

# Album Trends (1990-2021) Process Book

Group 9

Madeline Oliver and Kristina Zuniga

---

**Github Repository :** <https://github.com/madelineoliver/madelineoliver.github.io>

**Github Page:** <https://madelineoliver.github.io>

**Video Presentation:** <https://youtu.be/qYgfrcnVpf0>

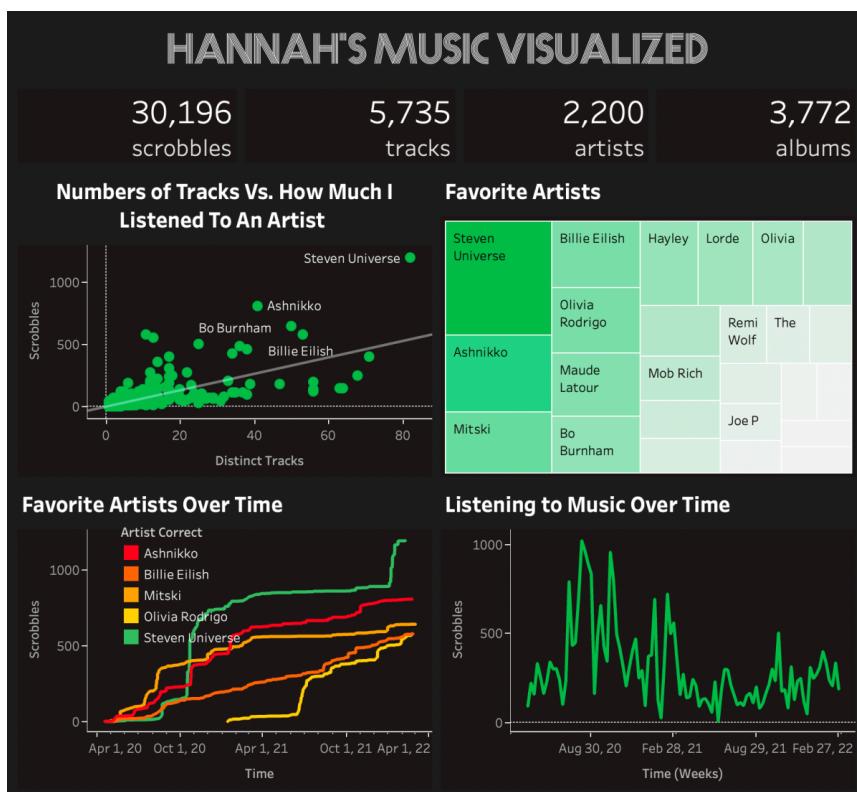
## Overview and Motivation

Our motivation for this project stemmed from our interest in music, more specifically we were interested in popular albums and what possible changes there have been to popular Album's composition over the past 31 years. In today's world, our attention spans have seemed to significantly decrease with the popularity of apps such as Tik Tok, which feature short clips designed to grab your attention. We were curious to see if there were any trends regarding the composition of popular albums over time, and as a result of this interest, we made it our goal to identify changes in genre, album length, track number, and overall sales of the Top Ten Albums per year from 1990-2021.

---

## Related Work

When thinking of the aesthetics and the data that we wanted to show, we looked at Spotify stats that have been visualized. Spotify is a music streaming app and every year does a 'spotify wrapped' which shows a user's streaming habits throughout the year. Looking at visualizations related to this, we thought it would be interesting to do something similar in terms of showing listening habits, but rather than for a particular user it would be more general (the whole world), and it would relate to albums rather than songs.



We thought the layout of a visualization like this would be interesting, showing different trends related to music but on a more general scale rather than for an individual person. We thought it was interesting in terms of how music can be visualized.

---

## Questions

Initially when looking at the dataset, the questions we sought to answer were more simplistic in nature:

- Has there been a noticeable trend in the length of albums in the past 31 years?
- Has there been a noticeable trend in the number of tracks of these albums in the past 31 years?
- Has there been a change in genre preference in the past 31 years?
- Is there any trend regarding the ranking of the albums and the length of them?

However, after examining the data and the different attributes, we realized that there were more related questions that we could explore with our visualizations. These questions evolved from our experimentation with the data, looking at different attributes, how they could be related, and how they make up the entire item in the dataset and present it in different ways. These questions helped shape our final visualizations into a related interactive system, rather than aggregated data:

- Was there any correlation between an album's genre and their length, were certain genres more likely to have shorter or longer album lengths?
- Was there a noticeable trend in the overall world wide sales of popular albums from 1990 - 2021?
- Was there a change in the number of tracks related to album length from 1990 - 2021 (ex. Did albums have less tracks with longer songs, the opposite, was there no correlation?)

---

## Data

Site we got our data from :

<https://www.kaggle.com/datasets/nickadair44/top-10-annual-best-selling-albums-by-length>

## Data Scraping

We did not scrap any data for this project. We considered scraping geographically data on the sales of albums per country, however we decided that the specific geographical information was not relevant to our bigger picture and the main question we wanted to answer, which was related more so to album length and genre. We thought that adding geographical data would expand the focus too much and lessen or make it more difficult to visualize our other attributes.

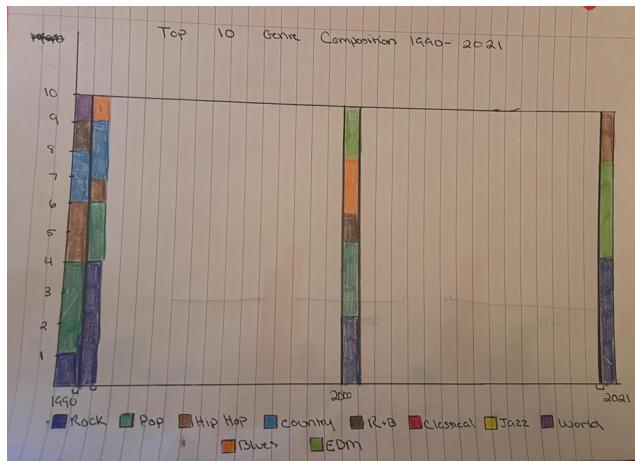
We used filters, arrays, and key value pairs to group data needed for each graph generated.

