



Max Branca - 2061432

# Data Warehouse Project

A multidimensional data warehouse to explore trends in audio features, genres, and commercial success across decades.

# Project Objectives

This project aims to design and implement a unified data warehouse that brings together diverse aspects of the music industry, including audio features, artist profiles, genre classifications, and chart performance. The goal is to create a flexible analytical platform capable of revealing long-term trends and correlations in popular music and how musical styles, popularity, and artist success evolve over time.

# Tools & Technologies

This project leveraged three main tools across different stages of the workflow:

## Kaggle

Used as the primary source for collecting the datasets. Kaggle offers a wide range of open datasets, including those related to music, artists, and song popularity.

## Python

Employed during the entire ETL process. Python was used to extract, clean, merge, and transform the data into a structured format ready for analysis.

## Tableau

Utilized for the analytical and visualization phase. Tableau allowed the creation of charts, graphs and tables to explore trends and insights.

# Dataset Selection

For this project, I selected **four datasets from Kaggle** to build a complete picture of the music world:

- Two datasets focus on song features: the **Song Features Dataset** and the **Spotify 1.2M+ Songs Dataset** provide information like song name, artist name and audio characteristics.
- One dataset covers artist information: the **Worldwide Music Artists Dataset** includes artist names, genres, and countries.
- One dataset tracks song rankings: the **Billboard Hot 100 Songs Dataset** contains weekly chart positions of the most popular songs in the U.S. since 1958.

# ETL Process Overview

The ETL (Extract, Transform, Load) process followed a structured approach to prepare the datasets for integration:

## Cleaning

Removed duplicates, handled missing values, and normalized text fields to ensure consistency across all datasets.

## Merging Song Features

The two datasets containing song-level audio features were combined into a single, unified dataset with consistent structure and attributes.

## Data Standardization

Standardized key fields such as `artist_id` and `song_id` across the merged datasets, the artist dataset, and the chart dataset. This ensured reliable joins between tables.

# ETL Results

**charts.csv**

Tipo	Nome campo
Abc	Song Id (Charts Final With Artist Id.Csv)
白	Date
#	Rank
Abc	Song
Abc	Artist
#	Peak-Rank
#	Weeks-On-Board
Abc	Artist Id (Charts Final With Artist Id.Csv)

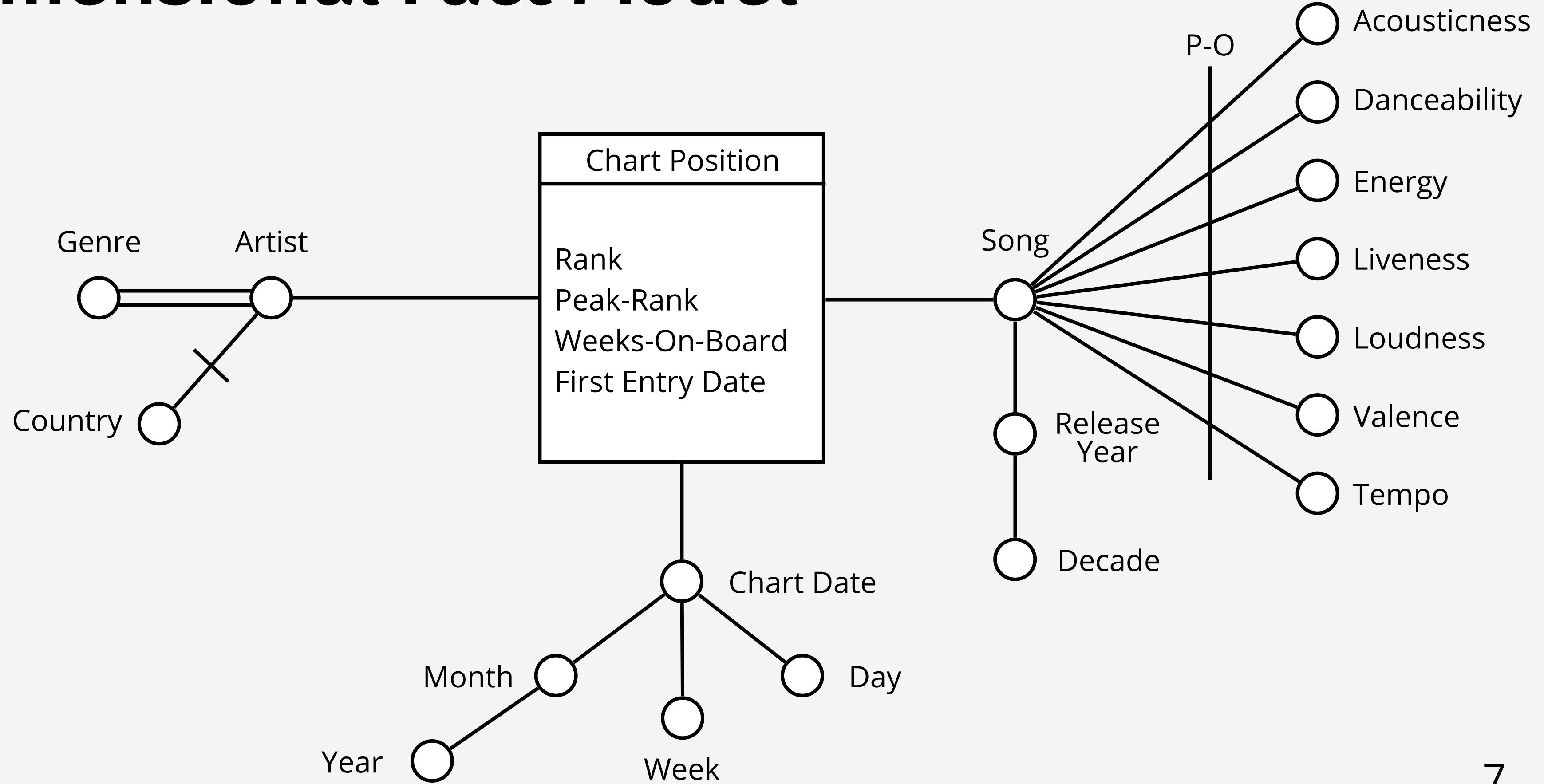
**tracks.csv**

Tipo	Nome campo
Abc	Song Id
Abc	Name
Abc	Artists
#	Danceability
#	Energy
#	Loudness
#	Acousticness
#	Liveness
#	Valence
#	Tempo
#	Year
Abc	Artist Id
=#	Decade

