

THE RISE OF HIP HOP

**A Visualization of the History & Evolution of Hip Hop Music
in the United States (1987-Present)**

University of Washington, HCDE 511, Spring 2023

Gabriella Cronk, Saumya Goyal, Amodini Khade, Tabitha Safari, Luka Sherman

TABLE OF CONTENTS

Table of Contents.....	2
Abstract.....	3
Introduction.....	3
Previous Work.....	4
Data Process.....	7
Step 1: Setup Spotify API App.....	7
Step 2: CSV of Every Billboard Hot 100 Song.....	7
Step 3: Create List of Hip Hop Artists.....	8
Step 4: Remove Billboard Hot 100 Songs by Non-Hip Hop Artists.....	8
Step 5: Use API Search to Find Track's Spotify ID.....	9
Step 6: Query Spotify API Data For Each Song and Add to New Columns.....	10
Sample Single Row of Final Data.....	11
Design Process.....	12
Initial Sketches.....	12
Design Iterations.....	15
Usability Study.....	23
Evaluation of Final Visualization.....	25
Overview and Detail Views.....	25
Interaction Techniques and System Features.....	26
Shneiderman's Information Visualization Tasks.....	27
Conclusion.....	28
Limitations of Tableau.....	28
Limitations of Data.....	28
Further Work.....	29
Acknowledgments.....	29
References.....	30
Appendix A: Python Script.....	32
Appendix B: Usability Questions.....	33
Appendix C: Final Visualization.....	34

ABSTRACT

Over the decades, Hip Hop culture and music have served as a powerful outlet for youth from marginalized backgrounds to express, share, and understand their voices, emotions, ideas, and experiences. This project aims to explore a customized dashboard that delves into the history and evolution of Hip Hop music in the United States, spanning from 1987 (the first day a hip hop song appeared on the Billboard Hot 100 chart) to present day. While time constraints and unforeseen limitations of our visualization software of choice prevented us from fully achieving our initial goal of visualizing the complete history of Hip Hop music, our interactive dashboard provides an informative overview of top-performing Hip Hop artists and songs during the Golden Age Era of Hip Hop music (1987-1997) according to one of the most ubiquitous and widely-utilized metrics of music popularity in the world—the Billboard weekly Hot 100 chart. Furthermore, our visualization allows users to delve deeper into the details of their favorite artists and songs, such as their longevity on the Billboard Hot 100 throughout the Golden Age Era and the various audio dimensions of their songs. By conducting this analysis, our intention is to illuminate the transformative power and resilience of Hip Hop as an art form deeply ingrained in the sociocultural fabric of the United States.

INTRODUCTION

“Hip Hop is a cultural form that attempts to negotiate the experiences of marginalization, brutality, truncated opportunity, and oppression within the cultural imperatives of African American and Caribbean history, identity, and community” (Rose, 1994).

Hip Hop music originated in New York City in the early 1970s and has since grown into one of the most popular and widely-consumed genres of music in the United States. Over the decades, Hip Hop music and culture have seen a massive evolution. Different styles, techniques, and subgenres have risen and fallen in popularity while the advent of new tools and technologies has brought forth new sounds, music distribution methods, and social platforms enabling artists to reach wider audiences. As a team of avid Hip Hop fans and followers, our primary goal was to explore the history and evolution of Hip Hop music in the United States from its origins in the late 1970s to today through data visualization.

PREVIOUS WORK

We referenced several related data visualizations throughout our project as sources of inspiration. A majority of these visualizations take the form of interactive timelines and temporal analyses that explore the evolution of specific genres, artists, or music landscapes over time.



Figure 1: Screenshots from Daniels, 2015. **Top Left:** Primary animated timeline view. **Top Right:** Filtered artist spotlight view. **Bottom:** Clicking on an artist image from the animated timeline view redirects you to the filtered artist spotlight view.

Our primary source of inspiration—**Figure 1** (Daniels, 2015)—plots the top ten Hip Hop tracks on Billboard’s weekly Hot Rap Chart from 1989-2015, taking users on an immersive auditory and visual journey exploring Hip Hop’s highest ranking tracks and artists. From the primary animated timeline view, users can click on the artist image associated with a given track to bring up a static visualization that displays a filtered selection of tracks by that artist that made it onto Billboard’s weekly Hot Rap Chart between 1989-2015. This action also switches the audio from playing the number one track (which is played by default on the primary animated timeline view) to a sample of the track selected on the filtered artist spotlight view. The soundtrack will then cycle through the filtered track list on the left starting with the song you clicked. Users can click on any of the songs on the left of the filtered visualization view to hear a sample from that song. Clicking anywhere on the static visualization takes you back to the animated timeline.

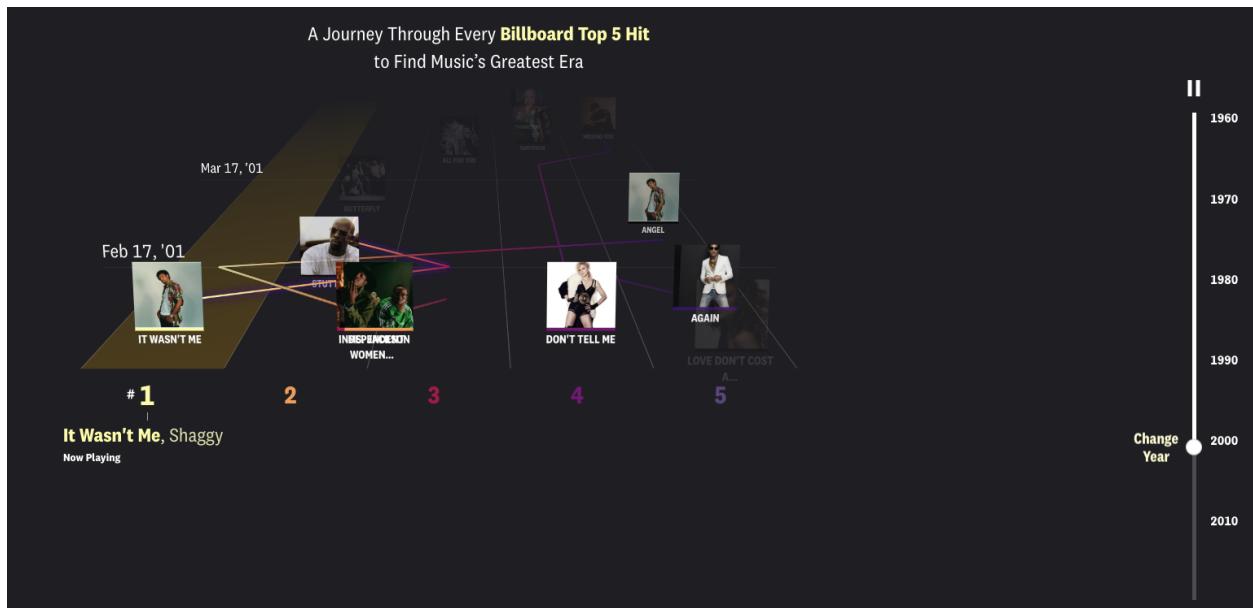


Figure 2: Screenshot from *A Journey through Every Billboard Top 5 Hit* (Best Year in Music, n.d.).

Figure 2 (Best Year in Music, 2019) uses a similar interactive animated timeline format featuring audio samples from every top five hit from Billboard’s weekly Hot 100 chart from 1958-2019 to uncover music’s greatest era. Furthermore, The Good, the Rad, and the Gnarly (Wilber, 2018) examines how the music of skateboarding changed over time, how different genres relate to particular styles of skating, which songs are associated with specific skaters, and popular artists within each genre that most commonly appear in skate videos.

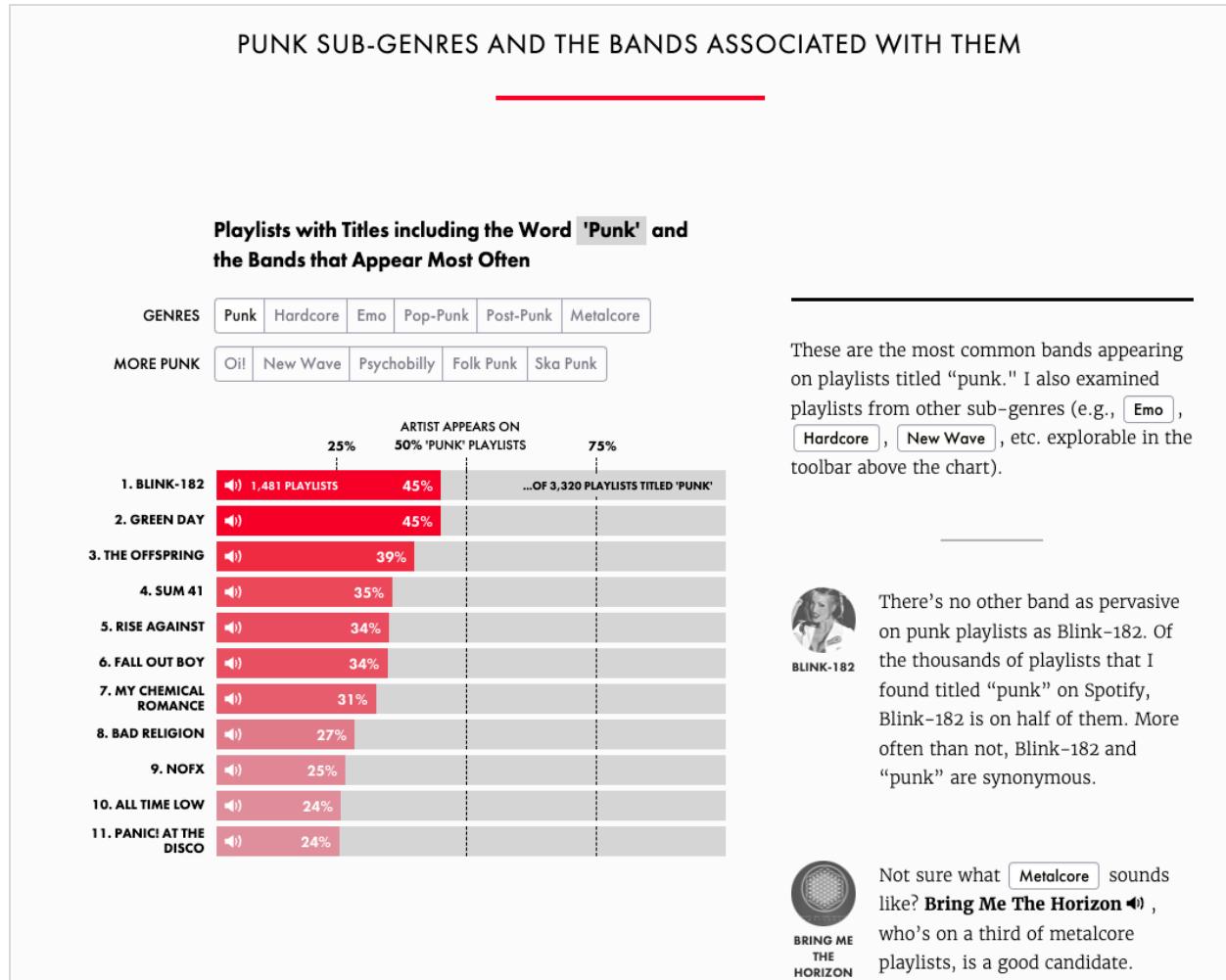


Figure 3: Screenshot from *Crowdsourcing the Definition of "Punk"* (Daniels, Crowdsourcing).