## Data Cleanup

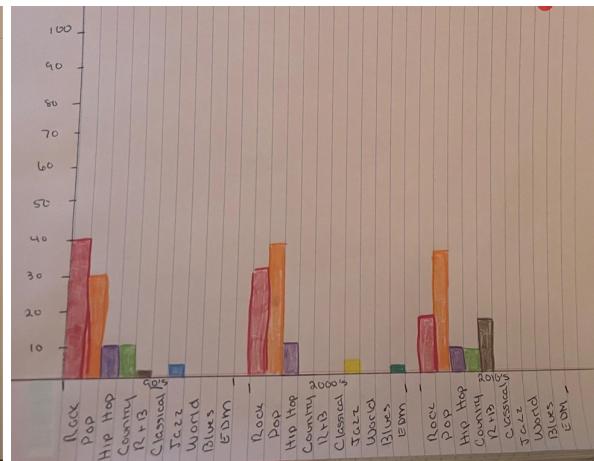
Our data and csv file was already well formatted, however we did make some adjustments to it. We took out the commas for the world wide sales attribute as well as took out the spaces for Album length, Hip hop, and world wide sales as well as changed "R&B" to "R and B", to make the data more accessible.

# Exploratory Data Analysis

Initial Visualizations:

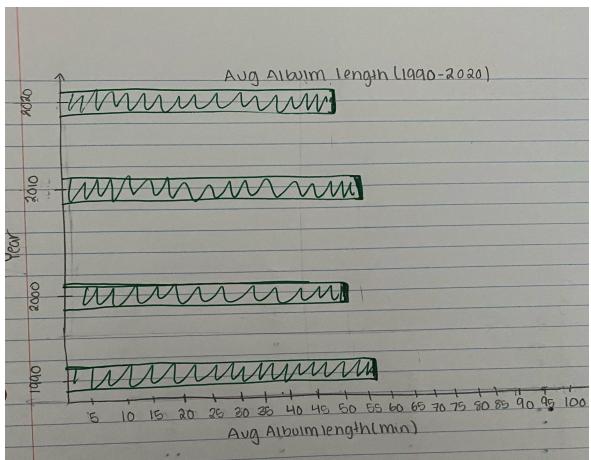


Img 1



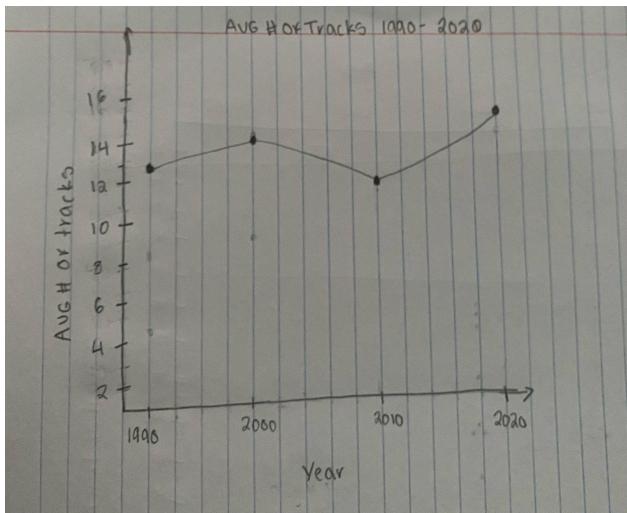
img 2

Initially we tried exploring different ways to display genre trends. We thought the best way to do that would be with a bar graph, since there were different categorical attributes (the genres) for each year, this visualization made sense to us. First we tried using a stacked bar graph(img 1) to show the different genres for each year, Y axis would be the associated ranking 1 to 10 of the genre and X axis would be years. Our next Idea for genres was to create a cluster bar graph. Y axis would be the number of albums in each genre and X axis would be cluttered by year and each bar would be a genre. From these two designs we came to the conclusion that a bar graph would be the best graph to use for the data.

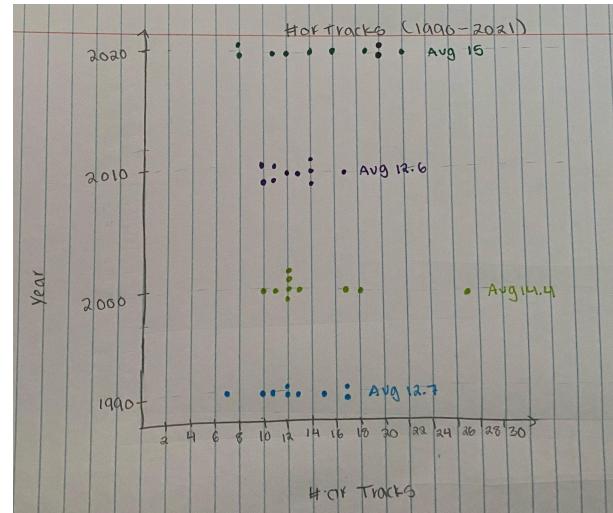


We passed ideas back and forth and considered displaying album length with a bar chart. Y axis would be Years and X axis would be Album length in minutes.

