Milestone 3 - Project Proposal

Basic Info

- Project Title: How Does Technology Advancement Affect Music Artist Popularity?
- Member Information:
    - Michael Harris | mah6@clemson.edu | C12977960
    - Nicolas Mowll | nmowll@clemson.edu | CUID
    - Thomas Personett | tgperso@clemson.edu | C11137890
- Link to Repository: https://github.com/tgperso/CPSC4120-Project
- Link to GitHub Pages: https://tgperso.github.io/CPSC4120-Project

Background and Motivation

The music industry has been through quite the technological revolution. From old-school vinyl records to the digital streaming era, the changes have been massive. And smack in the middle of this transformation is the Billboard Hot 100 chart, chugging along since 1958, telling us which songs are making waves in the U.S. Now, the way we measure song popularity has morphed over time, mainly due to technology. Back then, it was all about radio plays, but now, it's about streaming numbers from a wide array of platforms and media. Shoutout to Dylan Burati for the dataset from Kaggle, which gives us a solid chunk of Billboard Hot 100 data. Our mission here is to grasp how technology and music artist popularity are intertwined. We aim to dive into this data stash to extract trends, connections, and stories that showcase how technology shaped artists' careers, music genres, and the ebb and flow of popularity.

Our motivation is that we want to understand how technology influenced artists' journeys and share those insights. We all like both music and technology and find it very interesting to see exactly how our favorite artists raised themselves to the top of the charts.

Project Objectives

1. Artist Popularity Over Time: We're going to create visuals that illustrate how artist popularity has fluctuated over the past 80 years. Essentially, we want to see who was in the limelight and when they faded into the background.

2. The Tech-Popularity Connection: Explain how technology and artist popularity are closely linked. We will demonstrate how tech advances, like streaming, social media, and mediums have impacted artists' fame.

3. Social Media's Influence: Show how platforms like Twitter, Instagram, and TikTok have influenced artists' popularity.

4. Medium Influence: From records to cassettes to CDs to MP3s to streaming, explain how album release mediums have influenced artist popularity.

5. Connectivity: Explain how the world being more and more connected affects how artist popularity is distributed.

6. Visually Impactful: We'll use colors and designs that make the data easy on the eyes.

Data

The dataset used is a Full History of the BillBoard Hot 100 songs from 1958-2021. The Hot 100 acts as sort of a standard for how the U.S. measures song popularity, and has been a popular form of musical judgment for eighty years. BillBoard's data gathering methods have changed over the years, but generally they combine several components that contribute to the overall chart (now, they combine radio plays and streaming listens). Credit to Dylan Burati for the dataset *Billboard Hot 100 - Full History* downloaded off of Kaggle.com.

Source: https://www.kaggle.com/datasets/dylanburati/billboard-hot-100-full-history

Dataset:

Table: billboard-hot-100-credited-artists

Description: provides info on specific artists for each track

Items: Artists

Attributes:

- Song_id: FK - unique categorical identifier for each track

- Artist_id: PK - unique categorical identifier for each artist

- Artist: Categorical, credited artist's name

- Ordinal: Ordinal, index showing the position the artist was listed on the track

Table: billboard-hot-100-tracks

    Items: Song Tracks

    Attributes:

    Description: provides info on every ranked track

- Id: PK - unique categorical identifier for each track
- Rank: Quantitative - Rank of the track on the Hot 100
- Date: Categorical - Date of the chart
- Artist: Categorical - Credited artist(s) on the track
- Song: Categorical - Song title
- Last_week: Quantitative - Ranking on chart for the previous week
- Weeks_on_chart: Quantitative - Total # of weeks on the chart
- Peak_rank: Quantitative - Highest rank achieved

Data Processing

The dataset is already clean, but because of the few attributes, I derived two additional columns in order for us to reach our objective. For now, I used Tableau to calculate the values of the two new columns.

Derived Columns:
- Popularity Score - Quantitative - Popularity of the song
    - Because we need a metric to show popularity of an artist for a specific time frame, we need a "popularity score". The Billboard Hot 100 data consists of every week's top 100 songs, so the number of listens/streams is not applicable. Therefore, I decided to calculate a weighted popularity score for each song based on the ranking it was given.