Another subset of visualizations our team took inspiration from includes deep dives into specific artists and subgenres. For instance, *Crowdsourcing the Definition of "Punk"* in **Figure 3** delves into the influence of Blink-182 on punk culture by exploring Spotify and YouTube playlists with titles including the word 'Punk' and the bands that appear most frequently within them. Additionally, *Charting The Beatles - An Infographic Atlas Of Beatles Music* (Deal, n.d.) conducts extensive exploratory analyses into the music of the Beatles and includes several data visualizations on song structure, song rhythm and tone, tour dates and locations, and song authorship credits.

We enjoyed how fun, visual, auditory, and interactive these visualizations were to use, and the ability to instantly customize them to your preferences by filtering or searching for your favorite artists. We hoped to make something informative and engaging similar to these works, but viewed through the lens of mainstream popularity and with the addition of Spotify's advanced API data.

DATA PROCESS

We created a custom dataset by writing a Python script ([Appendix A](#)) with the pandas library to combine multiple data sources, in addition to employing manual cleaning methods.

Step 1: Setup Spotify API App

- A. Created a new Spotify API App from the developer dashboard, using “<https://localhost>” as the url, and obtained client id and secret codes (Spotify, n.d.).
- B. Set local environment variables with the app codes, redacted here for security.

```
export SPOTIPY_CLIENT_ID='your-spotify-client-id'
export SPOTIPY_CLIENT_SECRET='your-spotify-client-secret'
export SPOTIPY_REDIRECT_URI='http://localhost'
```

- C. Check that local variables were set correctly.

```
echo $SPOTIPY_CLIENT_ID='your-spotify-client-id'
echo $SPOTIPY_CLIENT_SECRET='your-spotify-client-secret'
echo $SPOTIPY_REDIRECT_URI='http://localhost'
```

- D. Install Spotipy, a Python library to access the Spotify API (Lamere, 2023).

```
pip install spotipy --upgrade
```

Step 2: CSV of Every Billboard Hot 100 Song

We started with an existing CSV file of every Billboard Hot 100 song from when the chart started in 1958 through the beginning of 2023 (Miller, 2023). This list contained songs from all genres, to show us the percentage of Hot 100 songs that were Hip Hop, and demonstrate mainstream popularity from an industry-established source. Billboard’s Hip Hop and rap charts didn’t start until the genre was more mainstream, and we were specifically interested in the mainstream rise of Hip Hop.

Rows: 336295, one for each Billboard chart instance

Columns: 13 total, 8 of them used

chart_position	chart_date	song	performer	time_on_chart	consecutive_weeks	peak_position	chart_debut
48	7/17/93	It Was A Good Day	Ice Cube	20	19	15	3/6/93

Step 3: Create List of Hip Hop Artists

We created a one-column csv file of Hip Hop artists from copy/pasting the text on two Wikipedia pages (*List of Hip Hop Groups*, 2023) (*List of Hip Hop Musicians*, 2023).

- A.** Copy and “Paste Without Formatting” both files one after the other into a blank Excel sheet.
- B.** Remove letter headings (e.g. "A[edit]", "B[edit]", etc) and Wikipedia page indicators (e.g. "Common[1]" -> "Common") through “Replace All” with regular expressions.
- C.** Manually searched for "(" to remove duplicates when artists have multiple language pages (e.g. "Shorty (American)", "Shorty (German)" -> "Shorty").
- D.** Removed all empty rows through “Find & Select” -> “Special” -> “Blanks” -> “Delete..” -> “Shift cells up.”
- E.** Cleaned each row in Python through converting to unicode, lowercase, spelling out symbols.

Rows: 4413, one for each Hip Hop artist

Columns: 1

hip hop artist
iayze
icecube
...

Step 4: Remove Billboard Hot 100 Songs by Non-Hip Hop Artists

- A.** Used a Python script to iterate over each row: If no artist from Billboard’s “performer” category (split by “featuring,” if present) is in the list of all Hip Hop artists, the row is removed. We also made a secondary dataset at this time, with only appending a column “is_hiphop” with a TRUE/FALSE value, to visualize the percentage of Hip Hop Hot 100 songs over time.
- B.** When filtering from the earliest year, we observed some non-Hip Hop songs from the 50s-80s present in the list. We determined this to be caused by non-Hip Hop artists having the same common name as artists on the Hip Hop Wikipedia pages. When manually looking up these early songs, if the artist wasn’t the same as the one indicated on the Wikipedia page (almost always apparent because the song was published before their “years active” range started), the row was removed. Artists removed this way were:

Cream, Tommy Cash, Dawn, Free, Exile, Angel, Circus, Focus, Cochise, Sweet Sensation, Saga, King, The Firm, Dragon, Kano, Georgio, Wax

- C.** The earliest several dozen songs were by James Brown and removed. Although he was the same artist linked on the Wikipedia page as an early influence to the genre, his Hot 100 songs were funk, not Hip Hop. Manually cleaning up these early songs was important to accurately highlight when Hip Hop was first on the charts and was very visible in our data. Although there were other artists listed who we might *not* consider Hip Hop (e.g. Mariah Carey), we left them to stay consistent to the crowdsourced Wikipedia definition of Hip Hop.

Rows: 42653, one for each Hip Hop Billboard chart instance

Columns: 13 total, same as Step 2

Step 5: Use API Search to Find Track's Spotify ID

Used Spotify API search() call to find the track in Spotify to get its Spotify ID to later access more specific Spotify datasets.

- A.** First, we searched automatically for the track, putting the “song” and “performer” through a string-sanitizing function to eliminate differences in punctuation and spelling (e.g. “1” versus “One”). Iterating through the list of potential tracks, if the sanitized artist and performer fields of the Billboard song were an exact match to that of the Spotify song, it was determined a match and the additional fields of the Spotify track were filled in.

```
d = (spotipy.Spotify()).search(q=song+"artist:"+ (performer) ,type="track")
```

- B.** If there was no exact match in the search results, the script printed the entire search results and prompted the user to manually select if any of them were a match. If none of them were, the user could manually search for the song on Spotify and enter the url, and the script would load the information for that track. This was common when there were differences in how curse words were censored or unexpected spelling and punctuation differences. This was also necessary when Spotify’s search blacklisted a song, for instance all of R. Kelly’s songs would not return via search even if on Spotify, as with some songs with derogatory themes.

- C.** Manually removed all songs from five artists not found on Spotify:

Cash Money Millionaires, Divine, Funkmaster Flex, Jesse Jaymes, Purple Ribbon All-Stars

- D.** Manually removed twenty six songs not found on Spotify:

Ching-A-Ling by Missy Elliott, Gotham City by R. Kelly, Hate Sleeping Alone by Drake, Hell Of A Life by T.I., Hoop In Yo Face (From "Sunset Park") by 69 Boyz Featuring Quad City DJ's, I'll Never Stop Loving You by J'Son, I'm Back by T.I., Just Another Girl by Monica, Just The Way You Like It by The S.O.S. Band, Last Chance by Ginuwine, Like That by Kris Wu, Redemption Song by Rihanna, Roberta by Bones, Shake Your Pom Pom by Missy Elliott, Soldier's Heart by R. Kelly, Someday At Christmas by Lizzo, Tell Me If You Still Care by The S.O.S. Band, The Finest by The

S.O.S. Band, The Heart Part 4 by Kendrick Lamar, The World's Greatest by R. Kelly, Warrior by Nelly, Way Too Cold by Kanye West Featuring DJ Khaled, We're Not Making Love No More by Dru Hill, Who Dat by J. Cole, Your Body's Callin' by R. Kelly, ZaZa by 6ix9ine

Rows: 41772, one for each Hip Hop Billboard chart instance with a Spotify record

Columns: 22, below are select columns added

ID	song	track_id	track_name	popularity	embed	duration_ms	explicit	number
18835	It Was A Good Day	2q0m7ukLyHUXWyr4ZWlx	It Was A Good Day	84	https://open.spotify.com/embed/track/2q0m7ukLyHUXWyr4ZWlx	260000	TRUE	7

Step 6: Query Spotify API Data For Each Song and Add to New Columns

- A.** For every artist in the track's "artist" key list, use the artist API call to get more information.

This shouldn't have been necessary as according to the documentation these fields should have been included in the original track() call, but the values always returned empty. We were initially cautious as to not exceed our maximum API calls with such a large dataset, but after we had enough initial data we did further tests and found that the artist() calls did have these fields populated. We did get temporarily blacklisted for exceeding our call quota but were eventually able to run these additional calls for all rows. One way we minimized calls was first creating a dictionary of every artist in the dataset, and then only calling once for each artist.

```
artist_dict = spotipy.Spotify().artist(artist_id)
```

- B.** For every artist in the track's "artist" key list, use the artist API call to get more information.

These fields were also present and should have been populated in the original track() call according to the API documentation but were only found to be populated in the album() call.

```
album_dict = spotipy.Spotify().album(spotify_album_id)
```

- C.** Get calculated audio features for each track.

```
track_audio_features_dict = spotipy.Spotify().audio_features(track_id)
```

- D.** Removed columns that were universally unpopulated (album_genres, music industry reference codes).

Rows: 41772, same as Step 5

Columns: 88 total, including 7 fields each for up to 6 artists, not commonly populated

Sample Single Row of Final Data

Column names abbreviated for clarity, some less significant columns omitted here for simplicity.

