Artist Popularity Trajectory Using Spotify API

David Gibson

The Question

Whenever I hear people talk about music and musicians I always hear "Their older stuff is better" or "I used to listen to them back when they were good." I have heard different explanations of why a musician's or band's older music is better including "They had more time to work on their first albums" and "That was when they had more control over the production." I want to find out if a musician's older music is "better" than there newer music.

The Data

I am an avid Spotify user, so I was immediately drawn to the Spotify API. Unfortunately, there is no true measure of how good or bad a song is. Fortunately, the API provides a popularity rating for each artist, track, and album. Spotify defines popularity as a value between 0 and 100, with 100 being the most popular. The popularity is calculated by an algorithm and is based, in the most part, on the total number of plays the track has had and how recent those plays are. For this analysis, popularity is not an ideal measure because it is not a true measure of how good a song is but there is still valuable insight to be gained. Keep this in mind when moving forward

As for which artists analysis, I decided to us all the artists on the Billboards Greatest Artists of All Time. This list includes 100 artists from a variety of genres and multiple decades.

Five artists are missing from the dataset because I was unable to retrieve their albums from the API. These artists include Elvis Presley, The Rolling Stones, The Beach Boys, James Brown and JAY Z.

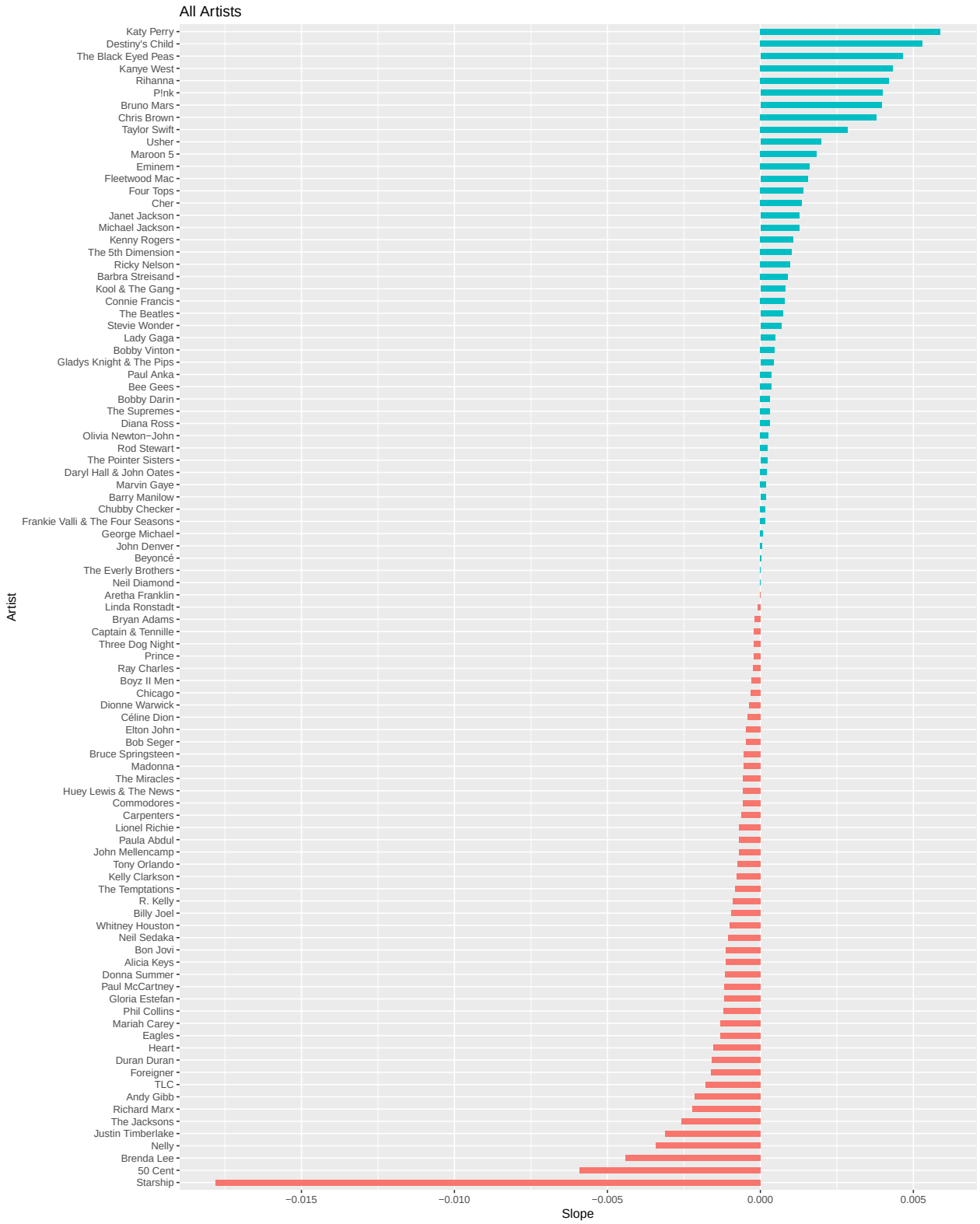
For each of the 95 artists, I pulled all their albums from Spotify. I then removed any live albums, remastered albums with a non-original release date and duplicate albums keeping the more popular album. For each artist, I also calculated their number of albums released and career length in years.