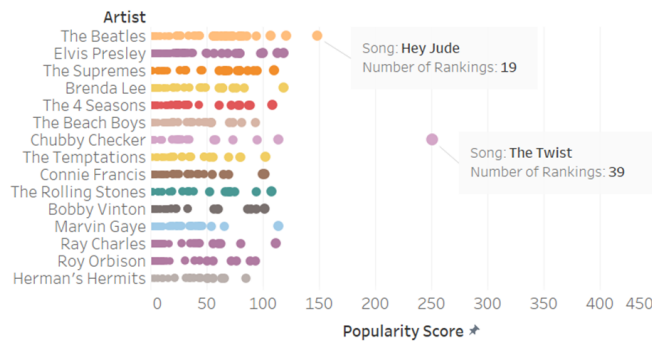   Popularity Score = (POWER(101-[Rank], 5))/1000000000

(The exponential inflates the importance of rank)

To then see an artist's popularity score, I sum up the popularity score for each of their ranked songs for the filtered time frame.
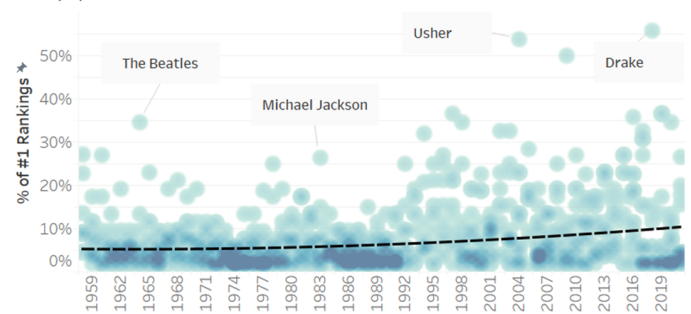
- TopRankedOnly - Quantitative - Shows 1 if song was ranked #1, else null
    - This column is useful for our visualization showing trends about only #1 ranked songs.
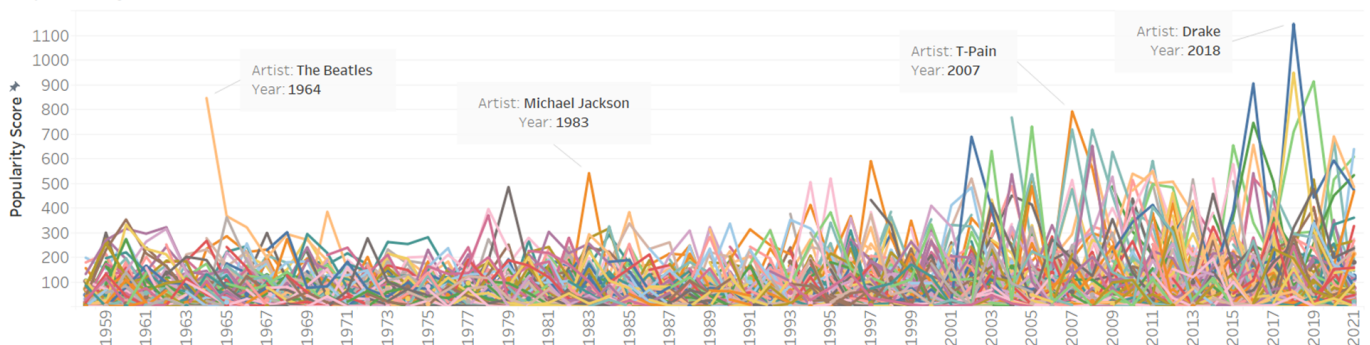
Visualization Design

Artist Popularity in the 60's/70's/80's/90's/2000's/2010's
- Dot plot ranking artists from most to least popular for a certain decade (60's listed here). Each dot is a song positioned at its popularity for each artist. One purpose of this chart is to display interesting details about the rankings of artists and how their songs contributed to their popularity. The main purpose is to show the

change in popularity distribution of top artists over each decade. As the user iterates through decades, songs become more and more popular, showing a wider distribution in artist popularity as a whole.