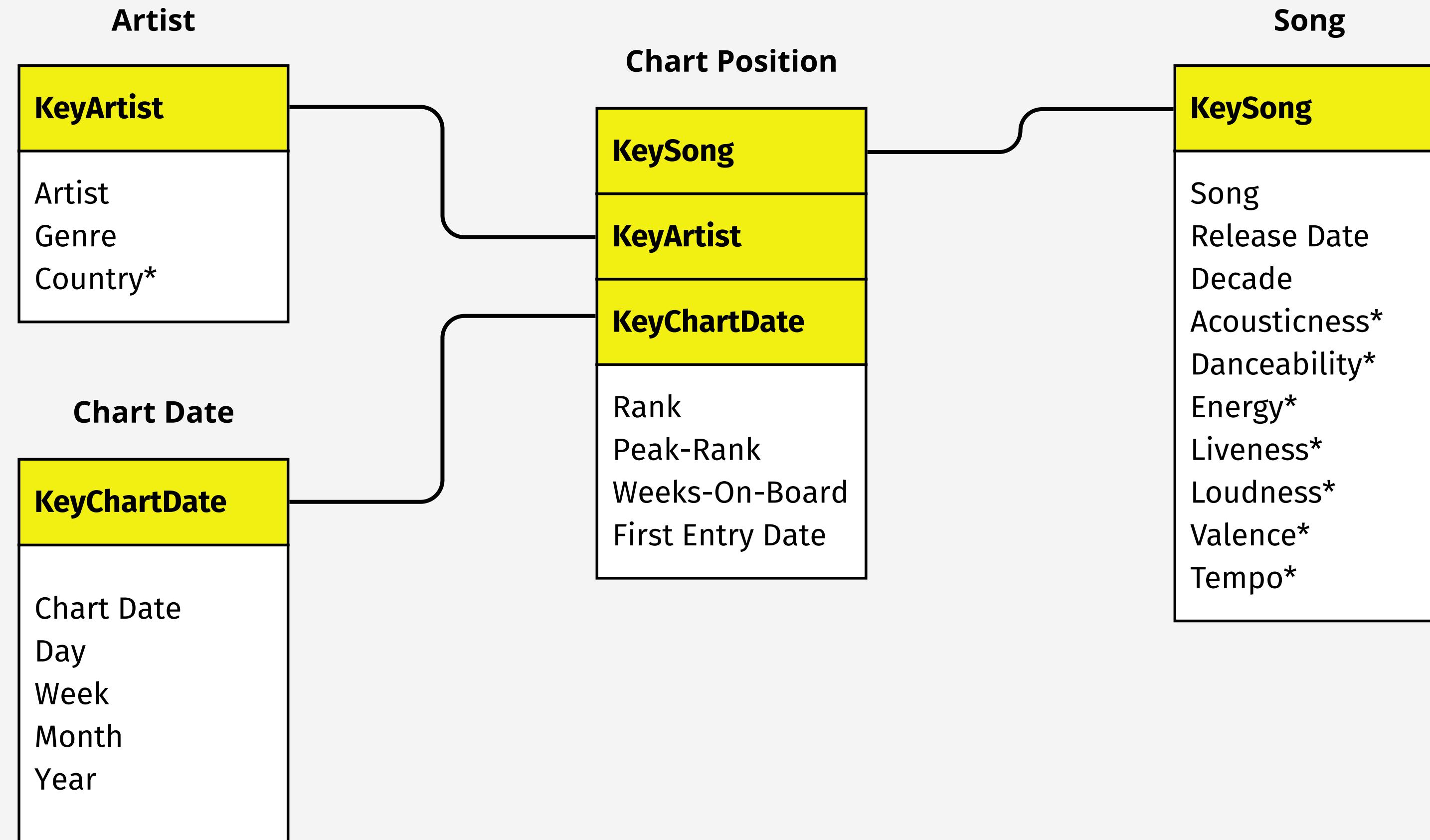
**artists.csv**

Tipo	Nome campo
Abc	Artist Id (Artists Final No Img.Csv)
Abc	Artist Name
Abc	Artist Genre
🌐	Country
=Abc	Main Genre

# Dimensional Fact Model



# Star Schema

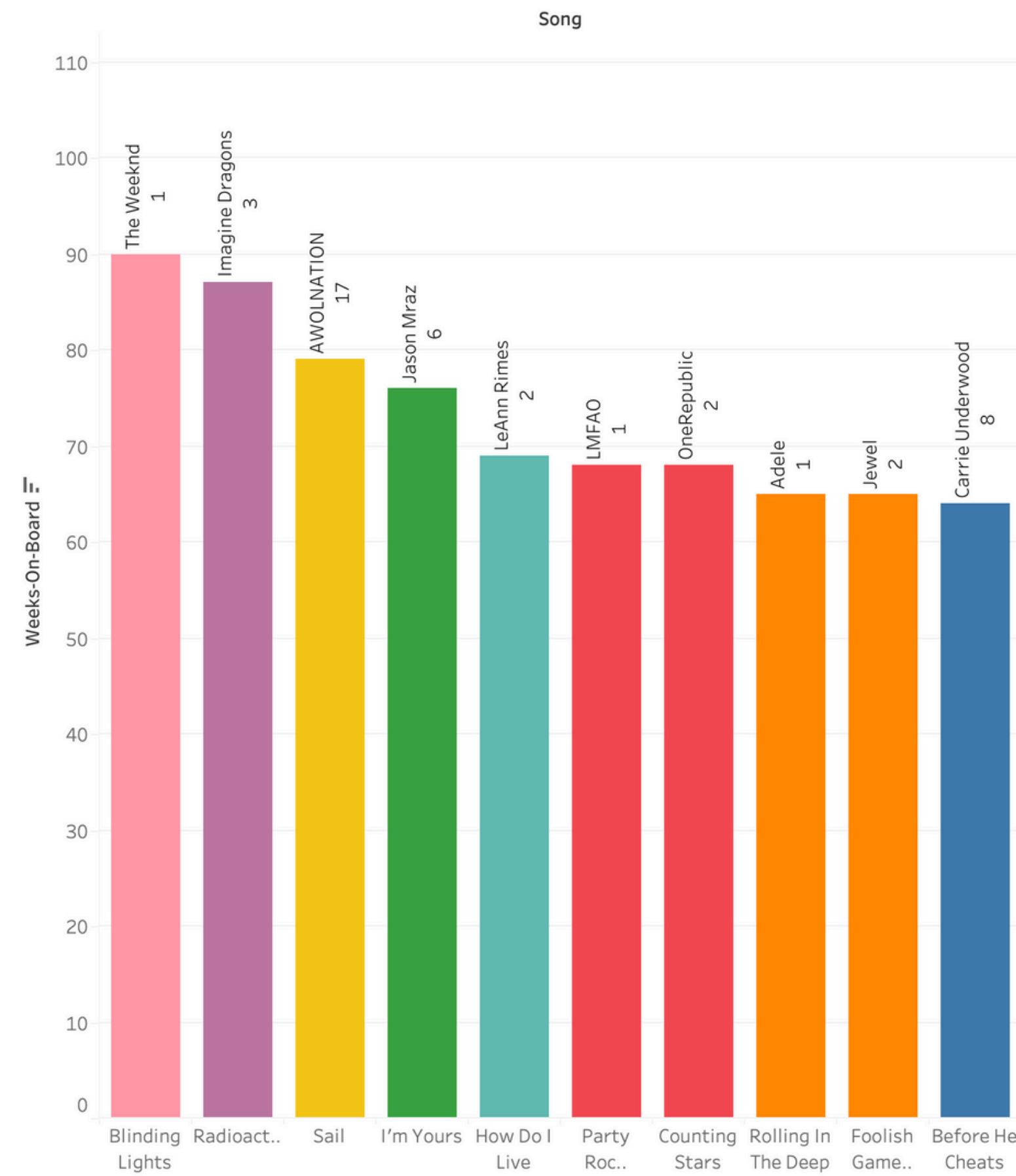


# Top 10 Longest Charting Songs on Billboard Hot 100

This bar chart displays the **Top 10 songs with the highest number of weeks on the Billboard Hot 100**, highlighting their lasting popularity over time.

- "**Blinding Lights**" by **The Weeknd** leads with **90 weeks**, making it the longest-charting song in Billboard history.
- Other notable long-runners include:
  - "Radioactive" by Imagine Dragons (87 weeks)
  - "Sail" by AWOLNATION (79 weeks)
  - "I'm Yours" by Jason Mraz (76 weeks)

These results showcase how certain songs maintain public attention and commercial relevance well beyond their initial release, offering insight into enduring musical appeal.