img 3



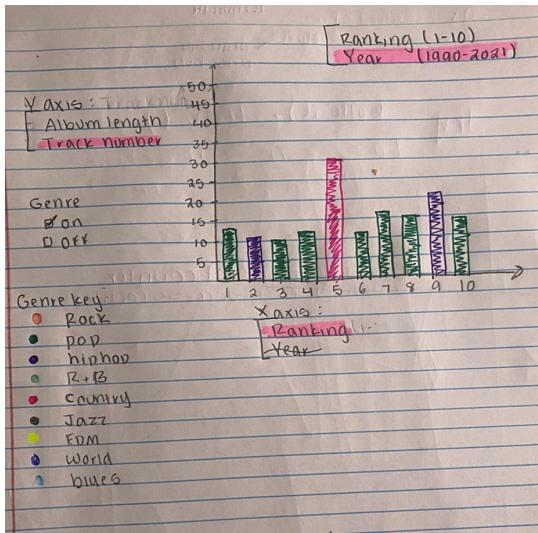
Img 4



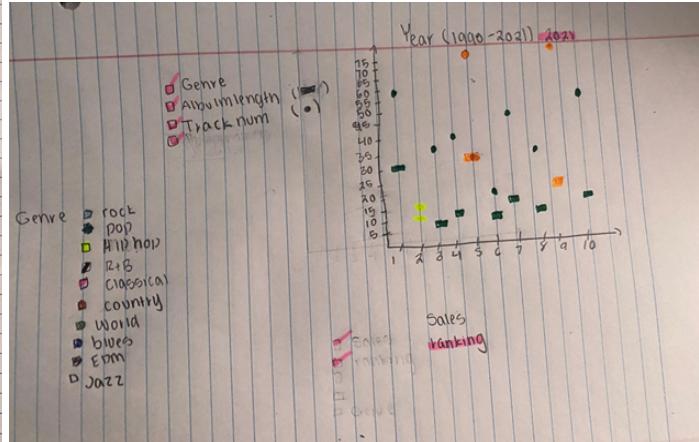
img 5

We explored different ways to display the number of tracks for each album. First we considered using a line graph to show the number of tracks over time. The Y axis would be the avg number of tracks and the x axis would be years. Our second idea was to use a scatter plot that would display all the albums tracks for each year instead of averaging it out. The Y axis would be years and the x axis would be the number of tracks. We concluded that we needed to decide if it was better to show all the data or average the data and determine which would best show the overall trend.

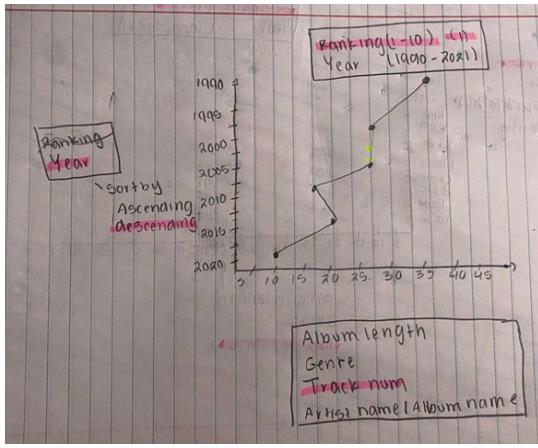
## Design Evolution:



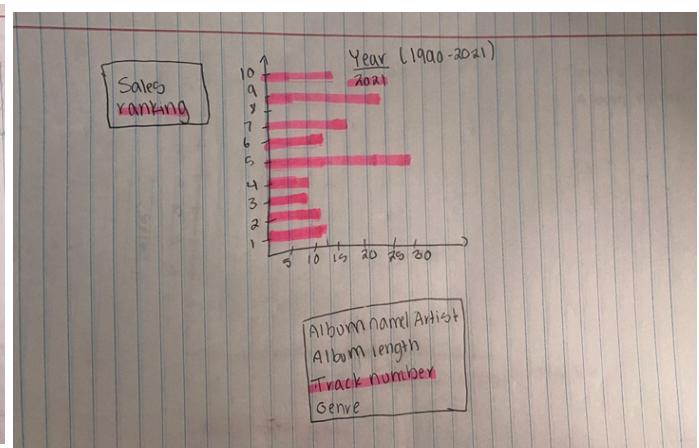
Img 6



Img 7



Img 9



img 10

## Milestone 3: Ideas

Originally when designing our data for milestone one we wanted to make one graph that would have filters that show different data sets. Images shown above (img 7-10)

---

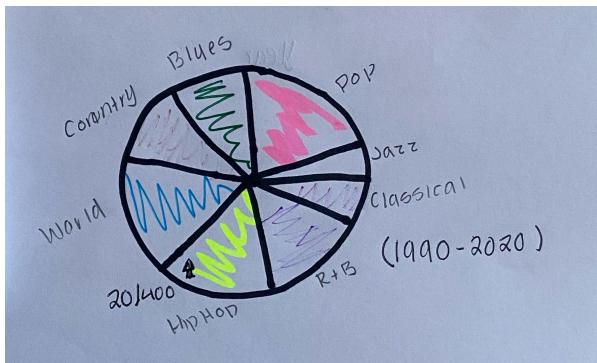
So you would have an x and y axis with different buttons on each axis that allows you to pick and choose what data you wanted to see. For example in img 6's graph on the Y axis you have album length or Track number and the X axis has the option of ranking or year. You also had the option to see genres if a filter for it was clicked and you could choose to display the x and y axis depending on year or ranking. Whichever you choose for this option would automatically cross out that option in the x axis. We had similar ideas for img 8-10 but the way the data was displayed changed. Image 8 displayed data with scatter points, Image 9 displayed data with a point line graph, and image 10 changed the bars to be horizontal. We also tried switching the data options on the X and Y axis. We thought all these ideas were creative at the time however, we later realized that all the filters made the graphs hard to read and we were not able to see all the trends we wanted to explore at once. Looking back I feel that although we created a jumbled mess of a graph, creating these graphs really helped us to visualize all our ideas and gain a better focus on what we actually wanted to show.

#### **Milestone 4: Prototype**

After doing milestone 2 and getting all our ideas on paper, we asked ourselves "How easy is it to extract information from our graph," and realized that there was too much going on, making the graph difficult to read. We then asked ourselves "how can we improve the design", and decided that the best course of action was to split up the graph into multiple graphs.

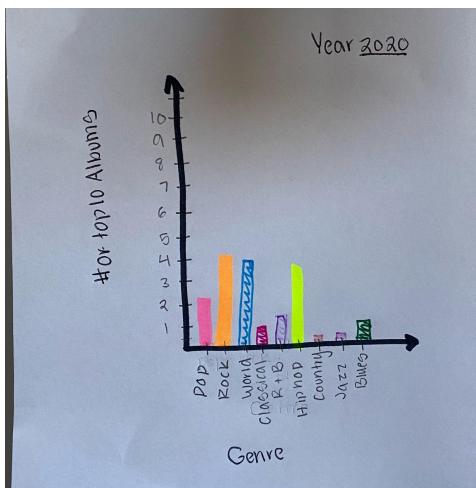
---

## Genre Graph evolution

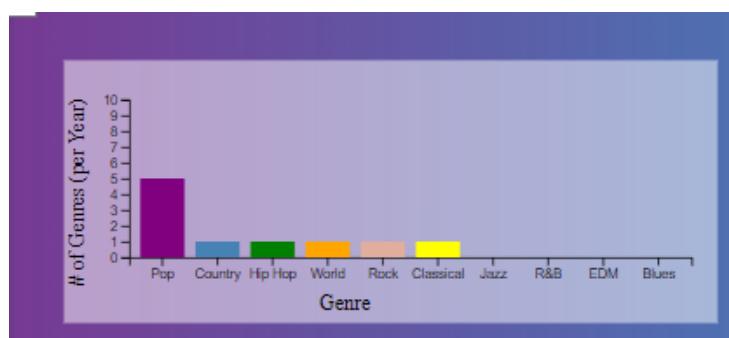


First we looked at the genres and considered putting it into a pie chart(img 11). The pie chart would have all the genre percentages from 1990-2021 in one graph. However we once again asked ourselves "How easy is it to extract information from our graph", and realized that the data may become distorted to the viewer if the slices are close in size. We thought about adding percentage labels but decided to use a bar chart instead( img 12).