- Mark: Points - Ranked Song Tracks
- Channels:
    - Position (horizontal) - Showing the popularity of the song
    - Position (vertical) - Showing the popularity of the associated artist
    - Color - Grouping songs together with their artist

Why this prototype?
- Initially, the prototype had an interactive year slider where the user could scroll through each year. While this is fun to see certain years with outliers, users cannot get a grasp of the overall trend we wish to demonstrate for our goal. Split into decades, the change in distribution of popularity is much more visible.

Distribution of #1 Rankings across Artists
*May become a colored Streamgraph showing distribution of #1 rankings

- Density plot showing the distribution of the percentage of #1 rankings that each artist received for each year. The purpose of this graph is to show the change of distribution of #1 rankings over time. The density of ownership over the #1 ranking spreads out over the course of history, showing a higher number of outliers and more extreme outlying values.

- Mark: Points - Artists
- Channels:
    - Position (horizontal) - The year that this artist received a percentage of #1 rankings
    - Position (vertical) - The percentage of #1 rankings the artist received

- Color - The density of the number of artists that exist in a certain percentage. Darker = more artists at the given percentage, Lighter = Less artists at a given percentage

Why this prototype?
- Initially, the prototype showed number of #1 rankings. While the visualization looked nearly identical, because there is a set amount of #1 rankings each year (52 for each week in the year), a percentage value is easier to grasp and makes more of a solid point.

Popularity of Artists over Time
**May become a colored Streamgraph showing distribution of popularity

- Line chart showing the popularity of every ranked artist for each year. The purpose of this graph is to notice the major outliers which we will address in our descriptions and analysis. Also, the user can once again see the change in popularity distribution between artists. This graph demonstrates that the top artists in recent years are much more popular than the top artists of older decades.

- Mark - Line - Artist popularity over time
- Channels:
    - Position (2D) - year and popularity of the artist
    - Color - Artist (meant to distinguish between lines)

Why this prototype?
- The other prototype version used density like in the #1 ranking distribution graph, however for this version, the line graph paints a better picture, offering the same trend of change in popularity distribution of artists but also more clearly shows the change in their popularity over time.

Must-Have Features

- Visualizations that show the trend of the distribution in artist popularity over the last 80 years
- Visualizations that express the correlation between technology advancement and saturation of artist popularity
- Visualizations that are easily digestible and understandable for a viewer
- Visualizations with interactive elements that make them more interesting and impactful
- The connection between technology advancement and and the change in music popularity should be clear and easy to follow
- Showing how social media has impacted the recent music industry and popularity
- Visualizations with clear titles and labels that allow the viewer to understand the data correctly
- Visualizations that use intuitive and impactful marks and designs to represent the data effectively
- Visualizations that use time to represent different cultural and musical trends

Optional Features

- Visualizations with filters of specific popular bands and artists that are more relatable to the viewer
- Breaking time data down further to show more specific trends over time, like a year long trend
- Import genre data and information to see trends throughout certain decades and music types, along with filters
- Use different types of marks with similar data to show more trends, like density and linear marks
- Visualizations that contain a high level of interaction especially with each band or artist

Project Schedule

- Week 1 (10/9-10/13):
    - Researching D3js and other visualization design tools
    - Use Tableau to format our ideas correctly

- Make sure data is correct and sufficient to complete our objectives
- Week 2 (10/16-10/20):
    - Each member gets one visualization to focus on
        - This will allow us to have 3 really great visualizations
    - Use D3js and other design tools to format and create designs
- Week 3 (10/23 - 10/27):
    - Continue creating visualizations and aim to have them finished by the end of this week
    - Results in three finished designs
- Week 4 (10/30-11/3):
    - Add visualizations/prototypes to website and wrap up design and formatting of web page
    - Turn in Milestone 4
- Week 5 + 6 (11/6-11/17):
    - Use feedback to update and improve existing design
    - Create interactive features of visualizations
    - Complete Milestone 5: Peer Review
- Week 7 + 8 (11/20-12/1):
    - Improve existing designs from peer review feedback
    - Create presentation of website and interactive visualizations
    - Practice the presentation and decide who will present what
- Week 9 (12/4-12/8):
    - Present presentation in front of class
    - Watch other presentations
    - Use presentation feedback for project screen
    - Wrap up project report and code
    - Complete peer assessment