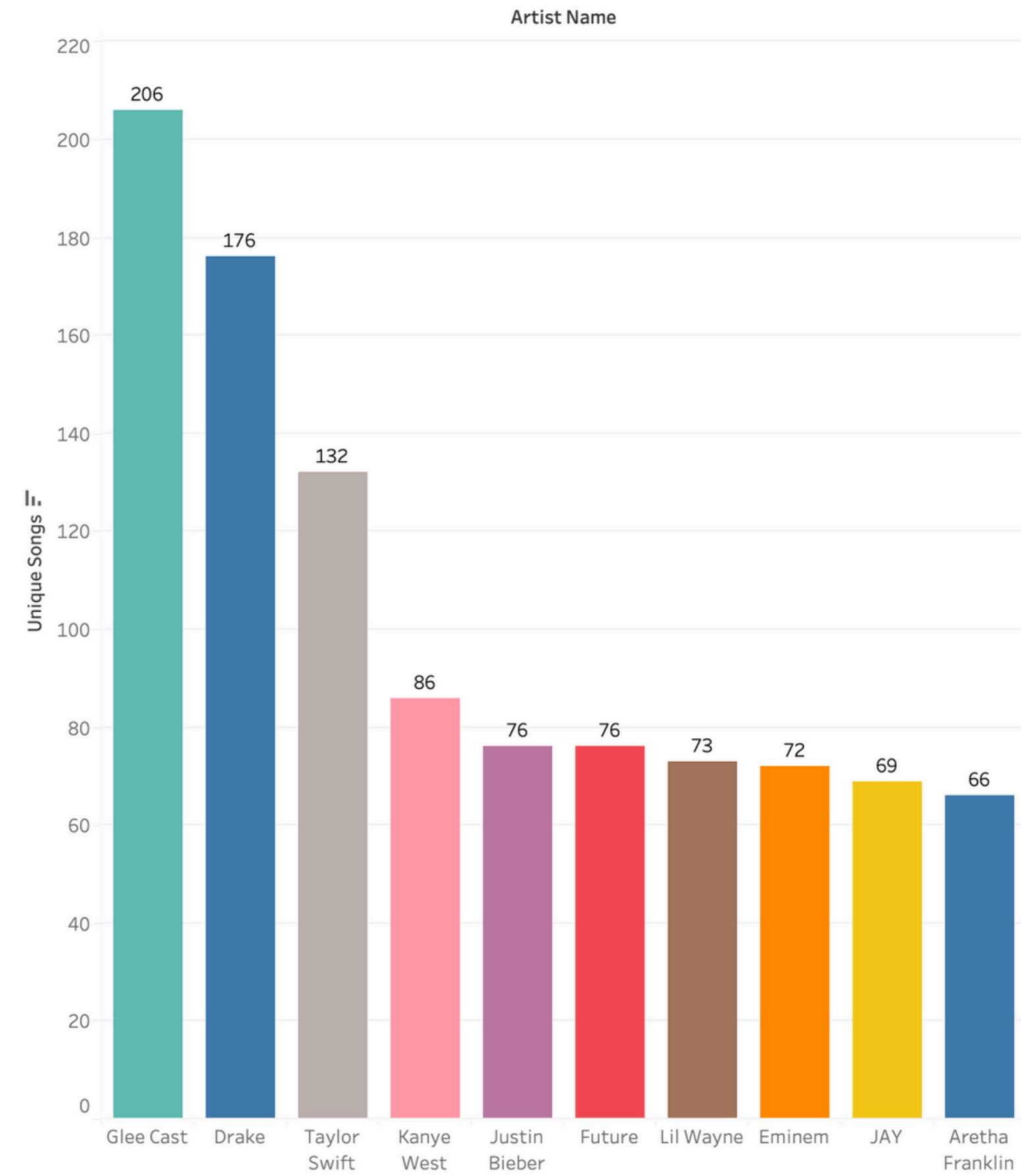


## Top 10 Artists with most Unique Songs on Billboard Hot 100

This bar chart highlights the **Top 10 artists with the highest number of unique songs that entered the Billboard Hot 100.**

- **Glee Cast** leads with an impressive **206** unique charting songs, driven by the show's weekly musical releases.
- **Drake** follows with **176** songs, reflecting his high volume of releases and frequent chart presence.
- Other prolific artists include:
  - Taylor Swift (132 songs)
  - Kanye West, Justin Bieber, and Future (all with 70+ songs)

This insight emphasizes the dominance of certain artists in popular music, not just through hits but through the quantity of songs that reach public attention.