[FROM STEP 2]

ID	chart_position	chart_date	song	performer	time_on_char_t	consecutive_weeks	peak_position	chart_debut
18835	48	7/17/93	It Was A Good Day	Ice Cube	20	19	15	3/6/93 →

[FROM STEP 5]

track_id	track_name	popularity	embed	duration_ms	explicit	number
2qOm7ukLyH UXWyR4ZWlxwA	It Was A Good Day	84	https://open.spotify.com/embed/track/2qOm7ukLyHUXWyR4ZWlxwA	260000	TRUE	7 →

[FROM STEP 6A, repeated for every credited artist]

artist_id	name	popularity	followers	genres	image_url	url
3Mcii5XWf 6E0lrY3Uk y4cA	Ice Cube	76	6909895	['conscious hip hop', 'g funk', 'gangster rap', 'hip hop', 'rap', 'west coast rap']	https://i.scdn.co/image/ab6761610000e5eb650e79afaf00492e33d7b8c	https://i.scdn.co/image/ab6761610000e5eb650e79afaf00492e33d7b8c →

[FROM STEP 6B]

album_id	album_name	release_date	total_tracks	type	label	popularity	url	image
71HM1CMYWeZ zws8pyiEn46	The Predator	11/17/92	16	album	Priority Records	77	https://open.spotify.com/album/71HM1CMYWeZws8pyiEn46	https://i.scdn.co/image/ab67616d0000b273994c319841a923465d62e271 →

[FROM STEP 6C]

acousticness	danceability	energy	instrumentalness	key	liveness	loudness	mode	speechiness	time_signature	valence
0.33	0.798	0.744	0.000106	7	0.292	-5.328	0	0.136	82.3564	0.794

DESIGN PROCESS

Initial Sketches

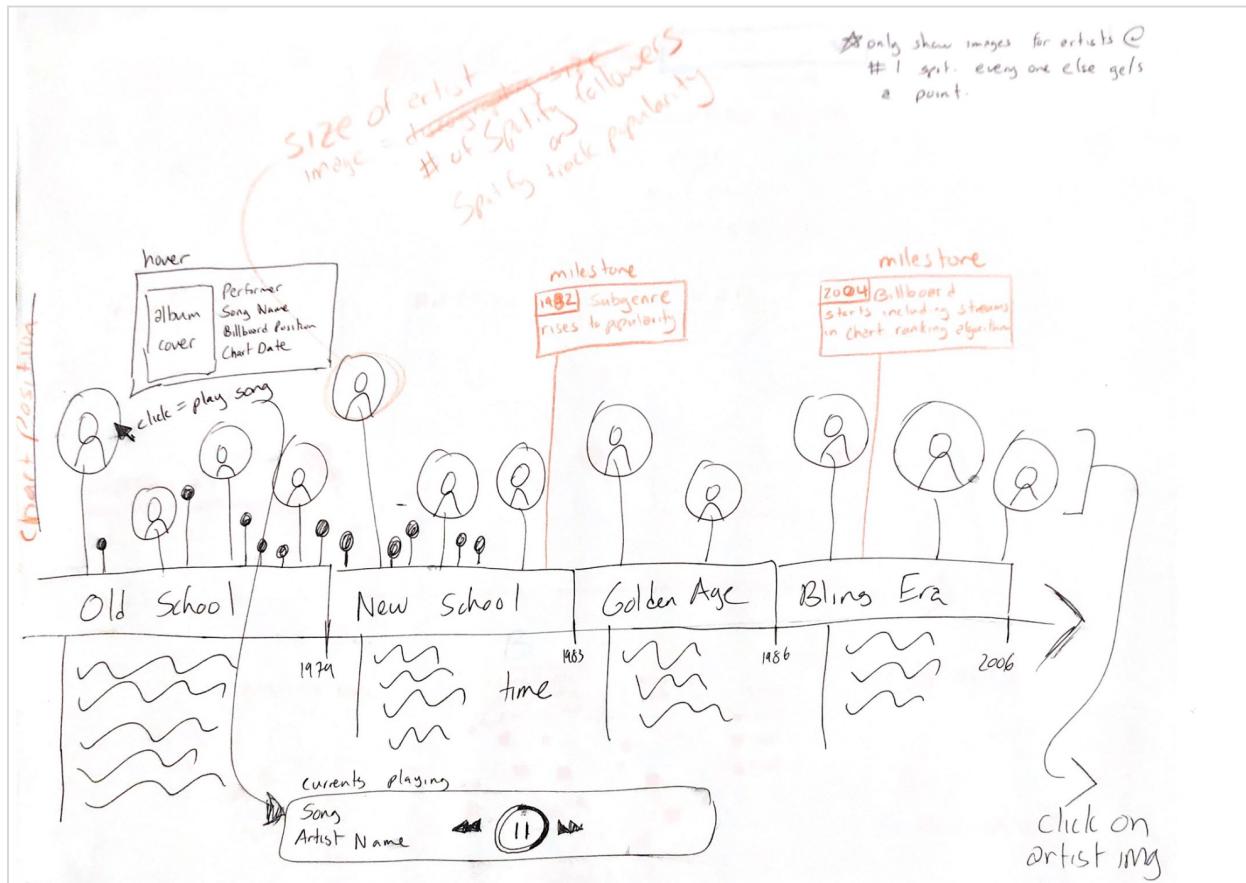


Figure 4: Sketch of Interactive Animated Timeline, the main visualization

Initially, our team attempted to create an interactive timeline depicting Hip Hop songs that charted in the top five positions of the Billboard Hot 100, spanning from the first instance of a Hip Hop song making it onto the chart to present day (**Figure 4**). Through this timeline, we also wanted to pinpoint important social, cultural, and technological milestones that led to widespread shifts in the culture, sounds, and aesthetic associated with Hip Hop music. Additionally, we envisioned embedding a Spotify music player within the visualization to allow users to play audio samples of any song. Any track that reached #1 on the Hot 100 would have been represented by a circular image of the artist associated with the track. Tracks that reached positions two through five would have been represented by a colored data point. Clicking on an artist name would take users to a visualization of a deep dive into the charted history of that artist (**Figure 5**).

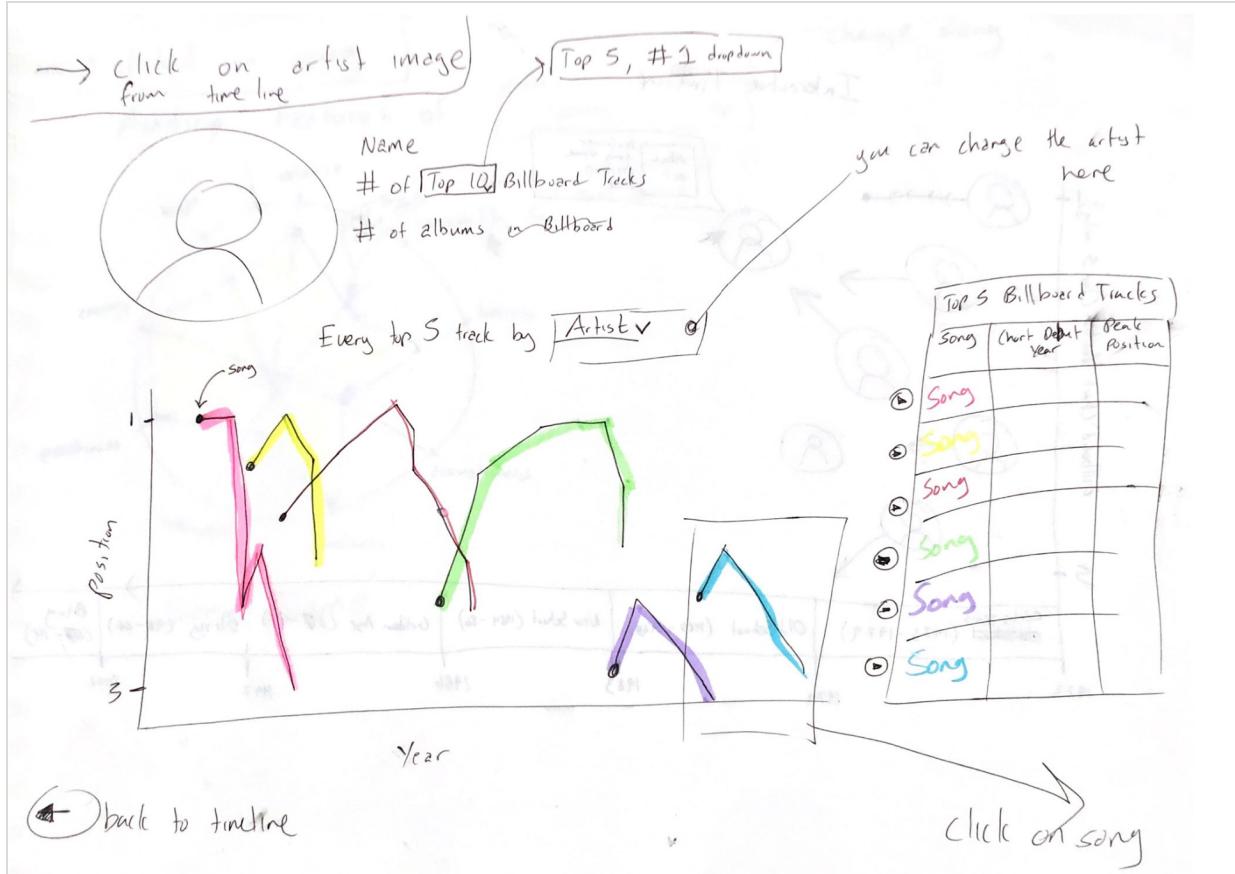


Figure 5: Sketch of Artist Deep Dive visualization. Songs by a selected artist that made it into the Billboard weekly Hot 100 are represented by different color lines. Each point along a single line represents the song's chart position (y-axis) on a given chart date (x-axis).

Our Artist Deep Dive visualization in **Figure 5** would feature a prominent image of the artist selected from the main interactive timeline as well as statistics and details about the artist, including but not limited to the total number of songs released by them that made it onto the Billboard Hot 100, the aggregate number of weeks the artist has been featured on the chart over the course of their career, and the the total number of Spotify followers they had as of April 2023 (when we last called the API data). Songs by the selected artist that made it into the Billboard Hot 100 are represented by differently-colored lines. Each point along a line represents the song's chart position (y-axis) at a particular point in time (x-axis). From here, we envisioned giving our users the ability to click on a song to be taken to a third visualization—a radial chart depicting the various audio dimensions of the song in **Figure 6**.

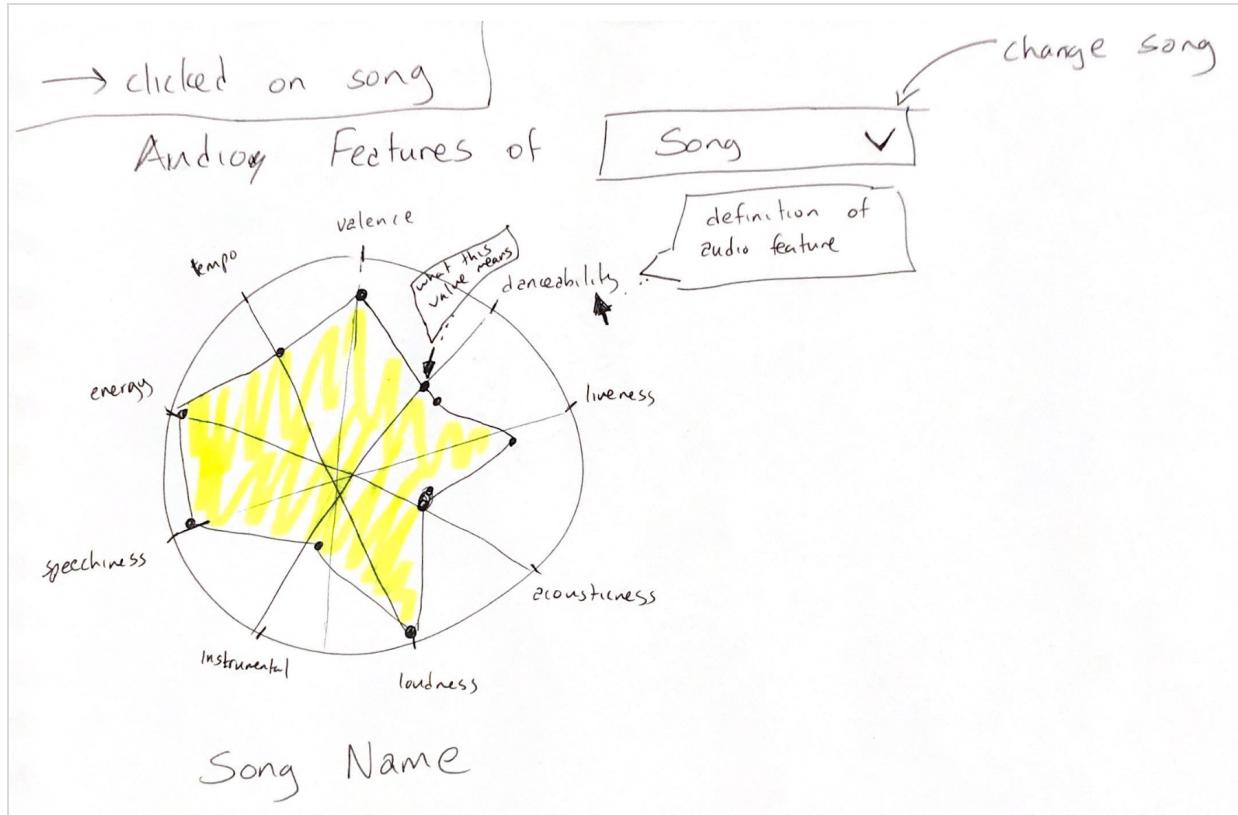


Figure 6: Sketch of radial Chart illustrating Spotify audio dimensions of the song selected from Figure 4.