Calculating Artist Popularity Trajectory

To determine if an artist's album popularity is increasing or decreasing over time I decided to run a linear regression between release date and album popularity for each artist. From these individual models, I took the slope of each of the trendline to determine an overall increase or decrease in popularity over time. A positive slope represents an artists popularity increasing over time and their "better" music being produced later in their career. A negative slope represents an artists popularity decreasing over time and their "better" music being produced earlier in their career.

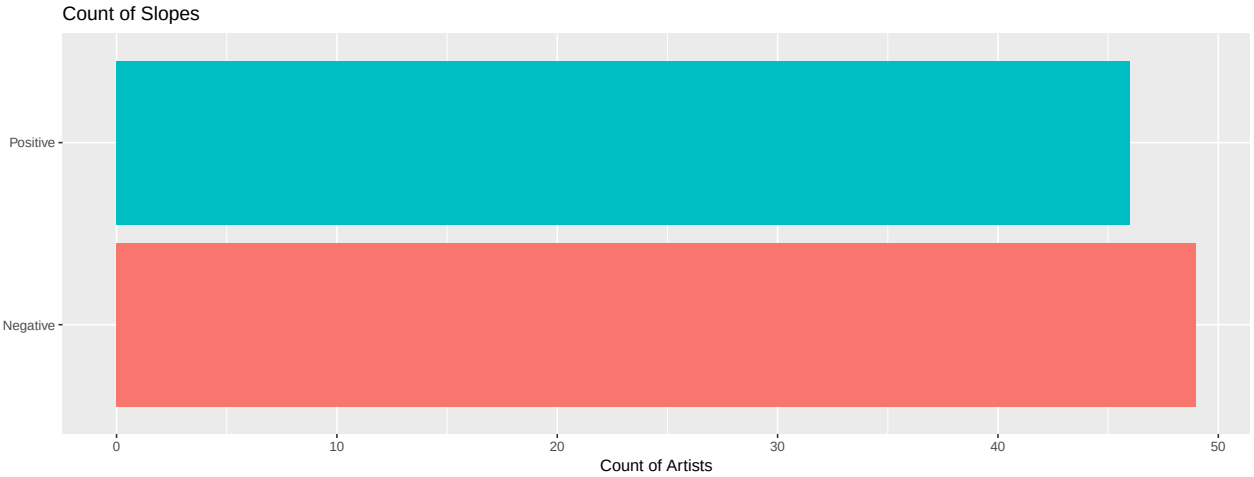
Note: Using linear regression is not ideal with time-series data, but I am currently looking into other techniques.

Below we can see all 95 artist's and their slopes.


All Artists
Artist
Katy Perry
Destiny's Child
The Black Eyed Peas
Kanye West
Rihanna
P!nk
Bruno Mars
Chris Brown
Taylor Swift
Usher
Maroon 5
Eminem
Fleetwood Mac
Four Tops
Cher
Janet Jackson
Michael Jackson
Kenny Rogers
The 5th Dimension
Ricky Nelson
Barbra Streisand
Kool & The Gang
Connie Francis
The Beatles
Stevie Wonder
Lady Gaga
Bobby Vinton
Gladys Knight & The Pips
Paul Anka
Bee Gees
Bobby Darin
The Supremes
Diana Ross
Olivia Newton−John
Rod Stewart
The Pointer Sisters
Daryl Hall & John Oates
Marvin Gaye
Barry Manilow
Chubby Checker
Frankie Valli & The Four Seasons
George Michael
John Denver
Beyoncé
The Everly Brothers
Neil Diamond
Aretha Franklin
Linda Ronstadt
Bryan Adams
Captain & Tennille
Three Dog Night
Prince
Ray Charles
Boyz II Men
Chicago
Dionne Warwick
Céline Dion
Elton John
Bob Seger
Bruce Springsteen
Madonna
The Miracles
Huey Lewis & The News
Commodores
Carpenters
Lionel Richie
Paula Abdul
John Mellencamp
Tony Orlando
Kelly Clarkson
The Temptations
R. Kelly
Billy Joel
Whitney Houston
Neil Sedaka
Bon Jovi
Alicia Keys
Donna Summer
Paul McCartney
Gloria Estefan
Phil Collins
Mariah Carey
Eagles
Heart
Duran Duran
Foreigner
TLC
Andy Gibb
Richard Marx
The Jacksons
Justin Timberlake
Nelly
Brenda Lee
50 Cent
Starship
−0.015
−0.010
−0.005
0.000
0.005
Slope