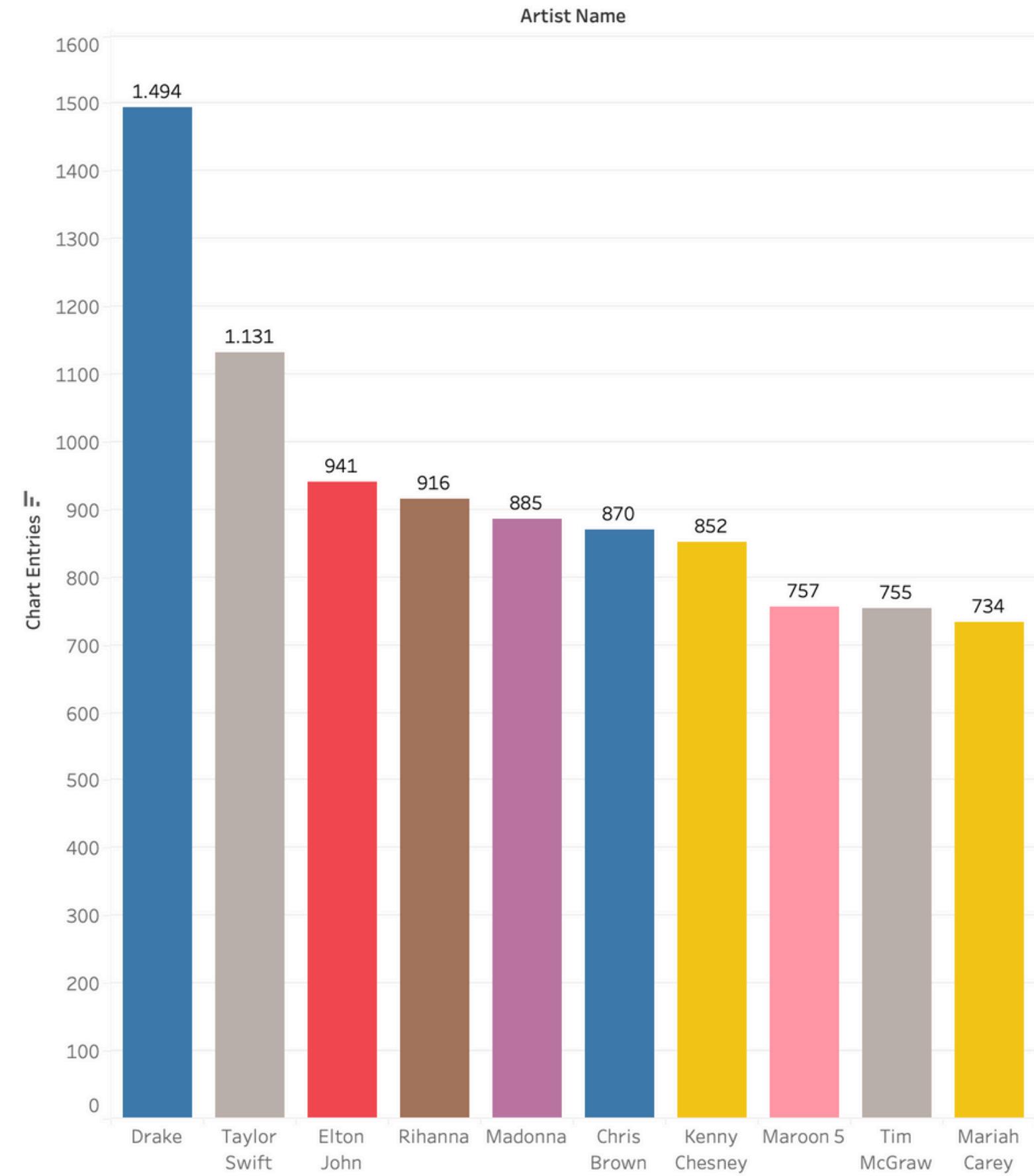


## Top 10 Artists by Chart Entries on Billboard Hot 100

This visualization shows the **Top 10 artists with the highest number of total chart entries on the Billboard Hot 100**, counting all weekly appearances.

- **Drake** dominates with **1,494** chart entries, highlighting his sustained presence and output in recent years.
- **Taylor Swift** follows with **1,131** entries, reinforcing her position as one of the most consistently successful artists.
- Other major contributors include:
  - Elton John (941)
  - Rihanna (916)
  - Madonna, Chris Brown, and Mariah Carey — all with over 700 entries

This metric captures not only the number of songs but also their longevity and recurrence on the charts.



## Top 10 Songs with Longest Consecutive Weeks at #1

This table shows the **Top 10 songs that held the #1 position on the Billboard Hot 100 for the most consecutive weeks:**

Song	Artist Name	
Old Town Road	Lil Nas X	19
Despacito	Luis Fonsi	16
One Sweet Day	Mariah Carey	16
Candle In The Wind 1997/Something About The Way You Look Tonight	Elton John	14
I Gotta Feeling	The Black Eyed Peas	14
I Will Always Love You	Whitney Houston	14
I'll Make Love To You	Boyz II Men	14
Macarena (Bayside Boys Mix)	Los Del Rio	14
Uptown Funk!	Mark Ronson	14
We Belong Together	Mariah Carey	14

- "**Old Town Road**" by **Lil Nas X** stands out with **19** consecutive weeks at #1 — a record-breaking run.
- "Despacito" (Luis Fonsi) and "One Sweet Day" (Mariah Carey) follow with 16 weeks each.
- Several iconic tracks stayed at the top for 14 consecutive weeks, including:
  - "I Gotta Feeling" – The Black Eyed Peas
  - "Uptown Funk!" – Mark Ronson
  - "We Belong Together" – Mariah Carey

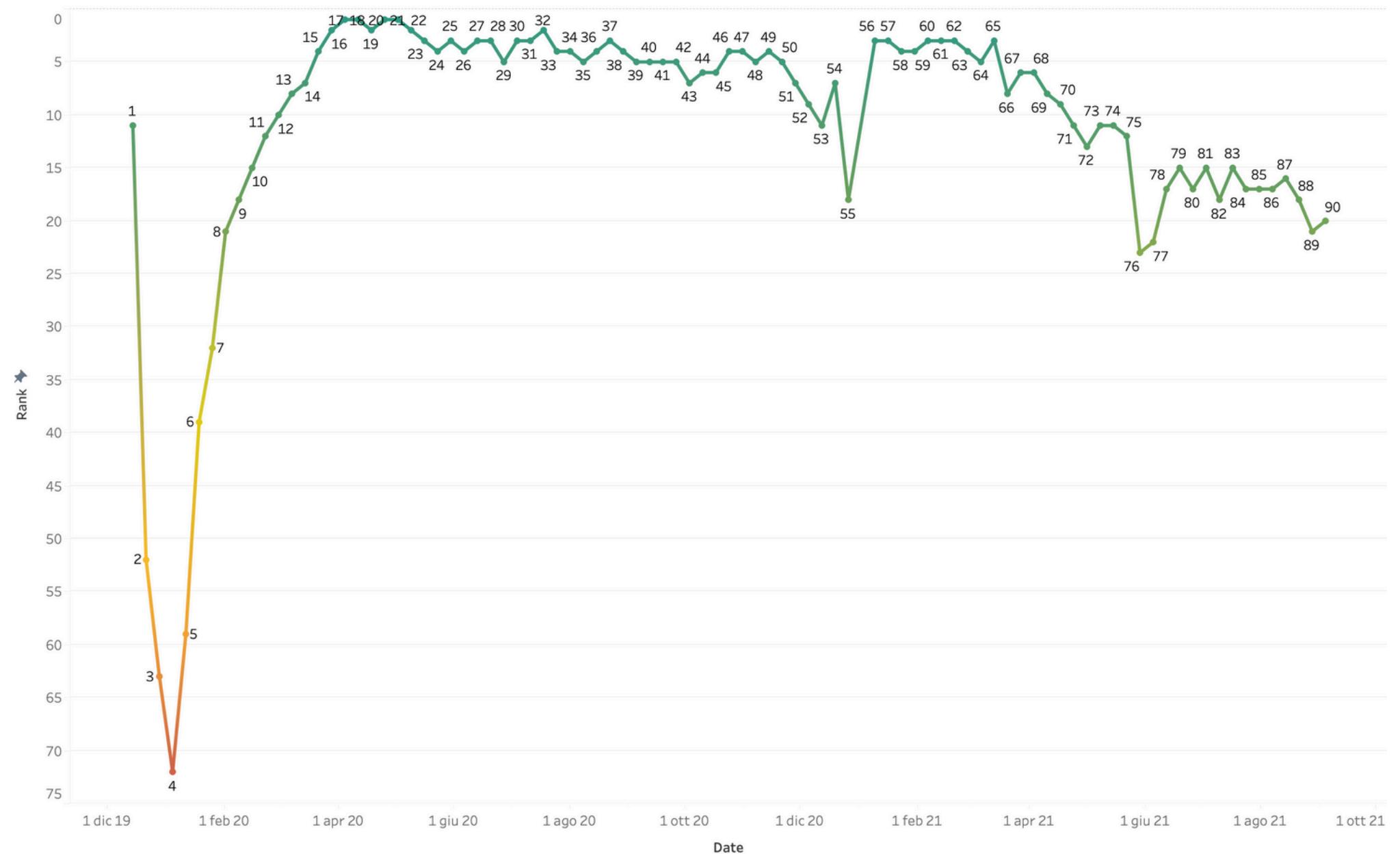
These songs achieved rare levels of popularity, dominating public attention over long periods.

## Chart Trajectory of Blinding Lights by The Weeknd

This line chart illustrates the **Billboard Hot 100 ranking trajectory of "Blinding Lights"** from its release through its chart lifespan.

- The song debuted outside the top 10, quickly rose to #1, and maintained a dominant position for a remarkably extended period.
- After peaking, it remained in the Top 10 for many weeks, showcasing long-lasting popularity.
- Even as it gradually descended, "Blinding Lights" demonstrated exceptional chart resilience, contributing to its **record-setting 90-week total run**.