On the radial chart visualization in **Figure 6**, users see the Spotify audio dimensions for the song they selected from **Figure 5**. Hovering over a dimension label would show the definition and examples of what a low and high valued-song sounds like.

Unfortunately, the system of visualizations we initially set out to achieve proved to be too complex for our team to achieve through Tableau. After consulting with a Tableau expert, we learned that many of the interactions and user flows we had envisioned—such as clicking on an artist image from the timeline in **Figure 4** to open up the Artist Deep Visualization in **Figure 5**—was impossible to achieve through Tableau. We also learned that radial charts are not a default chart type in Tableau and that if we wanted to include this type of chart in our visualization, we would essentially have to manually create them one by one. Fortunately, we had sufficient data to pivot our plan to an achievable solution.

Design Iterations

Before reaching the final visualizations, we iterated extensively on the layout, structure and type of visualization to use for our research questions.

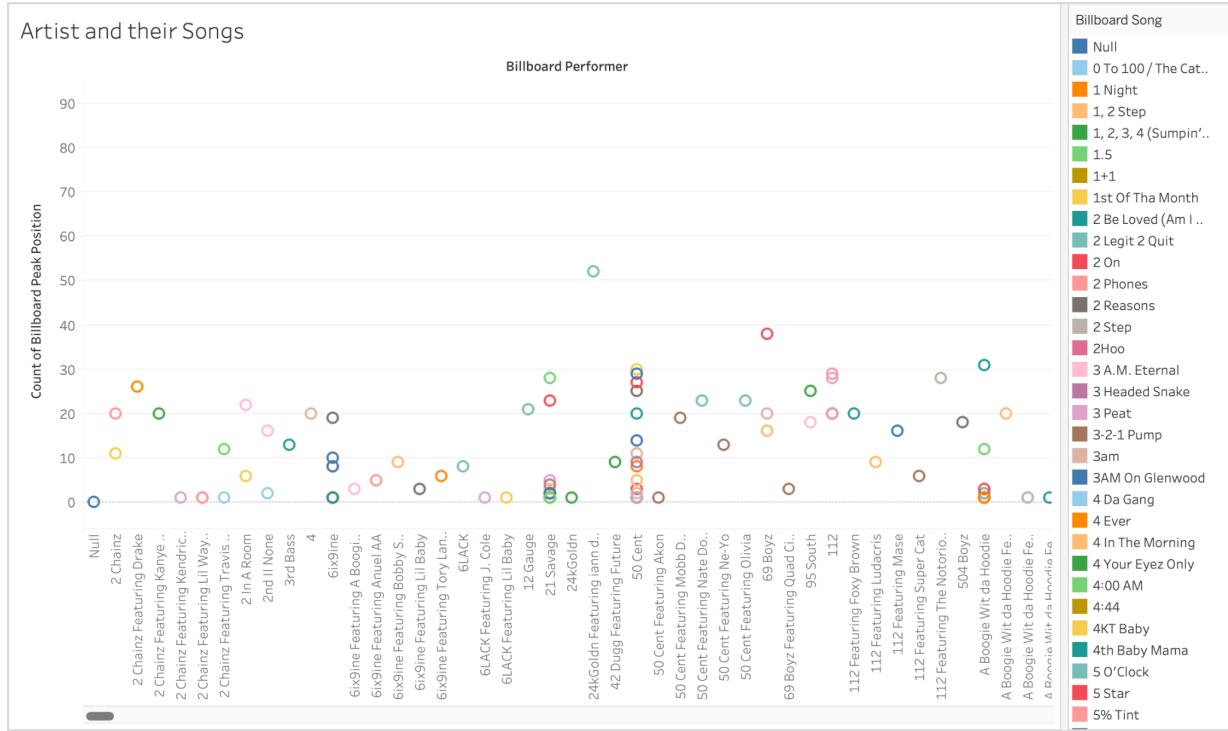


Figure 7: Artist (x-axis) and the peak Billboard Hot 100 position (y-axis) of their song (colored circles).

We began with a scatter plot (**Figure 7**) of the Artists and their Songs but this was vague to understand at a glance. The x-axis depicts the individual artist names and the y-axis depicts the Billboard Peak Position of the corresponding song. The filters did not make it easy to find a custom artist. The scatter plot was visually cluttered as there was too much going on between several data points. Additionally, the visualization required significant horizontal scrolling, inconveniencing users. Users didn't want to scroll alphabetically through such a large dataset.

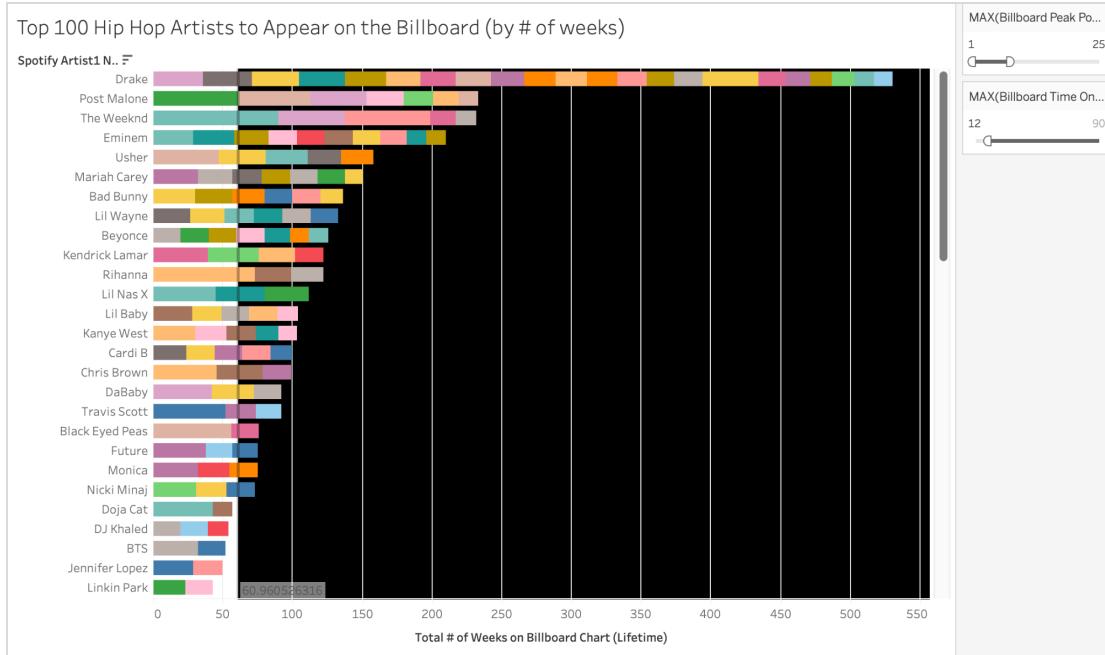


Figure 8: Top 100 Hip Hop Artists to Appear on the Billboard Hot 100 chart by Number of Weeks.

Figure 8 visualizes every artist to have appeared on the Billboard Hot 100 chart by the number of weeks their songs appeared on it. Each bar along the x-axis represents the length of their duration for which their song(s) occupied a position on the chart. This type of visualization required significant vertical scrolling so we decided to break our timeline into various recognized eras (*Hip Hop Music*, 2019).

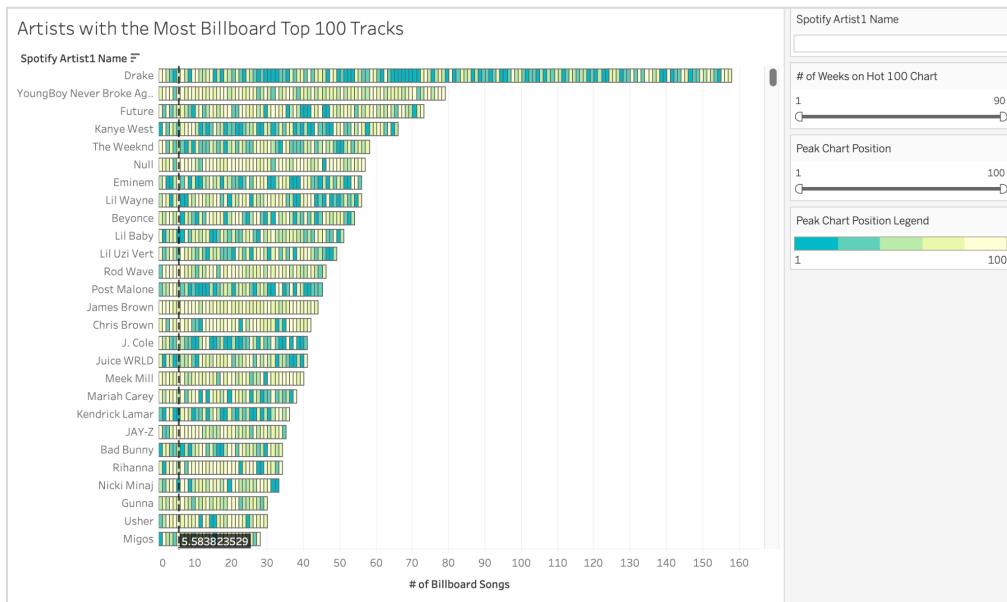


Figure 9: Artists with the Most Billboard Top 100 Tracks

In addition to exploring the time spent by each artist on the Billboard in terms of the number of weeks, we added indicators for how many songs they had on the chart (**Figure 9**). The y-axis represents the name of individual artists while the x-axis represents the number of songs by each corresponding artist. Users can view an artist of their choice with the search field to the right. They can also use the sliders to filter the artists by the number of weeks and the peak chart position. The legend on the right provides a color-coded representation of the ranking, with darker colors being on the top and lighter colors on the bottom.