img 11



img 12

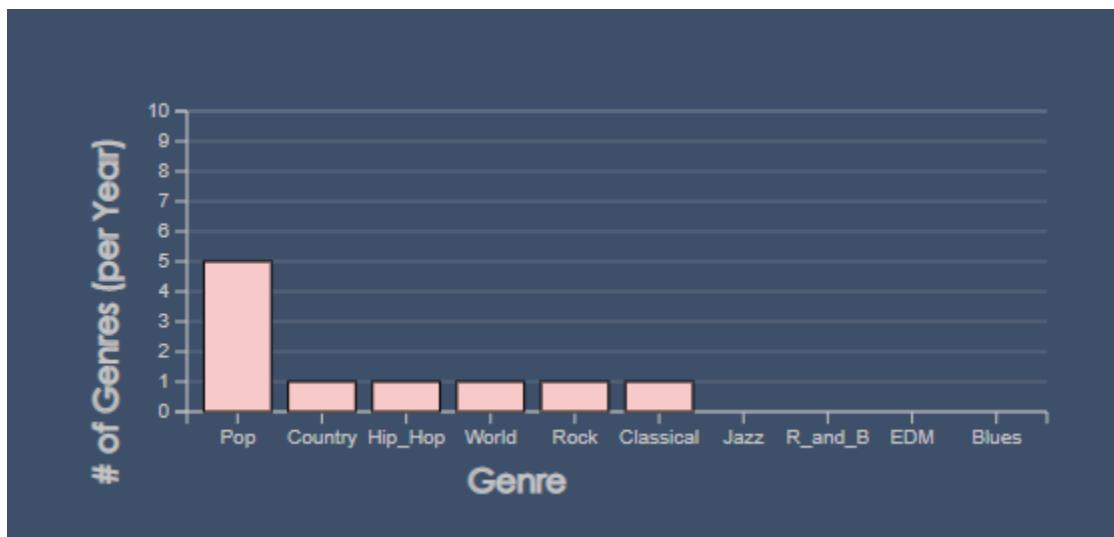


img 13

With our bar chart (img 12) the genres would go at the bottom and the number of top 10 albums in that genre would go on the Y axis. We also decided to display the graph by year instead of trying to put all the data for 31 years into one graph. We also decided to color code the genres to make them easier to distinguish. Once we had the idea down we started implementing and ran into some issues.

---

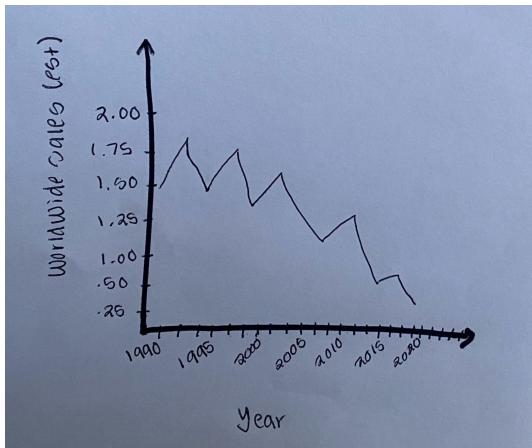
The first issue we had was finding a way to group Genres, genre number and years together since we wanted a drop down to be able to change the data displayed. We solved this problem by making an array of years with a key value pair, the key was the genre and the value would be the length of the genre( the amount of genres in the top 10 albums). The second issue we had was figuring out how to update the bars. When grouping the data, the genres that did not show up that year would not show up in the array at all. As a result, if Genre Pop was 2 for a year and the year it was switched to was 0, the bar would not change because there was no value in the array for pop since it did not appear that year. Our solution to this problem was to create a bunch of if statements that added the key genre and value 0 to each year even if it did not appear so there would be a placeholder for every genre. As a result of our efforts, we ended up with the graph in img 13.



This Bar graph shows the number of items with a specified genre per year. We decided to not do a color system for the genres as it would clash with the color system for the rankings when the visualizations are used together and would create confusion.