The trajectory reflects not just a hit song, but a cultural phenomenon with sustained impact over nearly two years.

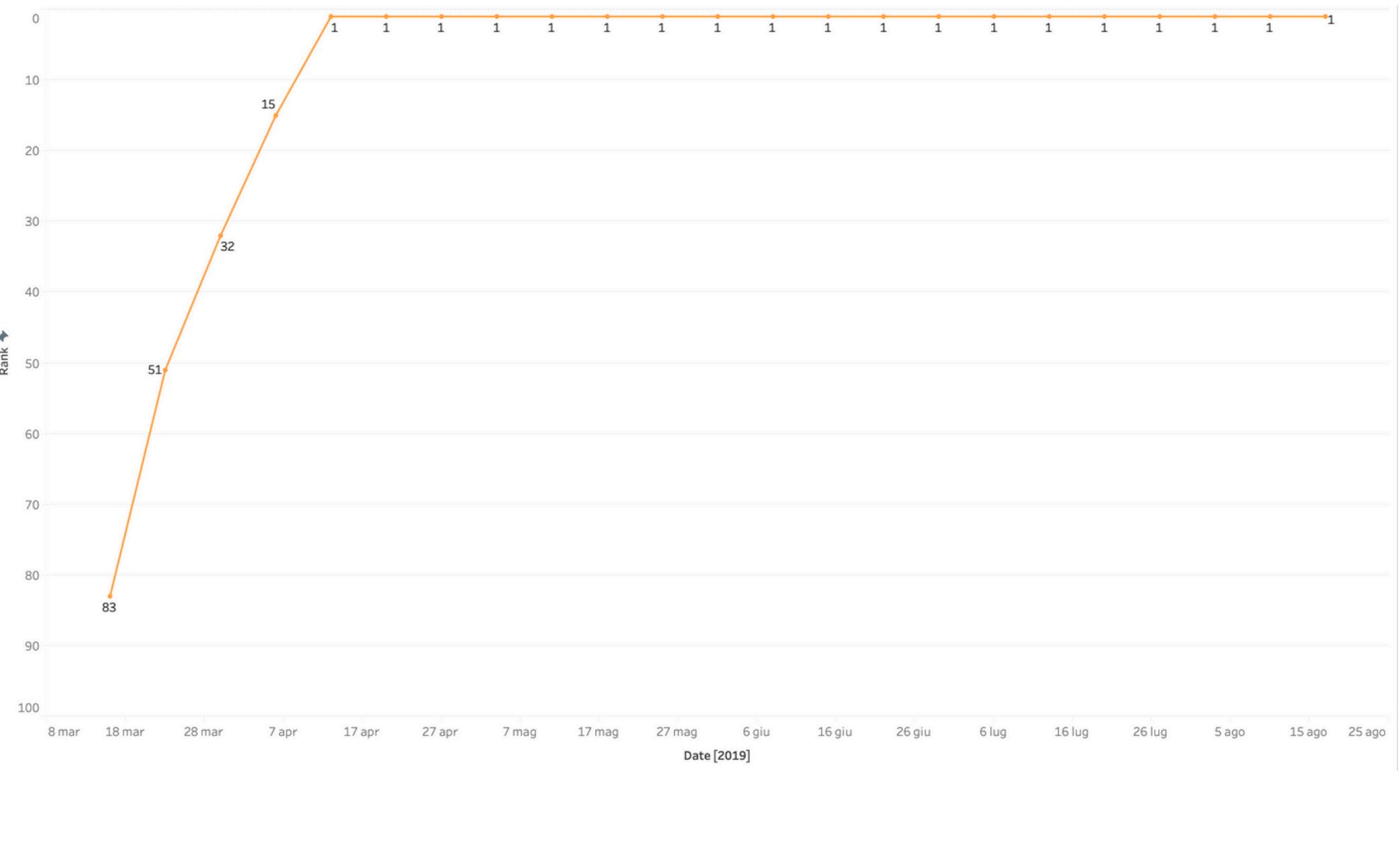


# Chart Trajectory of Old Town Road by Lil Nas X

This chart tracks the rise and dominance of **"Old Town Road"** on the **Billboard Hot 100**.

- The song entered the chart at position #83, then climbed rapidly week by week.
- **Within just four weeks, it reached the #1 spot, where it remained for 19 consecutive weeks** — setting the all-time Billboard record for the longest run at #1.
- Its trajectory illustrates a meteoric rise and exceptional staying power, driven by virality, remixes, and cross-genre appeal.

This case demonstrates how modern distribution channels and internet culture can amplify a song's impact beyond traditional industry patterns.



## Instant Hits: Songs That Debuted at #1

This table lists **songs that entered the Billboard Hot 100 directly at the #1 position**, marking an immediate and powerful impact on release.

- Examples of instant chart-toppers include:
  - **"7 Rings" – Ariana Grande (2019)**
  - **"Drivers License" – Olivia Rodrigo (2021)**
  - **"God's Plan" – Drake (2018)**
- In addition to the song title and artist name, the table also displays the total number of weeks each song remained on the Billboard chart and the week when the song first charted.

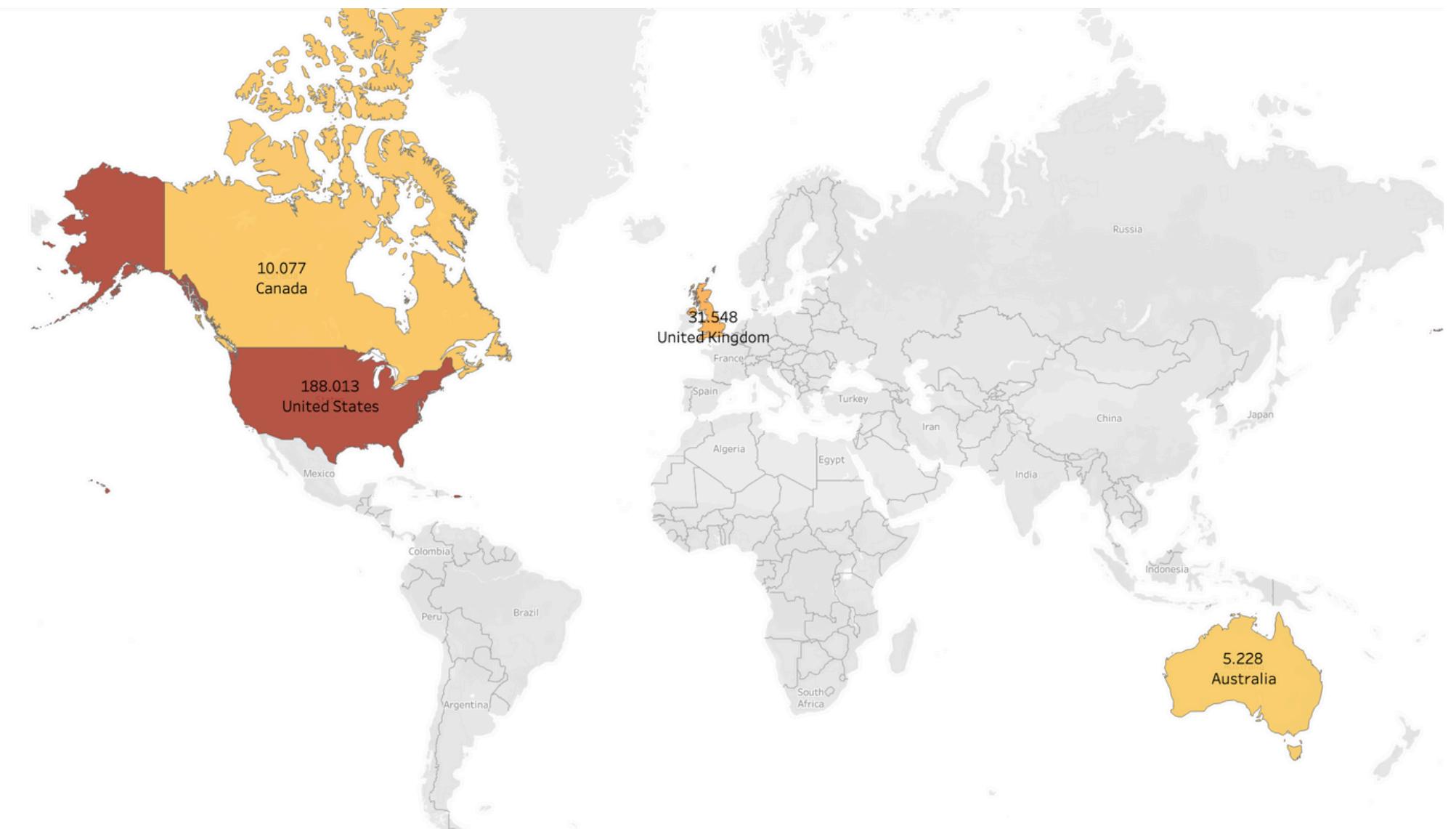
This phenomenon reflects the influence of digital releases, strong fanbases, and media anticipation in driving songs to debut at the very top of the charts.