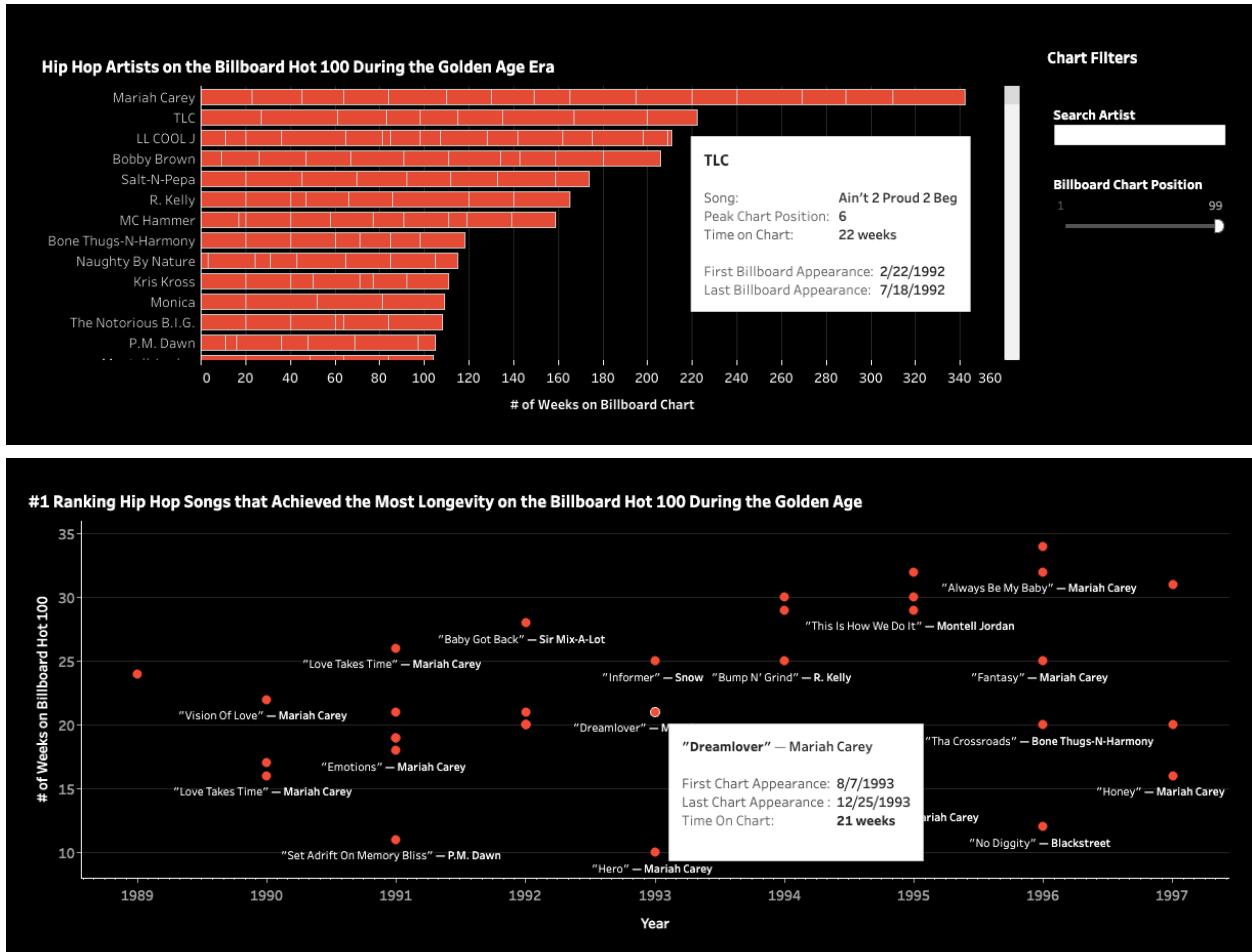


Figure 10: Bar chart (top): Hip Hop Artists on the Billboard Hot 100 During the Golden Age Era. Scatter plot (bottom): #1 Ranking Hip Hop Songs that Achieved the Most Longevity on the Billboard Hot 100 During the Golden Age

The bar chart (top, **Figure 10**) above displays every Hip Hop artist who made it onto the Billboard Hot 100 during the Golden Age Era of Hip Hop music (1987-1997) listed in order of most to least weeks spent on the chart. Each bar corresponding to an artist is broken into smaller segments representing each track and the weeks it spent on the chart. Users can search for an artist name to the right and see if their favorite artist made it onto the Billboard Hot 100 during the Golden Age. The Billboard Chart

Position slider allows users to filter the chart view to display only songs that made it to certain positions on the Billboard Hot 100 (e.g. Top 10, Top 5).

The scatter plot (bottom, **Figure 10**) serves a similar purpose to the bar chart by allowing users to explore the best performing Hip Hop songs of the Golden Age Era (1987-1997). The primary difference between these visualizations is that the bar chart focuses more on popular artists while the scatter plot focuses more on popular tracks by displaying Hip Hop songs that charted at number one and achieved the most longevity on the Billboard Hot 100 during the Golden Age. The position of each data point corresponds to the date when the track first debuted on the Billboard Hot 100 (x-axis) and the total number of weeks that the song spent on the chart (y-axis).

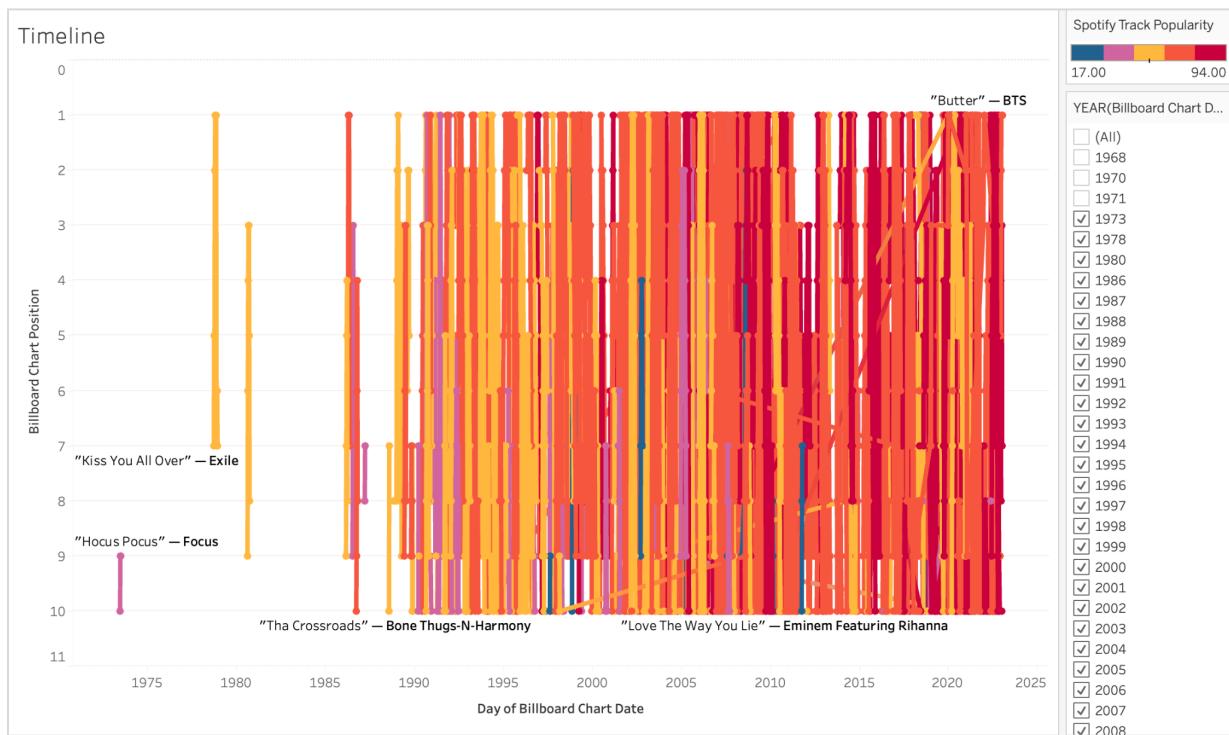


Figure 11: Timeline showing the changing chart positions of each song, colored by current Spotify popularity..

The line chart of **Figure 11** shows songs and their chart positions over a customized timespan, in relation to their current Spotify popularity. We concluded that the timeline would be too visually cluttered on Tableau unless we enabled horizontal scrolling, which would have been too crowded to navigate. We decided to pivot away from the timeline visualization and focus on breaking down different Hip Hop eras instead.

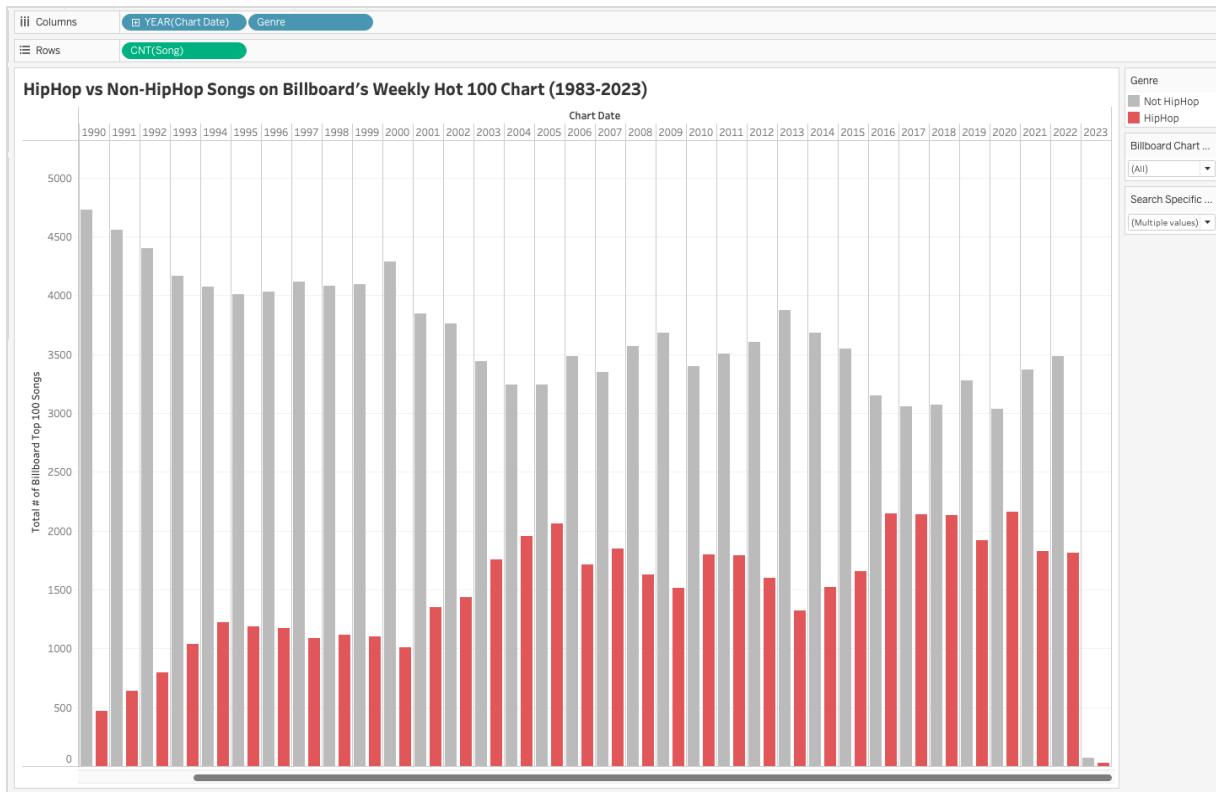


Figure 12: Hip Hop vs Non- Hip Hop Songs on the Billboard's Weekly Hot 100 Chart from 1983 to 2023.