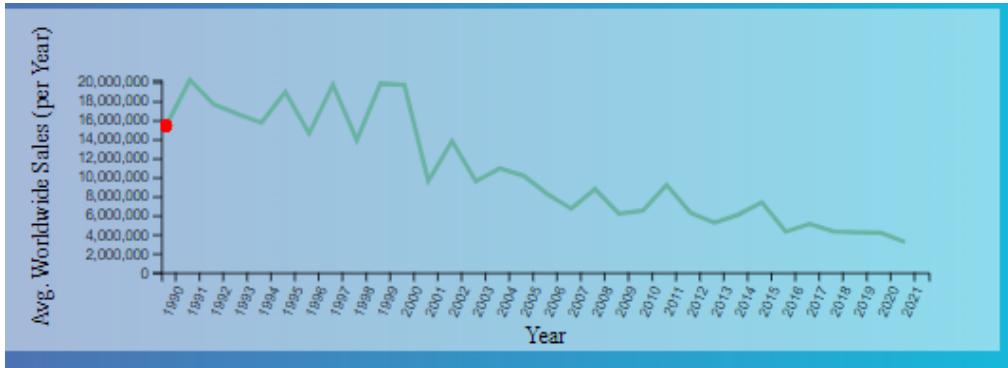
---

## World Wide sales Graph evolution



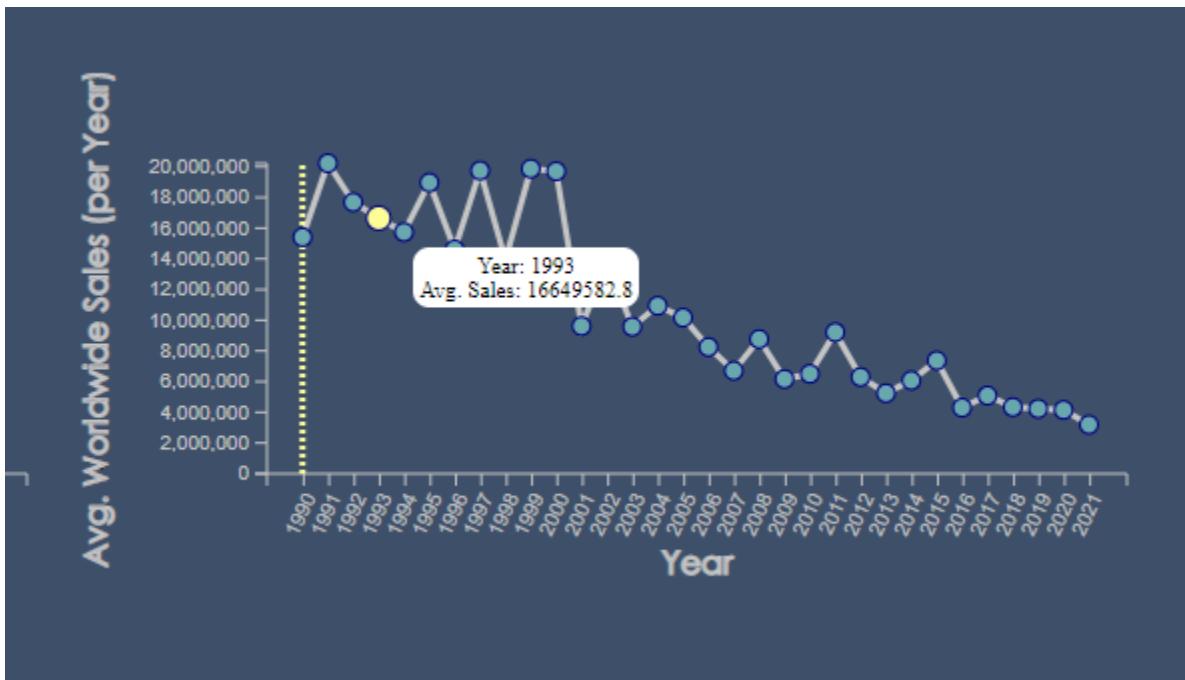
Our initial Idea was to create a graph with the average world wide sales of the 10 albums in each year instead of trying to fit 320 lines of data into one graph (img 14). So, the Y axis would be average world wide sales and the x axis would be years.

Img 14



img 15

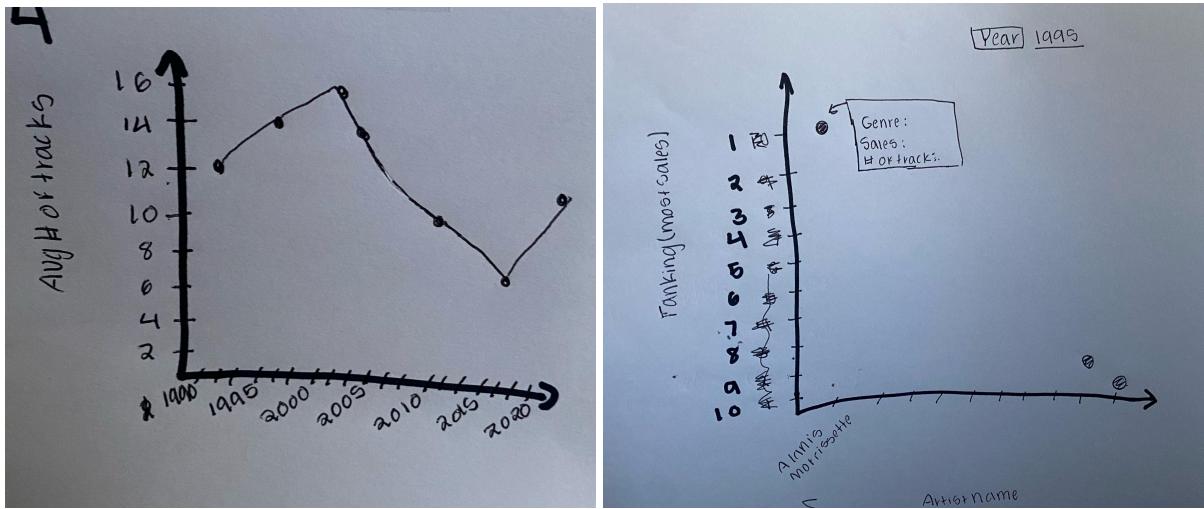
With our initial idea in mind, we began implementing our idea. The data for the Y and X axis of img15 remained the same, but we wanted to somehow connect this graph with the dropdown that would be used for the genres. Our first idea was to have a dot on the line that would move to the year that was selected in the drop down. However we could not get the previous dot to disappear after the new year was selected. We also wanted all the dots to show for each year and we could not get just a single dot to highlight or change color without all the other dots disappearing. Our solution to this problem was found in img 16.