Song	Artist Name	Max. Week..	
3	Britney Spears	20	24/10/2009
7 Rings	Ariana Grande	33	02/02/2019
Born This Way	Lady Gaga	20	26/02/2011
Butter	BTS	20	05/06/2021
Can't Stop The Feeling!	Justin Timberlake	52	28/05/2016
Candle In The Wind 1997/Somethin..	Elton John	42	11/10/1997
Cardigan	Taylor Swift	14	08/08/2020
Do I Make You Proud	Taylor Hicks	8	01/07/2006
Doo Wop (That Thing)	Lauryn Hill	21	14/11/1998
Drivers License	Olivia Rodrigo	28	23/01/2021
Dynamite	BTS	32	05/09/2020
Exhale (Shoop Shoop) (From "Waiti..	Whitney Houston	21	25/11/1995
Fantasy	Mariah Carey	25	30/09/1995
Franchise	Travis Scott	9	10/10/2020
God's Plan	Drake	36	03/02/2018
Good 4 U	Olivia Rodrigo	24	29/05/2021
Harlem Shake	Baauer	20	02/03/2013
HIGHEST IN THE ROOM	Travis Scott	22	19/10/2019
Hold It Against Me	Britney Spears	17	29/01/2011
I Believe	Fantasia	10	10/07/2004
I Don't Want To Miss A Thing	Aerosmith	20	05/09/1998
I'll Be Missing You	Puff Daddy	33	14/06/1997
I'm The One	DJ Khaled	22	20/05/2017
I'm Your Angel	R. Kelly	18	05/12/1998
Inside Your Heaven	Carrie Underwood	12	02/07/2005
Life Goes On	BTS	3	05/12/2020
Montero (Call Me By Your Name)	Lil Nas X	31	10/04/2021
My Heart Will Go On	Celine Dion	20	28/02/1998
My Universe	Coldplay	5	09/10/2021
Nice For What	Drake	25	21/04/2018
Not Afraid	Eminem	25	22/05/2010
One Sweet Day	Mariah Carey	27	02/12/1995
Part Of Me	Katy Perry	22	03/03/2012
Peaches	Justin Bieber	30	03/04/2021
Permission To Dance	BTS	7	24/07/2021
Pillowtalk	ZAYN	24	20/02/2016
Poor Little Fool	Ricky Nelson	11	04/08/1958

## Top 4 Contributing Countries to the Billboard Hot 100

This map highlights the top four countries by the total number of songs their artists have contributed to the Billboard Hot 100.

- **United States is by far the dominant contributor with 188,013 songs**, reflecting its central role in the global music industry.
- **United Kingdom follows with 31,548 songs**, known for consistently exporting successful artists and genres.
- **Canada contributed 10,077 songs**, boosted by artists like Drake, The Weeknd, and Justin Bieber.
- **Australia adds 5,228 songs**, underscoring its steady presence in international pop and rock scenes.

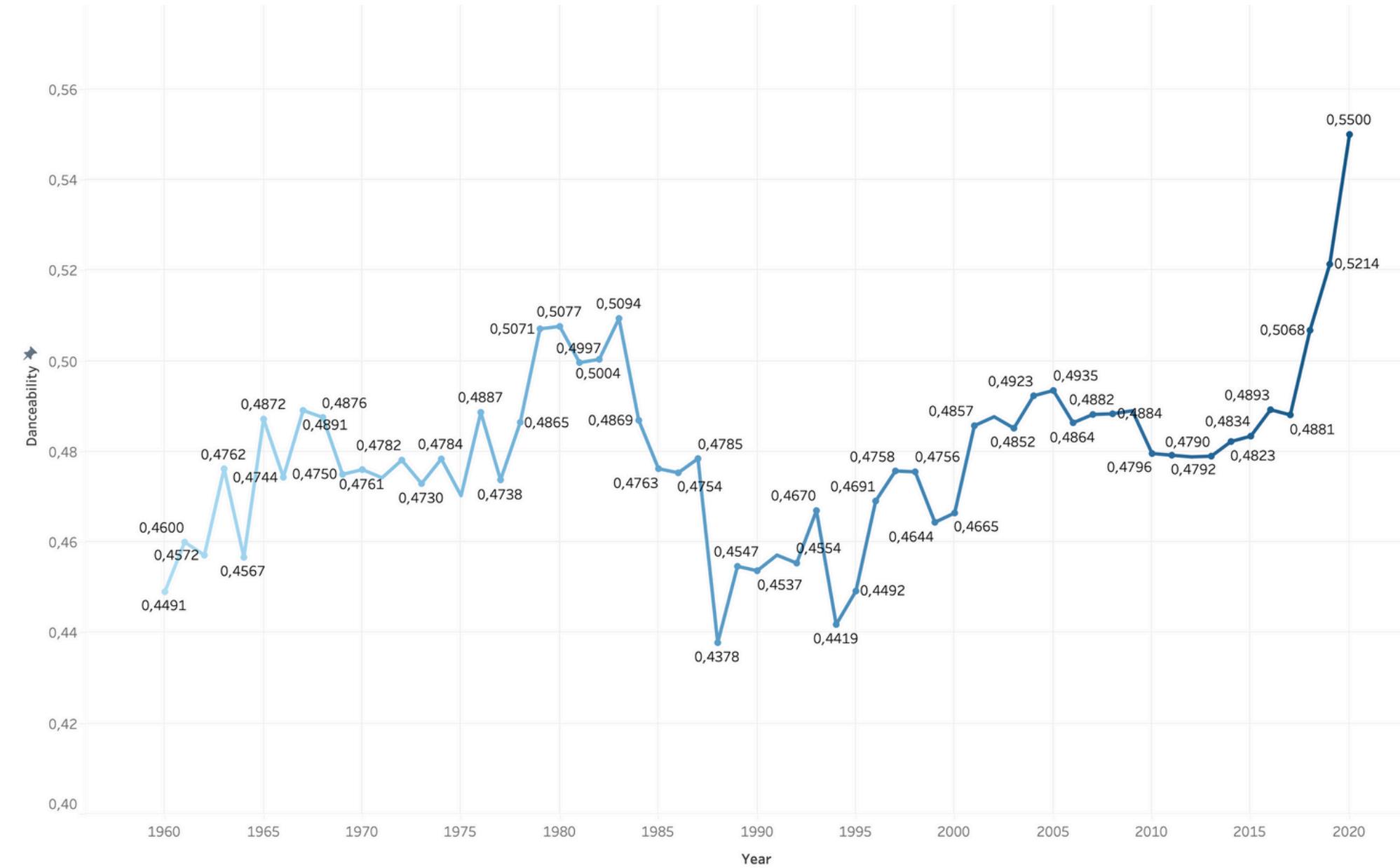


## Average Danceability of Songs (1960-2020) according to Spotify

This line chart shows the **yearly average danceability score of songs from 1960 to 2020**, based on audio feature data across the entire dataset.