The bar graph in **Figure 12** featuring side-by-side bars depicts the Year of Billboard Chart Date on the x-axis, the Billboard Hot 100 Song Count on the y-axis and the Song Genre (Hip Hop or Not Hip Hop) by bar color with red representing a Hip Hop song. This visualization allowed us to see how the number of Hip Hop songs in the Billboard Hot 100 has been increasing over the years, with some fluctuations. However, we were unable to fit all the data into a single view since it required horizontal scrolling. Furthermore, the chart was visually cluttered and it was difficult to keep track of which bars correspond to each year.

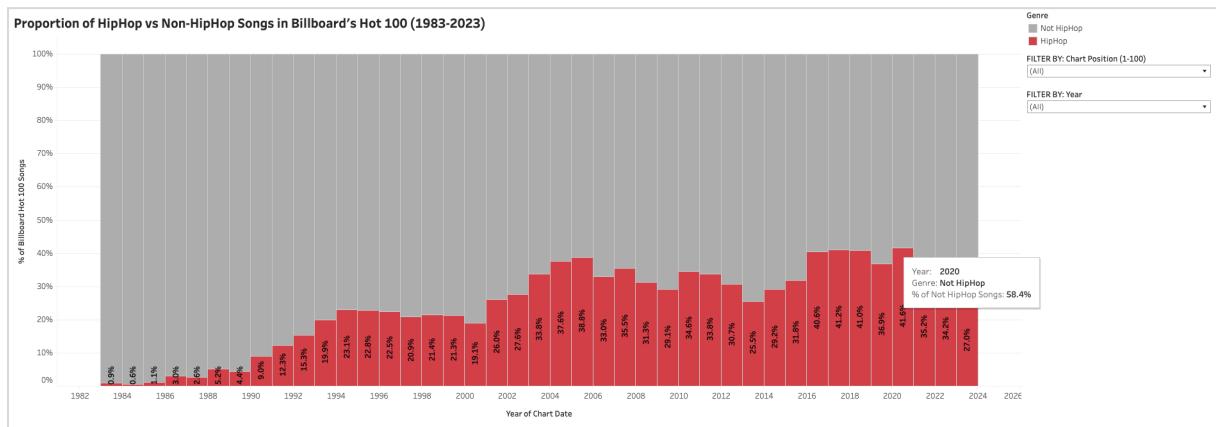


Figure 13: Hip Hop vs Non- Hip Hop Songs on the Billboard's Weekly Hot 100 Chart from 1983 to 2023.

We modified **Figure 12** by converting it to a stacked bar graph in **Figure 13**. Users can filter the Chart Position on the right to include only songs that are ranked at certain positions, and can filter by year.

Converting the year axis from a discrete to continuous field allowed us to easily fit the data into a single view without visual clutter. Calculating the genre distribution for each year as percentages of the total song count allowed us to standardize measures using a scale of zero to 100 so Hip Hop vs. non-Hip Hop data trends could be quickly, easily, and more accurately compared. Using percentages instead of total number of song counts also allowed us to fairly add our limited 2023 data back into the visualization.

However, we were not able to specify the position of the data labels added to the Hip Hop bars in Tableau, which added some minor visual clutter. Because there were very few Hip Hop songs in the Billboard Hot 100 in the first seven years of the timeline, the data labels corresponding to those years were not able to fit into their respective bars. We debated removing the data labels from those specific years, but felt it was more important for users to readily observe the percentage of Hip Hop songs across all years rather than forcing them to hover over those bars to access the data.

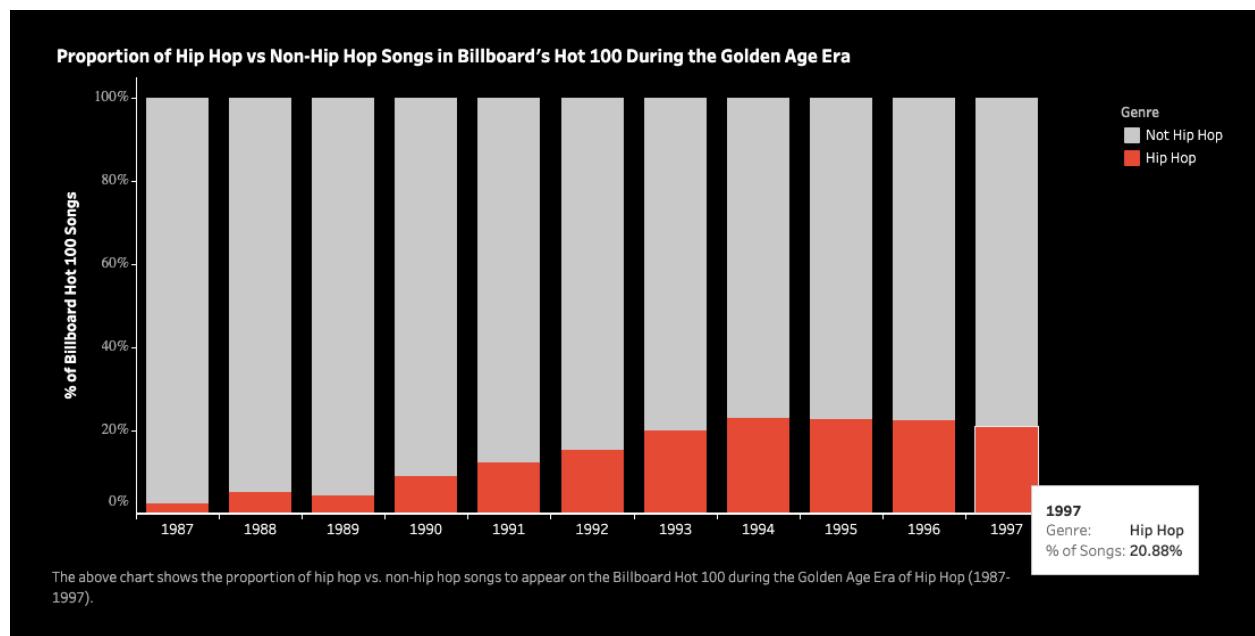


Figure 14: Bar Graph showing Proportion of Hip Hop vs Non- Hip Hop Songs on the Billboard's Weekly Hot 100 During the Golden Age Era

The bar graph in **Figure 14** is the section of **Figure 13** that corresponds to the Golden Age Era of Hip Hop. The breakdown of eras provided us with the ease of representing data in a more concise manner. The users could grasp the information at a glance, and did not need to scroll horizontally, thanks to the chunking of the large timeline into smaller, easily consumable sections.

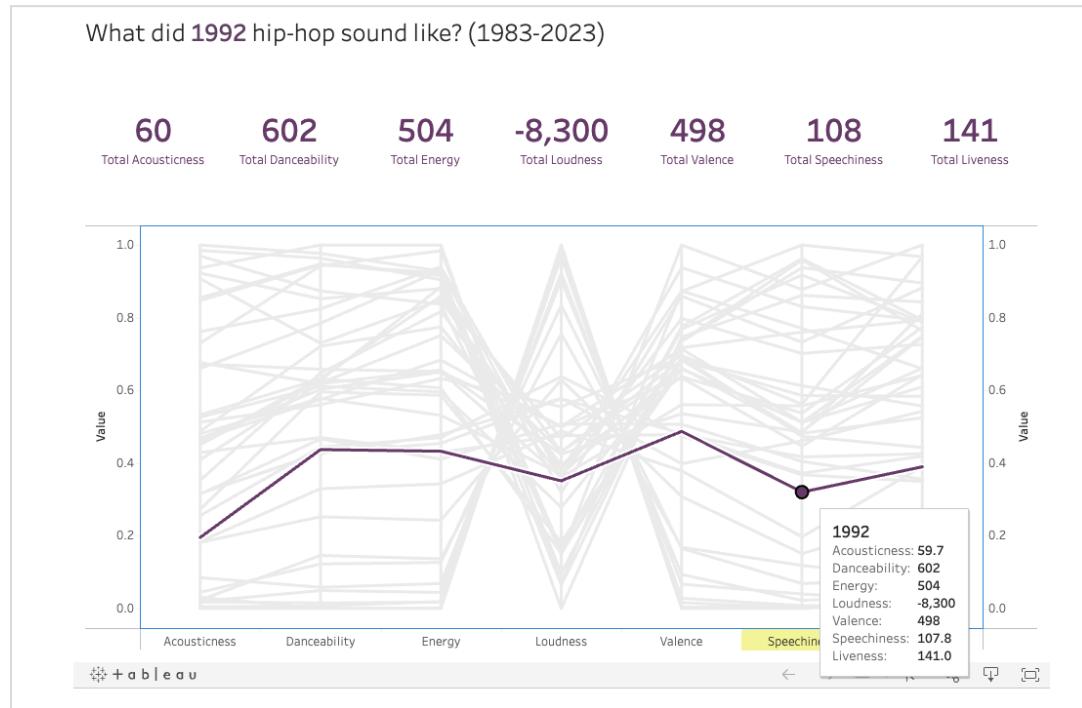


Figure 15: Line Chart of average Audio Features of the 1992 Billboard Hot 100 Hip Hop Songs

The line chart in **Figure 15** represents the audio dimensions of the 1992 Hip Hop songs. These dimensions include the collective acousticness, danceability, energy, loudness, valence, speechieness and liveness of the songs of the era on the x-axis, and the y-axis denotes their values. While this graph tries to depict all these parameters at once, the consumability of this information is difficult due to the overwhelming nature of the graph and unfamiliar metrics.

