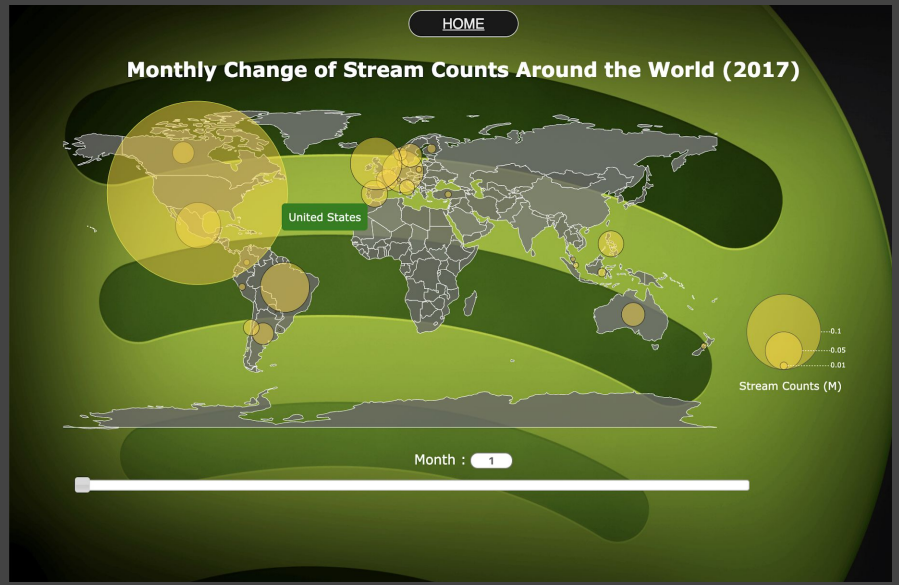
Strengths and Limitations

- **Strengths**
  - Let users explore how time signatures affect musicality of songs
  - Aggregate songs by artists, and let users choose the artists they are interested
  - Bubbles have transition animations when users switch artists for better UX
- **Limitations**
  - Bubbles (songs) that have low popularity index or stacked together may be hard to spot on the chart
  - Too many artists in the drop-down menu may be annoying for users to find a specific artist
- **Alternatives**
  - Density Plot

Approach 3: Bubble Map

The bubble map presents the average streaming counts of top songs in 2017 by month across different countries - the corresponding country names will show up when mouse hovering over the bubbles.

Users can either input month in the box or drag the bar to view the monthly streaming count changes over 2017.



Marks

- Circle - represents different regions

Channels

- Size - indicates regional streaming counts

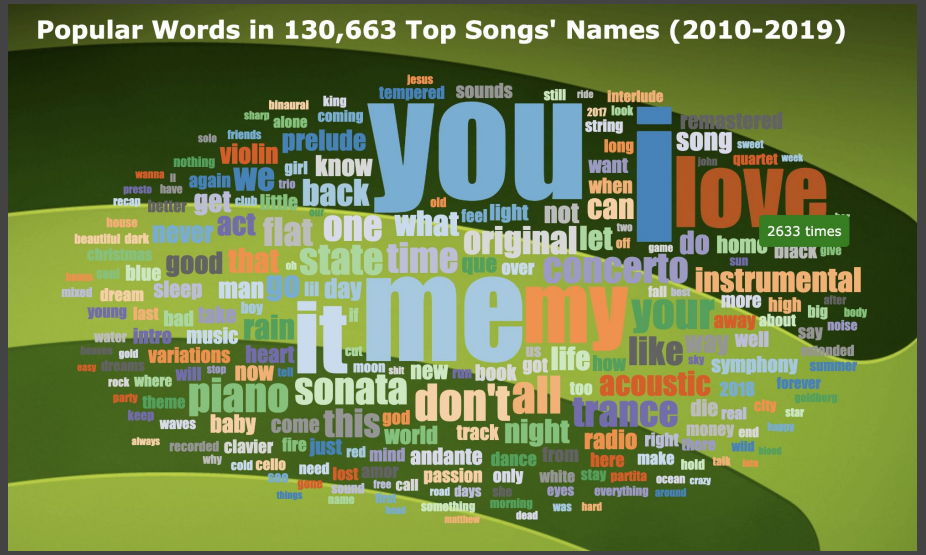
Strengths and Limitations

- **Strengths**
  - Users can clearly see the change and trend of stream amounts across months
- **Limitations**
  - Hard to spot the regions with low streaming counts;
  - Some countries were not mapped to its corresponding longitude and latitude due to software limitations, and hence their streams were not included in the map;
  - Might be more helpful if the map is at song level rather than aggregated level;
  - Relatively small changes in the bubble size cannot be easily captured by eyes;
  - Large bubbles covering other regions may be misleading (eg. US covers canada and Mexico)
- **Alternatives** - Interactive line chart where users can select the month manually

Approach 4: Word Cloud

The word cloud presents the most popular words that were used in the titles of 130 thousand top songs.

When the mouse hovers over a word, it will show the frequency of the corresponding word appeared in all song titles.



Marks

- Text - represents the word

Channels

- Size - indicates words' popularity

Strengths and Limitations

- **Strengths**
  - Novelty visual representation of text data, can be used to keyword searching
  - Give greater prominence to words that appear more frequently which is easy to understand and visually appealing
- **Limitations**
  - Words' colors and locations in the cloud has no meanings;
  - Longer words are easier to be spotted and have delusional larger size;
  - Hard to distinguish between the sizes of less frequent (smaller size) words;
  - For our data, except for the top most frequent words, the rest of the words have very similar frequencies, which makes the text sizes difficult to scale;
- **Alternatives** - treemap and bar charts

Insights from the system

From Radar Chart:

- Songs with similar popularity may have drastically different musicality measures.

From Bubble Chart:

- Most pop songs have time signature of 4/4, while classical music have more variations in time signatures
- Many songs with high popularity also have high danceability and valance scores

From Bubble Map:

- Stream counts are highest in holiday season (December, January), and lowest in February
- Different countries may have opposite direction of change in stream counts over the months

From Word Cloud:

- Songs with the word “Me”, “You”, “Love”, “I” appear the most

Overall Evaluation

The objective to design this visualization system is for Spotify users to better understand their music habits as well as the global trends among the huge amount of music Spotify provides.

The system provides robust interactive visualizations through the radar chart, bubble chart, bubble map and word cloud. Deep analysis is conducted and available for users to learn more about individual songs and artists. However, limitations do exist. It does not perform well on visualized global trends. Further development and optimization could be implemented.