- For most years, average danceability values range between 0.45 and 0.50, suggesting consistent musical rhythm and tempo over time.
- Starting around **2015**, there's a **sharp increase**, reaching an all-time high of 0.5500 in 2020.

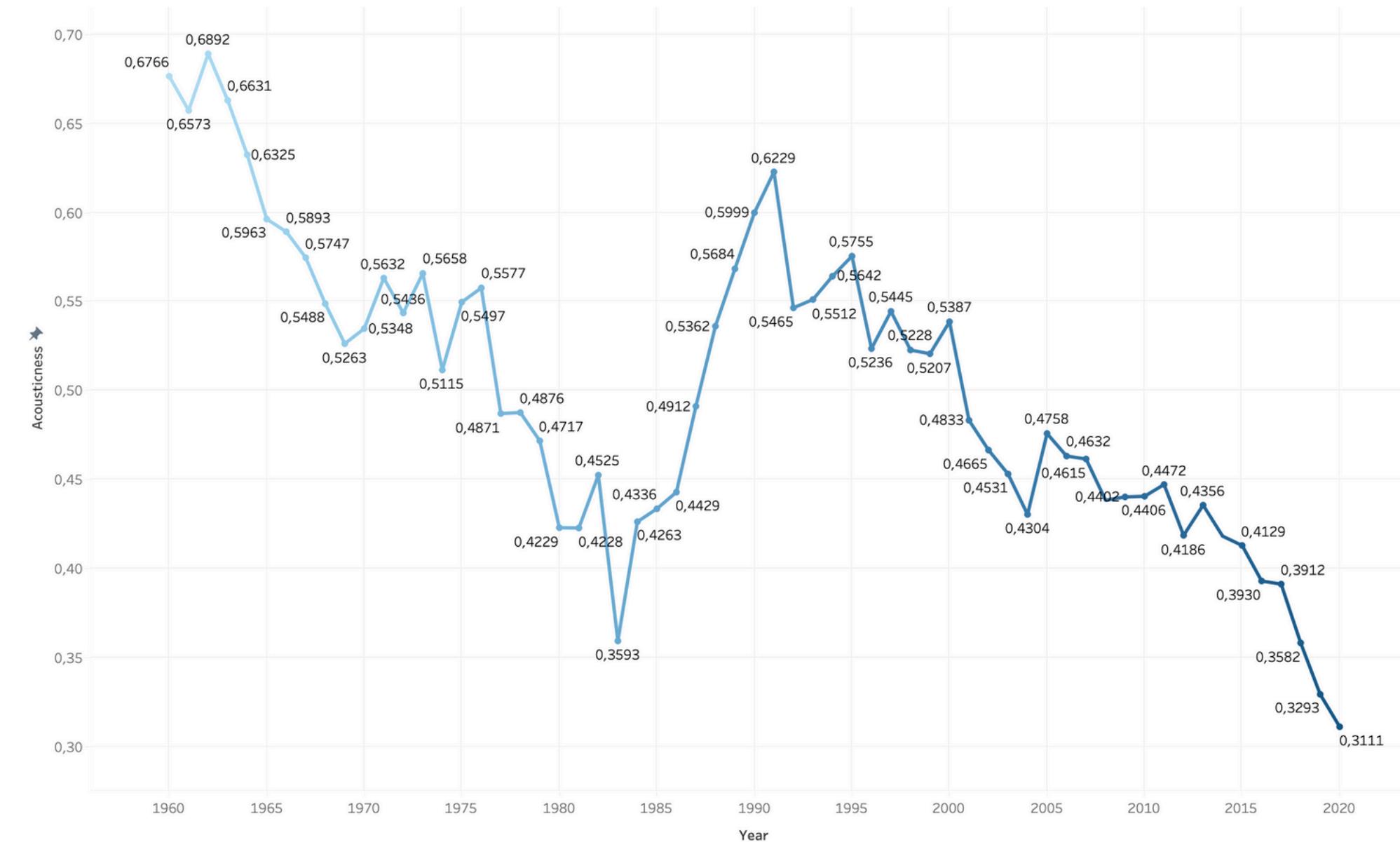
This upward trend indicates that **modern songs are becoming more danceable, likely due to the influence of electronic, hip-hop, and global pop styles** emphasizing strong beats and groove-friendly rhythms..



## Average Acousticness of Songs (1960-2020) according to Spotify

This line chart tracks the **yearly average acousticness of songs from 1960 to 2020** across the entire dataset. The acousticness metric reflects the likelihood that a track is acoustic, featuring natural instruments and minimal electronic processing.

- In the 1960s and early 1970s, acousticness values were high, often above 0.60, indicating a **strong presence of organic, instrument-driven music**.
- A steady **decline begins in the late 1970s**, coinciding with the rise of synthesized and digitally produced music.
- Since 2000, acousticness has dropped significantly, reaching a low of 0.3111 in 2020, marking a clear shift toward electronic and digitally enhanced soundscapes.