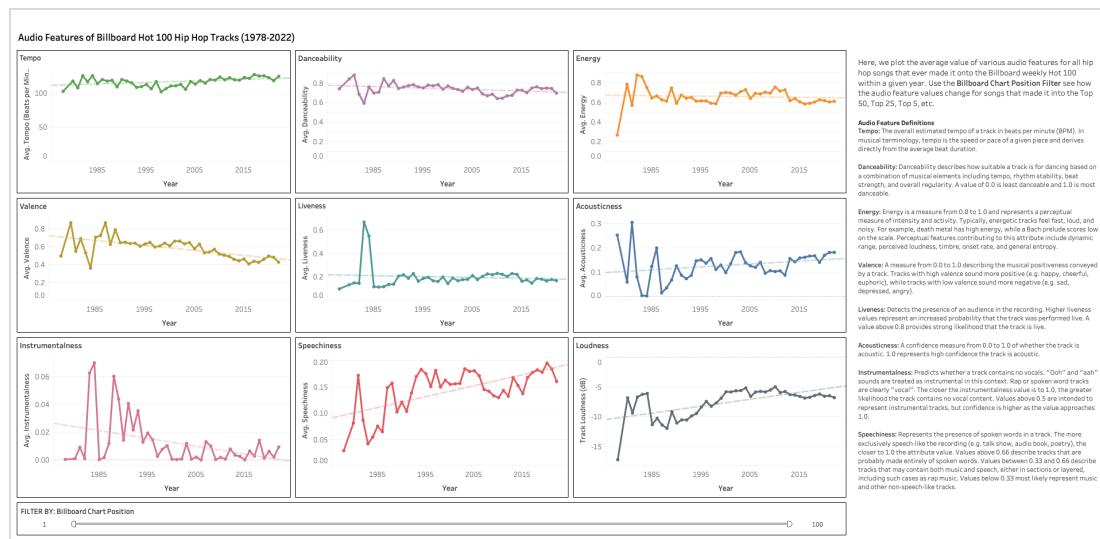


Figure 16: Individual Line Charts showing Audio Dimensions of Billboard Hot 100 Hip Hop Tracks (1978-2022)

For our second iteration of the audio dimensions visualization, we opted for a simpler line graph and plotted each audio dimension within their own individual graphs, organized in a matrix view in **Figure 16**. In each graph, we plotted the average value of the corresponding audio features for all Hip Hop songs that ever made it onto the Billboard weekly Hot 100 within a given year. The x-axis, for all the graphs, represents the year. The y-axis represents the value of the chart's specific audio feature. The Billboard Chart Position slider below the matrix acts as a filtering mechanism that allows users to see how patterns changed among a specific subset of the Top 100 songs (e.g. Top 5, Top 10).

In this version, we added definitions to each audio feature in the captions and made each definition available in the tooltip pop-up when users hovered over a data point. We also addressed confusion caused by the original aggregate view by plotting each audio feature on the y-axis with years on the x-axis within their own individual charts, which allowed users to clearly see trends in patterns in how the audio qualities of Hip Hop songs have changed over time. Furthermore, we added a filter so users could see how patterns changed among a specific subset of the Top 100 songs (e.g. Top 5, Top 10).

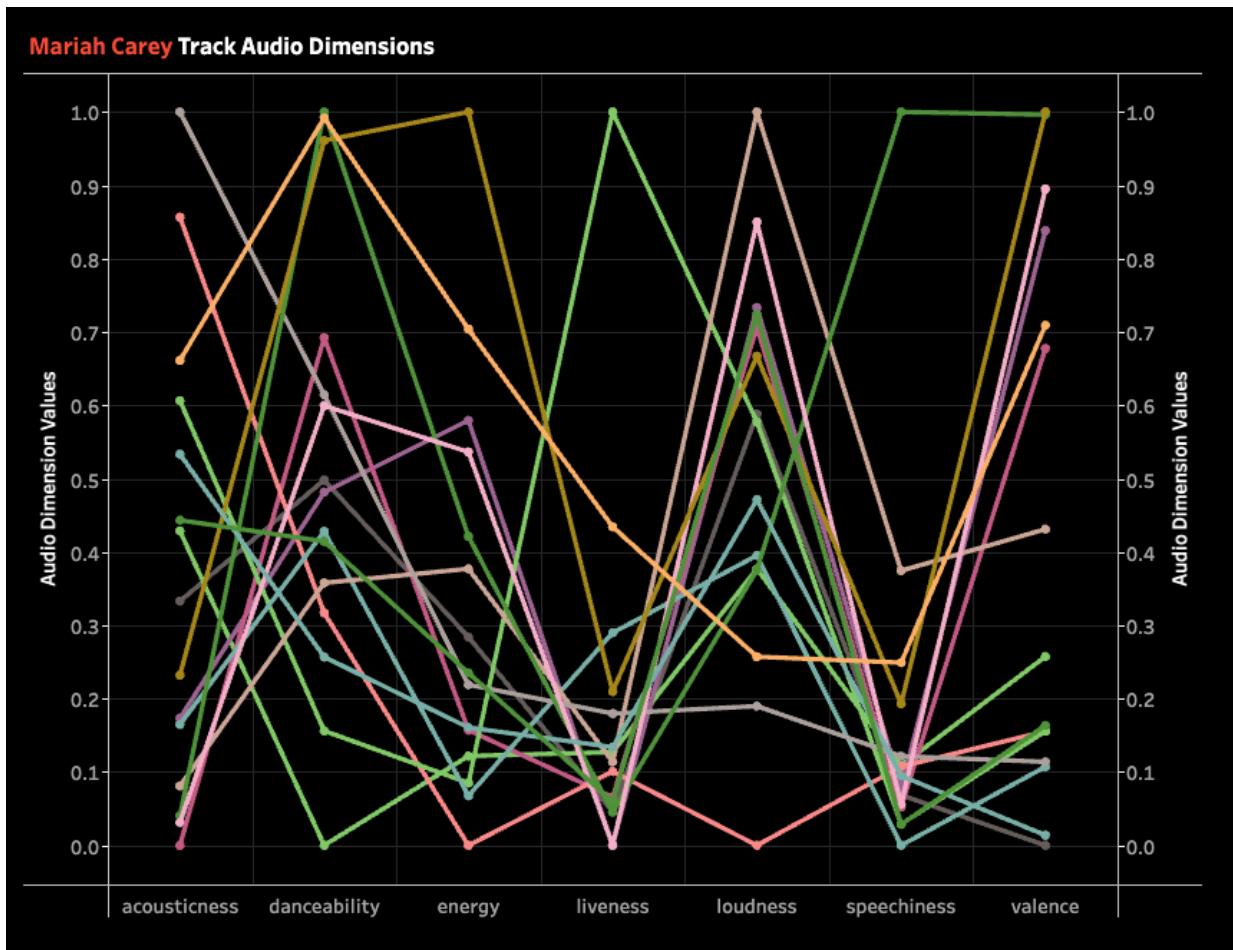


Figure 17: Final Version: Audio dimensions of a specific artist's tracks.

For our final audio dimensions visualization, we decided that visualizing an aggregate view of all track's from an artist would be more interesting to users. We undertook this decision because it was too complicated to display the parallel coordinates for multiple tracks on one graph, and our users struggled to comprehend the data. We intended to do radial plots but that was very challenging with Tableau and for our data being interlinked to the other views.

Usability Study

Five participants were shown two prototype visualizations—one visualization was about the proportion of Hip Hop vs. non-Hip Hop songs (**Figure 12**) and the other was about the different sound attributes such as valence, loudness, etc. of Hip Hop songs (**Figure 16**). The participants were asked the same set of questions (full list of questions in [Appendix B](#)) for the two visualizations and each interview lasted about twenty minutes. We recorded the observations and notes in a spreadsheet and analyzed the responses with qualitative coding to understand the key findings:

USABILITY FEEDBACK #1	
User Need	4 of 5 users wanted a deep dive into the songs & artists
Example Quote	"I want to see what songs these were! What Hip Hop or pop/rock songs were popular these years, this (proportion of Hip Hop vs. non-Hip Hop) doesn't tell me that much by itself as-is, not sure what to do with this." — P4
Changes Made	We added new visualizations that display the name of the artists and songs as well as song thumbnails of their Spotify album cover during the different Hip Hop eras. We also added more interactivity so that users could search for their favorite artists and explore how that artist's songs performed on the chart and learn about their audio dimensions.

USABILITY FEEDBACK #2	
User Need	3 of 5 were confused by the chart position and year filters
Example Quote	"It was confusing when I selected 2010 and it removed that year from the calendar, instead of excluding all the other years." — P2
Changes Made	We edited the filters so that they now show the year that is clicked on and excludes the others.

USABILITY FEEDBACK #3	
User Need	5 of 5 expected to see definitions of each audio feature
Example Quotes	<p>"I don't know what danceability means so I don't know what I'm looking at." — P1</p> <p>"Listening to the sound corresponding to the attribute would be helpful for someone who doesn't understand the value." — P4</p>
Changes Made	We made a new legend visualization that displays the definitions of the different sound attributes as well as examples of songs that represent low and high values of that attribute.

We also incorporated other changes that our users mentioned. One of the users said, *"Maybe if I clicked into a year and were able to follow the trend of every Hip Hop song that hit the chart and see how their position changed from week to week, that would be really cool."* We included this feature by showing how the chart position of songs for different artists has changed during the course of an era.

A user also mentioned that it would be helpful to get some context for the era. We incorporated that in one of our timelines for the Golden Era where we provided the overview of what the golden age of Hip Hop entailed. There was also a user expectation to see the top artist at a given time. We incorporated this by adding a song thumbnail for the top Spotify album during an era. We also have this in our "Best Performing Hip Hop Singles of the Year" visualization.

EVALUATION OF FINAL VISUALIZATION

Overview and Detail Views



Figure 18: The Golden Age Era Overview Dashboard

The Golden Age Era overview dashboard allows users to zoom in and explore specific time periods of interest, providing more detailed information and insights about the Hip Hop genre during those periods.