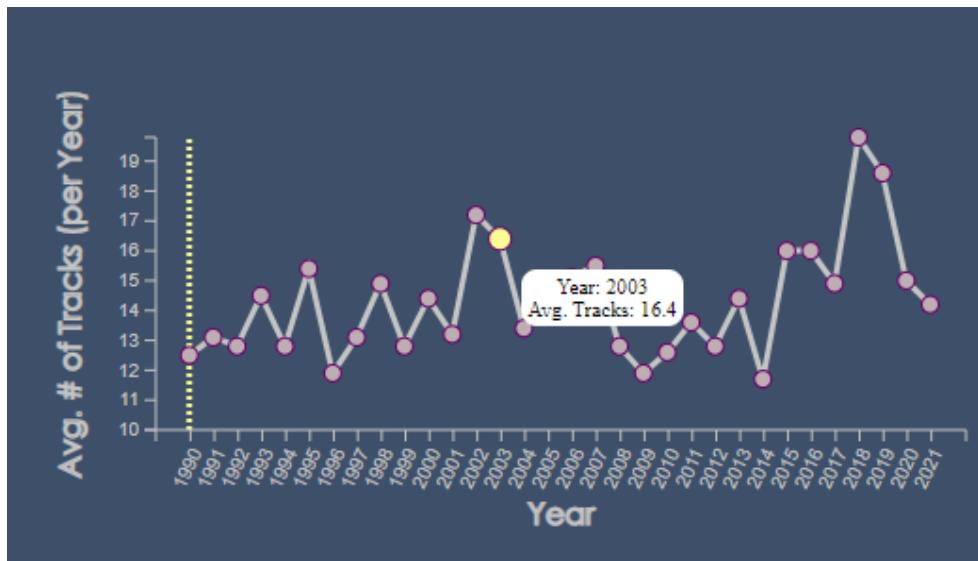
img 16

Our solution to the moving dot problem was to not mess with the dots and instead create a line that would go over the dot of the year selected. We then asked ourselves "how can we improve our visualization" and " how can we decrease distortion of data". Our solution was to create a mouseover and mouse off function that would show the specific year and the associated avg world wide sales for that year. All this combined resulted in our final graph shown above (img 16).

## Track graph evolution



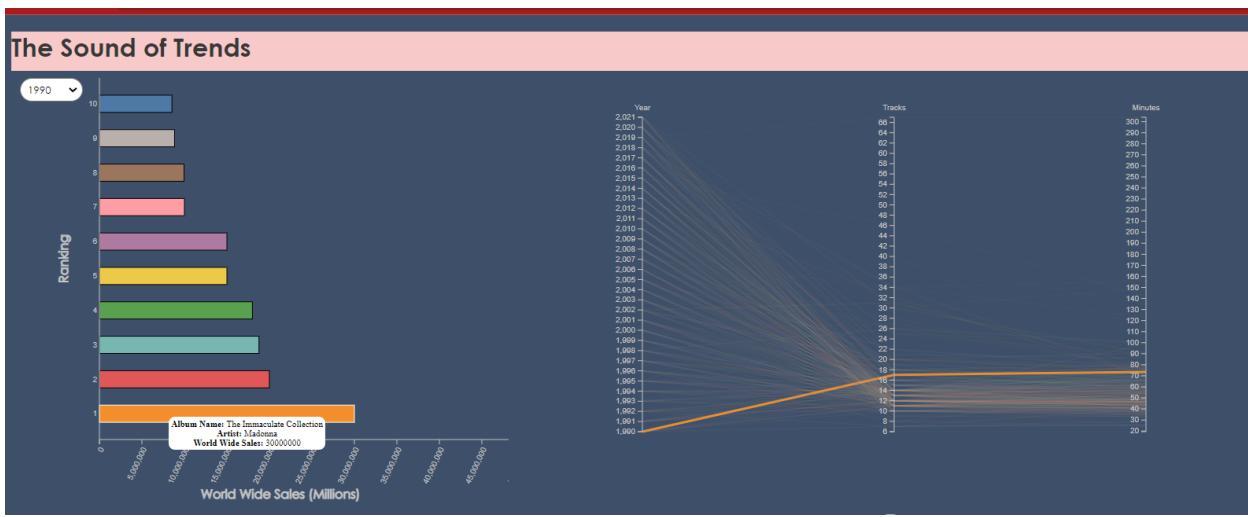
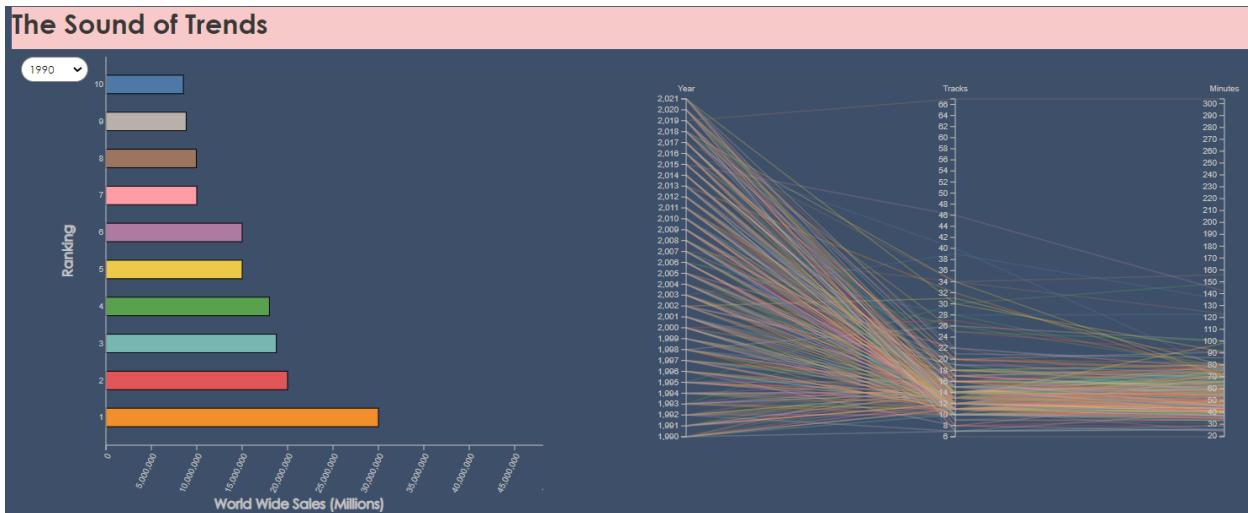
The average track line chart was very similar to the average world wide sales line chart. We initially explored two different possibilities, but in the end implemented it the same as the average world wide sales graph, with the average tracks on the y axis and the year on the x axis. We implemented the same system as with the world wide sales graph, with the line indicating the year selected and when moused over, the specific year and average number of tracks for that point were shown.



# Implementation

The ultimate goal of our interactive visualization system was to answer our initial and evolving questions as well as give an interactive view of the elements in our dataset.