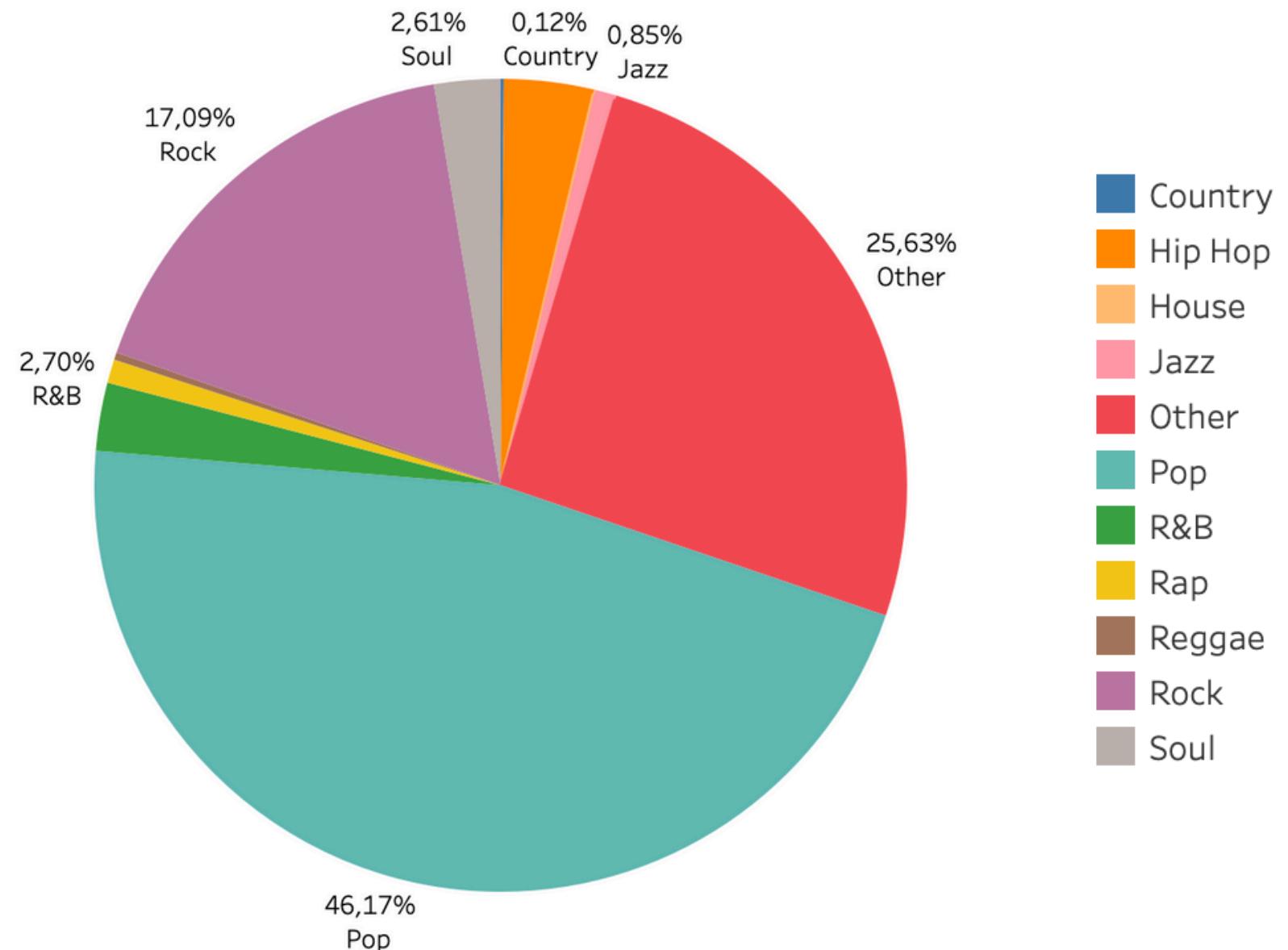


# Genre Distribution of #1 Billboard Hot 100 Hits

This pie chart shows the **distribution of main music genres among songs that have reached the #1 position** on the Billboard Hot 100.

- **Pop music dominates the chart, accounting for 46.17% of all #1 hits**, confirming its broad appeal and market reach.
- **Rock follows with 17.09%**, reflecting its long-standing influence, especially in earlier decades.

This chart illustrates how **Pop and Rock have consistently led mainstream success**, while other genres have made more niche but impactful contributions.

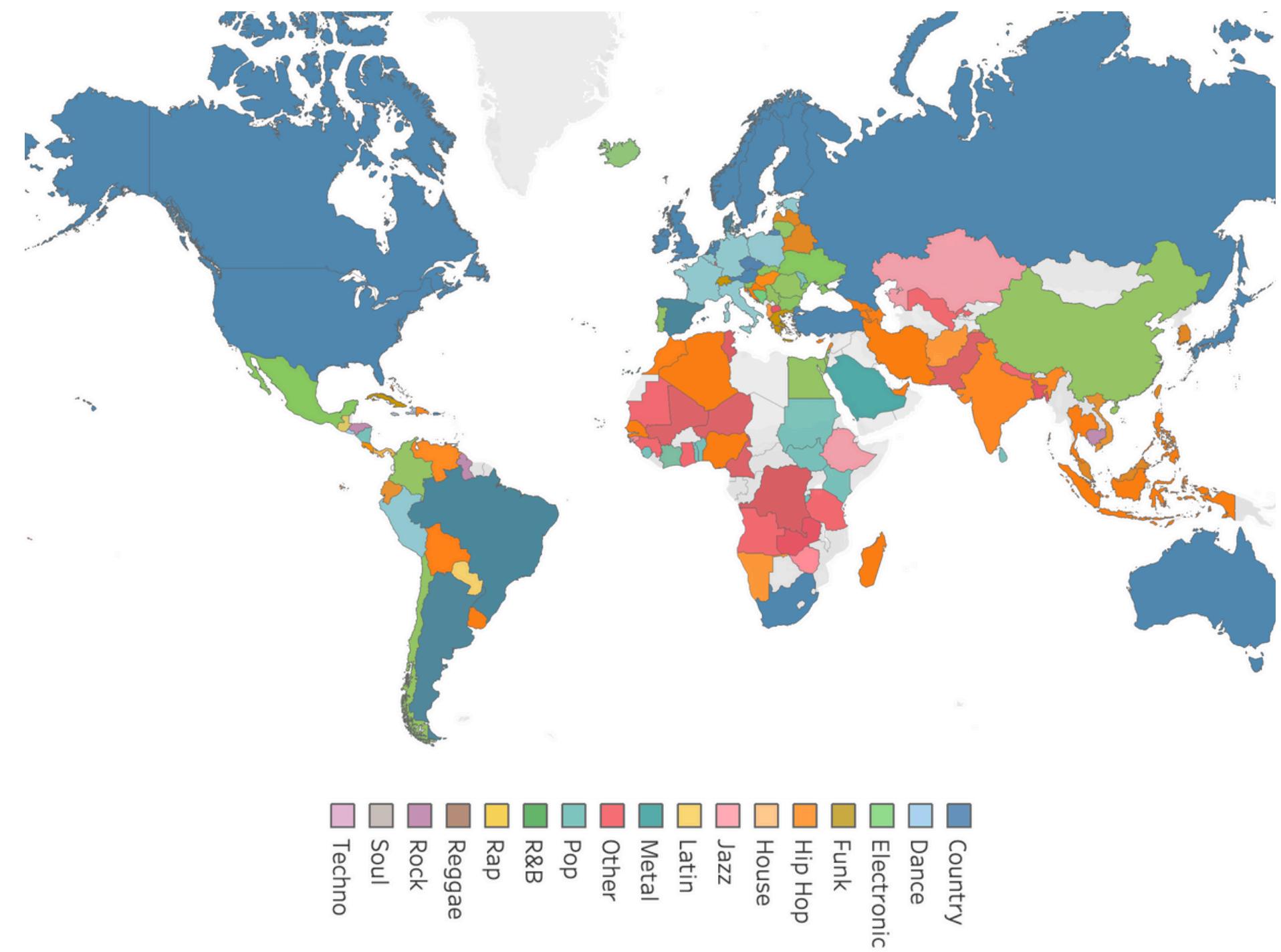


## Most Represented Genre per Country (by Artist Count)

This map displays the **most common music genre per country, based on the number of artists** associated with each genre.

- In countries like the **United States, Canada and Russia**, the **most represented genre is Country**. This doesn't mean that Country is the most popular genre in terms of audience or commercial success, but it indicates that this genre has the highest number of registered artists in the dataset.
- **Pop and Hip Hop** emerge as the **dominant genres across many countries, especially in Europe**.

This visualization reflects **artist representation, not popularity**, offering an interesting look at the diversity of genres cultivated across different countries.





Max Branca - 2061432

Thank you for  
your attention!