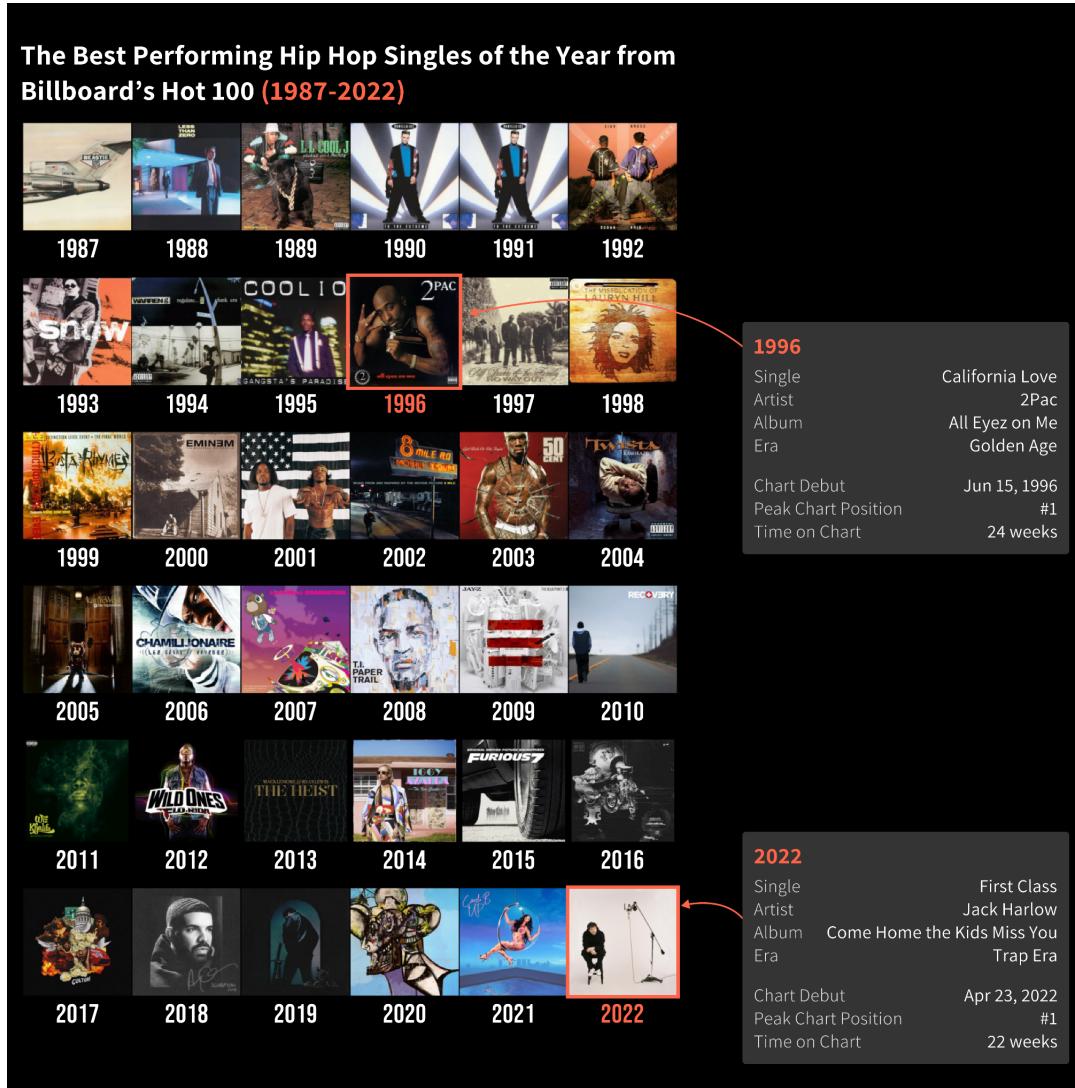


Figure 19: Tableau story of year's top Hip Hop albums

The inclusion of interactive album art that leads to corresponding Wikipedia pages provides users with additional information about the albums and helps deepen their understanding of the artists' discography.

Interaction Techniques and System Features

- The hover feature that displays details about the song, artist, and chart position when mousing over a data point provides users with on-demand details, enabling them to access specific information about individual data points without cluttering the main view.

- The ‘Hip Hop Artist on the Billboard Hot 100’ bar chart in the Golden Age dashboard is divided into the songs for that artist and clicking on a specific ‘box’ within that bar chart changes the other two graphs so that the user can quickly see the chart position and audio dimensions for the song or artist that they choose to explore.
- The search feature in the Billboard Hot 100 visualization allows users to explore their favorite artists easily, enhancing the usability and personalization of the visualization.
- The audio dimensions legend with song thumbnails and examples of different sounds provides a useful reference and interactive element for users to understand and explore the characteristics of Hip Hop songs in different eras.
- The filter functionality for Hip Hop vs. non-Hip Hop percentage allows users to focus on specific subsets of data, enabling them to analyze and compare the performance of Hip Hop songs against other genres.

Shneiderman's Information Visualization Tasks

We assessed our visualization against standardized tasks (Shneiderman, 1996, 337).

- **Overview:** The “Best Performing Hip Hop Singles of the Year” visualization provides an overview of the popular Hip Hop track of every year from 1987-2022. The album covers provide a relatable way for users to explore the songs and artists that define Hip Hop.
- **Zoom:** The breakdown of eras allows users to zoom in and explore a specific time period in more detail.
- **Filter:** The filter functionality enables users to focus on specific subsets of data, such as Hip Hop vs. non-Hip Hop songs and the search feature allows users to filter by their favorite artists.
- **Details-on-Demand:** The hover feature provides users with on-demand details about specific data points such as chart position, number of weeks on the Billboard chart, and the audio dimensions for the song.

It would be interesting to explore the relation task to compare the different Hip Hop eras. This could involve comparing audio dimensions of songs from different eras, or which artists have songs that have spanned different eras and continue to hit the charts. We could not do this within the constraints of this project, but it would be a fascinating exploratory tool.

The link to our Tableau Workbook is in [Appendix C](#).

CONCLUSION

Limitations of Tableau

Tableau was a useful tool to understand overall patterns and trends in our data. It was intuitive to use when uploading data files, dragging and dropping, and adding filters to explore the data through various dimensions. The sheets provided a way to experiment with different visualizations and test our ideas. The undo feature was also helpful as it allowed us to experiment without the anxiety of losing our previous ideas. The separation of the data by its type – numerical, date, etc. – was helpful to quickly eyeball which kind of comparisons we would like to make.

In terms of features that we wish Tableau had, a collaborative feature in the free student version would be valuable to understand each other's progress and share ideas easily. It would also be helpful to see a "History" tab similar to Photoshop so that we can clearly document all steps taken to get to the final visualization and duplicate if necessary.

It was challenging to dynamically incorporate images and Spotify embeds. We wanted our Tableau story to be very visual and engaging through the heavy use of these elements, but were unable to incorporate as many as planned because they needed to be statically and individually added, formatted, and positioned.

Our initial sketches utilized multiple linked visualizations. We were advised this was impossible with Tableau, so we modified our story to have some linked visualizations on the same page, but were unable to build out our sketches exactly as ideated, and compromised on functionality and flow.

Limitations of Data

Due to the imperfections of matching the artist name in the future it would be helpful to change the scope if necessary and use Hip Hop-specific charts to improve accuracy and reduce the manual cleaning load. Even as-is, there is the potential for the dataset to contain non-Hip Hop songs from artists with common names. The ones removed were visible and apparent as they were in the 50s and 60s, one way to do this would be to employ web scraping to get the "years active" field from Wikipedia and automatically discount any tracks from the supposed artist outside of that range. One could also increase certainty by seeing if each song was in that artist's discography.

Popular tracks, as the ones in the Hot 100 chart, often have multiple releases in Spotify from being rerecorded, a radio edit, explicit version, both an original single and occurrences in compilation albums, etc. Spotify's track ID only references one specific instance of the track, generally the most popular. This could reflect in fields such as popularity, explicitness, and acoustic features being non-comprehensive for all released versions of the song.

Overall with creating the dataset through custom algorithms and cleaning, there is room for future improvement on sanitizing the data.

Further Work

We hope to replicate the process of creating an overview for each Hip Hop era as we did for the Golden Age Overview visualization. There are several aspects of the design system that we would like to improve upon. For example, we could use a better relation functionality that allows users to compare different eras. It would also be interesting to note the political and major event milestones that shaped the lyrics of Hip Hop. We would like to explore some geographical aspects such as how the sound of Hip Hop differs across the different areas of the US.

Mouseover Takeaways: Implement a tooltip feature that displays additional information when the user hovers over a specific year or data point. This tooltip can include key highlights, such as notable artists, their hit songs, chart records, or subgenre shifts. For example, when hovering over 2011, the tooltip could show information like "Drake hits #1 with [song] for the first time and stays for 8 weeks."

Subgenre Shifts: Create a separate visualization or section dedicated to showcasing subgenre shifts within Hip Hop over the years. This can include a line chart or stacked bar chart that illustrates the prominence and evolution of various subgenres. When the user hovers over a specific year or subgenre, relevant information about that shift can be displayed in a tooltip or accompanying text. This can help users understand the changing landscape of Hip Hop and its sub-genres over time.

Regional Impact: Incorporate a geographical map or interactive region-based visualization to showcase the regional impact on Hip Hop over the years. When the user hovers over a specific region or year, relevant highlights and key artists from that region can be displayed. This provides users with insights into how different regions influenced the Hip Hop genre and its evolution.

Acknowledgments

Thanks to Negeen Aghassaiabake, University of Washington's Data Visualization Librarian, for consulting with us on Tableau. None of us had previous experience with Tableau and it was incredibly helpful. We'd also like to thank our five usability study participants for their time to help us iterate early on in the process, and our teaching instructors Cecilia Aragon and Andrea Figueroa, and teaching assistant Sourojit Ghosh.

REFERENCES

Best Year in Music. (2019). The Pudding. Retrieved May 5, 2023, from

<https://pudding.cool/projects/music-history/>

Daniels, M. (n.d.). *Crowdsourcing the Definition of "Punk".* The Pudding. Retrieved June 5, 2023, from

<https://pudding.cool/2017/03/punk/index.html>

Daniels, M., White, E., & Anderson, T. (2015). *This is what Hip Hop's Billboard top 10 sounded like, back in 1995.* Polygraph. Retrieved May 5, 2023, from <http://poly-graph.co/billboard/>

Deal, M. (n.d.). *Charting the Beatles / Infographic atlas of Beatles music.* Retrieved June 5, 2023, from
<https://chartingthebeatles.net>

Hip hop music. (2019, April 19). Wikipedia. Retrieved May 5, 2023, from

https://en.wikipedia.org/wiki/Hip_hop_music

Lamere, P. (2023, April 7). *Spotify.* Welcome to Spotify! — spotify 2.0 documentation. Retrieved April 17, 2023, from <https://spotipy.readthedocs.io/en/2.22.1/#>

List of hip hop groups. (2023, April 19). Wikipedia. Retrieved April 20, 2023, from
https://en.wikipedia.org/wiki/List_of_hip_hop_groups

List of hip hop musicians. (2023, April 19). Wikipedia. Retrieved April 20, 2023, from
https://en.wikipedia.org/wiki/List_of_hip_hop_musicians

Miller, S. (2023, January 9). *Billboard-Hot-100.* Github. Retrieved 20 April, 2023, from
<https://github.com/HipsterVizNinja/random-data/tree/main/Music/hot-100>

Rose, T. (1994). *Black Noise: Rap Music and Black Culture in Contemporary America.* Wesleyan University Press.

Shneiderman, B. (1996). The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations.

IEEE, 336-343. 0-8186-7469-5/96

Spotify. (n.d.). *Spotify API*. Spotify for Developers. <https://developer.spotify.com/dashboard>

Wilber, J. (2018, June 5). *Best Year in Music*. The Pudding. Retrieved June 5, 2023, from

<https://pudding.cool/projects/music-history/>

APPENDIX A: PYTHON SCRIPT

Link to Python Script for building dataset: <https://github.com/luka-sherman/hiphopdataviz>

APPENDIX B: USABILITY QUESTIONS

1. What do you think this visualization is about?
2. Take a few minutes to explore this visualization and tell me everything you can find.
3. Could you tell me the proportion of Hip Hop and non-Hip Hop songs in 2010?
4. Which part of the visualization was the easiest to navigate?
5. Which part of the visualization was the hardest to navigate?
6. What data dimensions would you like to explore further in this dataset?
7. What would make the visualization more interesting to you?

APPENDIX C: FINAL VISUALIZATION

Link to Tableau Public:

https://public.tableau.com/app/profile/gabriella.cronk/viz/511VizTheRiseandEvolutionofHipHop6_5/E RASFINAL?publish=yes



Figure 20: Screenshot of Final Visualization hosted on Tableau Public.