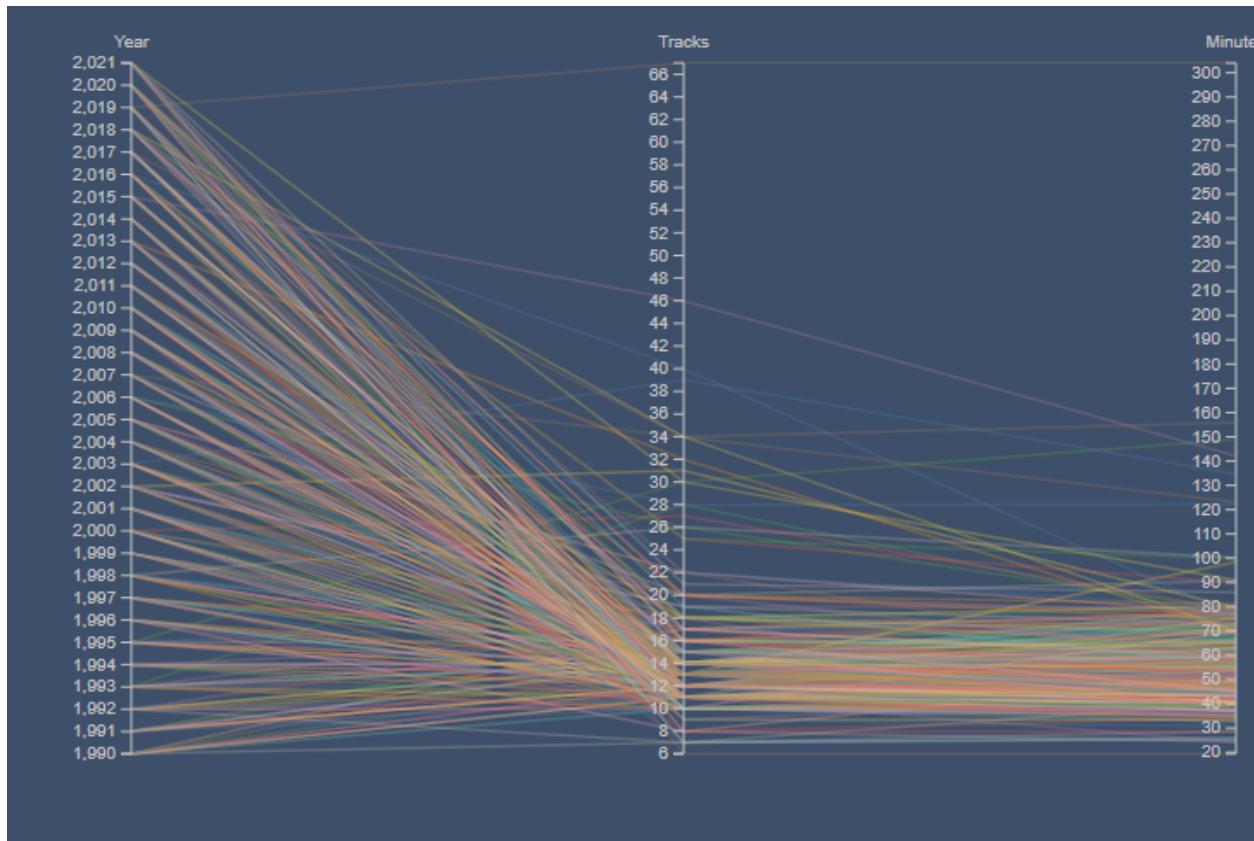
## Rankings Horizontal Bar Chart

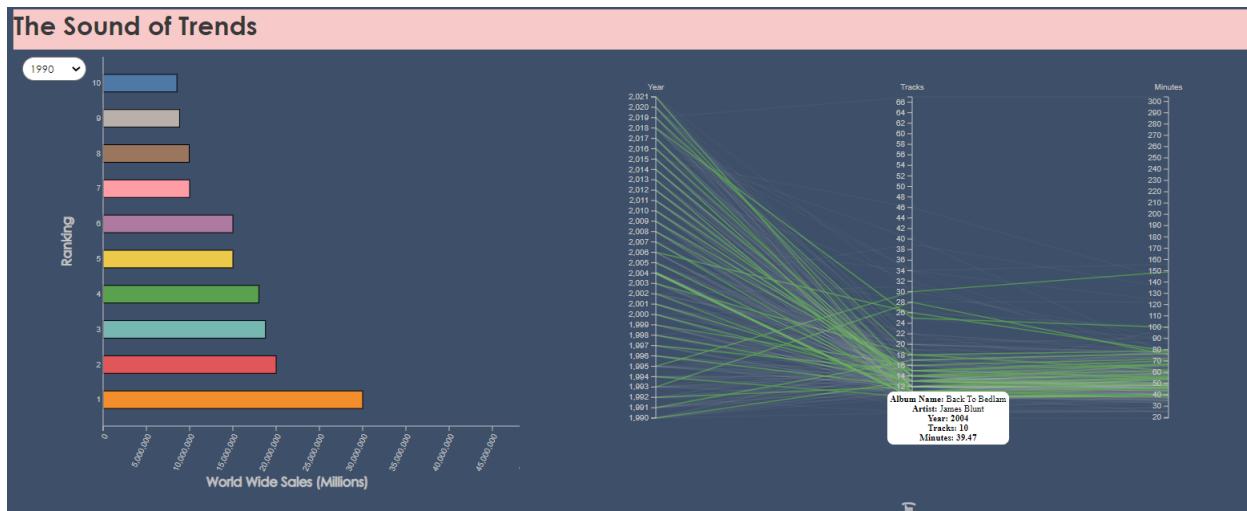


The ranking graph displays the top 10 albums per year designated by the year drop down menu. When a year is selected, the bars are updated to reflect the new data, which are the albums ordered by their rank and on the x axis is their world wide

sales in millions. When a particular bar in the graph is selected, it is highlighted, and the name of the artist, the album, and the specific number of world wide sales are shown. Also when a bar is selected, the album's tracks and length in minutes is highlighted on the parallel line chart, where you can see the selected content in comparison to the other data. Each ranking has its own color, and these colors coordinate to the parallel line graph to indicate the rank of each item.

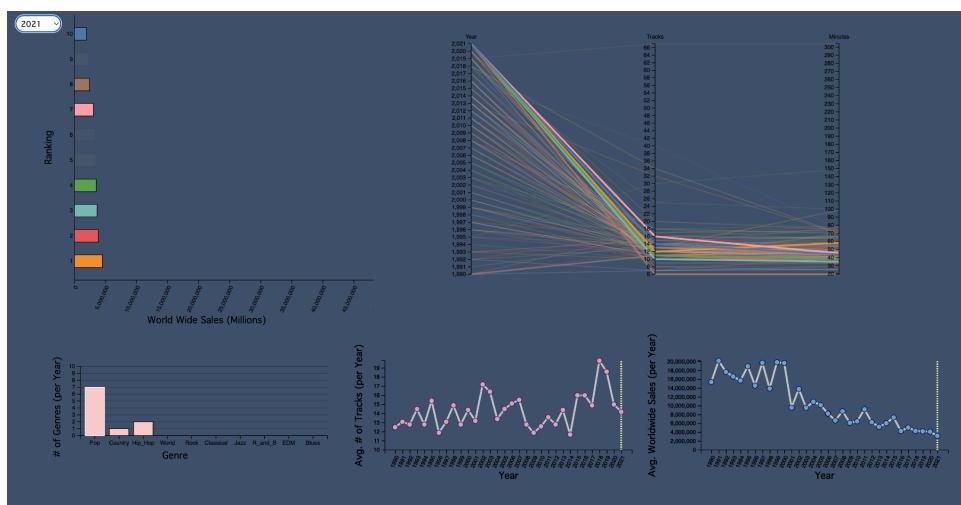
## Parallel Coordinate graph

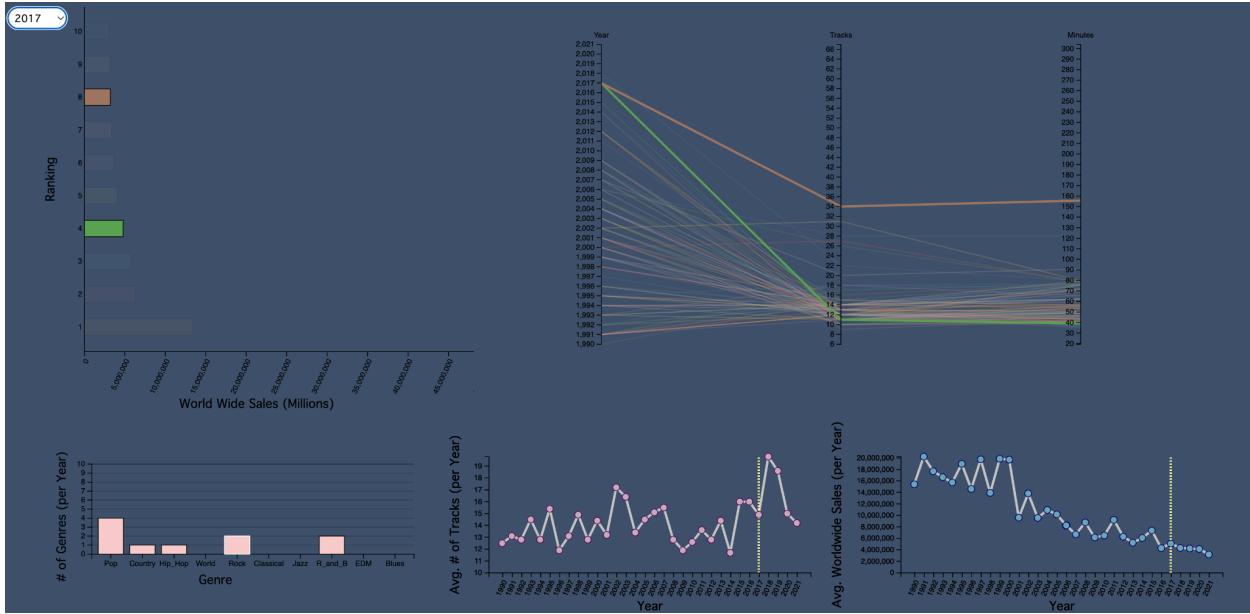




The parallel coordinate chart shows every item in the dataset and plots it with the year, number of tracks, and length in minutes. When selecting an element, you are able to see its color, which corresponds with its rank (shown on the rank horizontal bar chart), and it also highlights all of the other items with the same rank, so it is easy to see the comparison between these items in terms of their rank, year, number of tracks, and length. Also, when an item is scrolled over, it is enlarged so it is easier to see the particular item and the album name, Artist, Year, Tracks, and minutes are displayed.

## Genre Bar Chart





The genre bar chart has interactions with both the ranking horizontal bar chart and the parallel coordinate graph. When a genre is selected for the year, it is highlighted so it is easy to see the selection. On the ranking graph, all of the items with the genre attribute selected on the genre graph are shown, and the others without the specified genre fade, allowing the user to easily see which albums have that particular genre. On the parallel coordinate graph, only the items with the specified genre are shown, and the items for the current year from the drop down menu are bolded and slightly more opaque to show the data for the selected year and genre versus all of the other years.

---

## **Evaluation:**

From our visualizations we were able to gain more insight into our data on the top 10 ranking albums from 1990- 2021 and the trends in their genre, length, tracks and world wide sales. We learned valuable insight regarding these different trends, including that world wide sales have been at a decline these past 31 years, and heavily declined beginning in 2015. We also learned that the number of tracks per album seemed to be up and down through the 31 years, with the lowest average tracks in 2014 and the highest average tracks in 2019. Regarding genre, interestingly we were able to find a difference in album tracks and lengths for different genres, for example Hip Hop Albums seem to have a higher number of tracks, usually around 18 -20, where other popular genres such as pop and rock are around 10 -14 tracks.

We believe that our visualization works well, we think that it is engaging for the user and also answers the questions that we initially wanted to explore, as well as reveals other trends in the data.

Though we are happy with our visualization, we believe that there is always room for improvement. Having the date drop down made things a bit confusing to look at, so we would maybe look at turning it into a slider instead that the viewer can slide across to see the changing data in the graphs instead of having to click a different date to see the change and trends. Also, it would have been neat to explore with different types of visualizations for genre, to better show the 'parts of a whole', concerning the genres amount for the year versus in total for all 31 years.