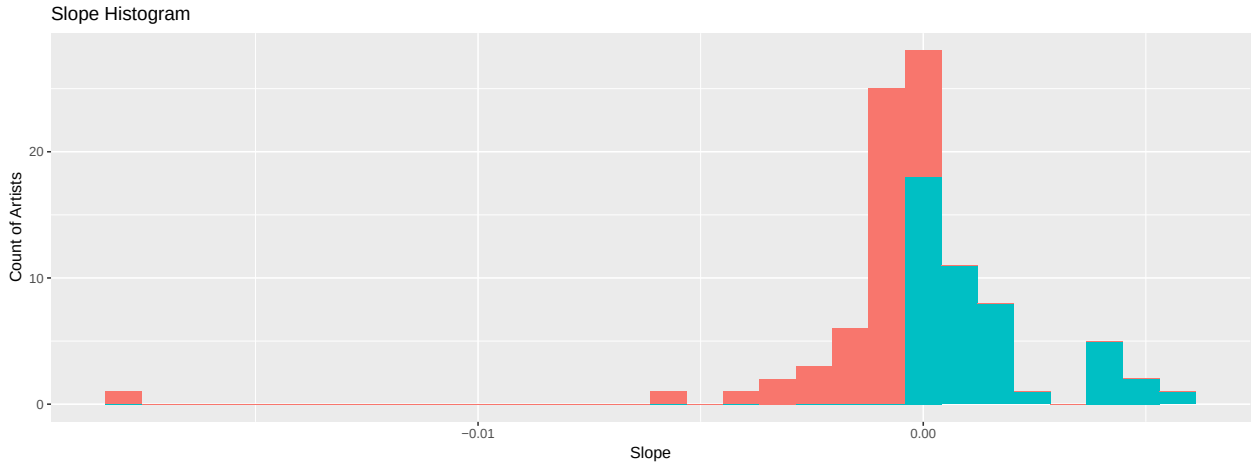
Initial Findings

The first thing to look at is to see how many negative and positive slopes there are out of the 95 artists.



As you can see the majority of artists have a negative slope meaning that most artist's more popular music is earlier in their career. Specifically, 55 out of the 95 artists have a negative slope. This would support the hypothesis I have heard about artists making "better" music earlier in their career.

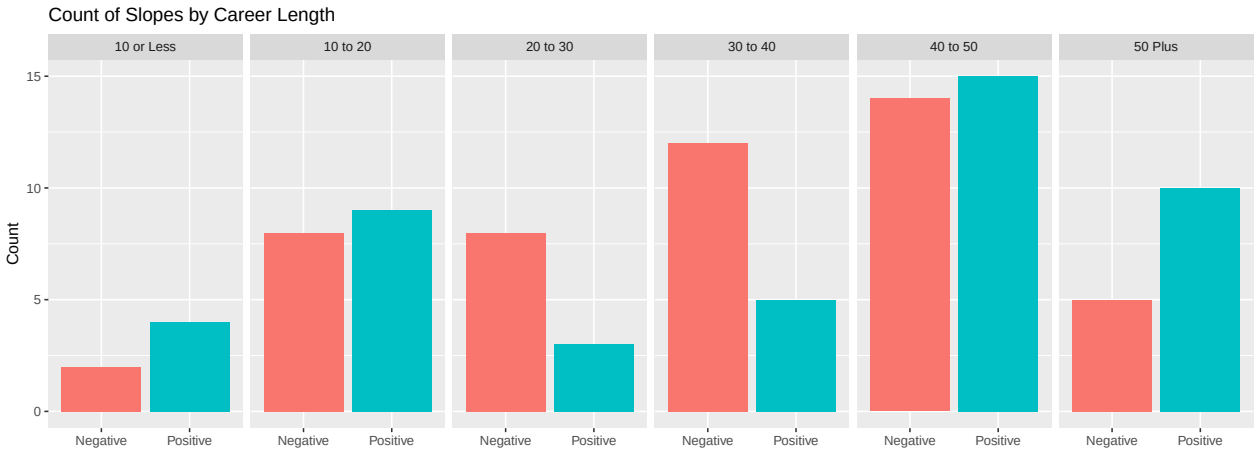
Another critical measure is to look at how distributed the slopes are. This can be achieved with a histogram.



From the histogram above we see that most of the slopes are around 0 with the majority being negative. What this tells us is that there is usually a steady increase or decrease in popularity over time. Regarding music, this may indicate a gradual change in sound or no change at all, but it is unlikely that these artists had a drastic change in sound.

Career Length

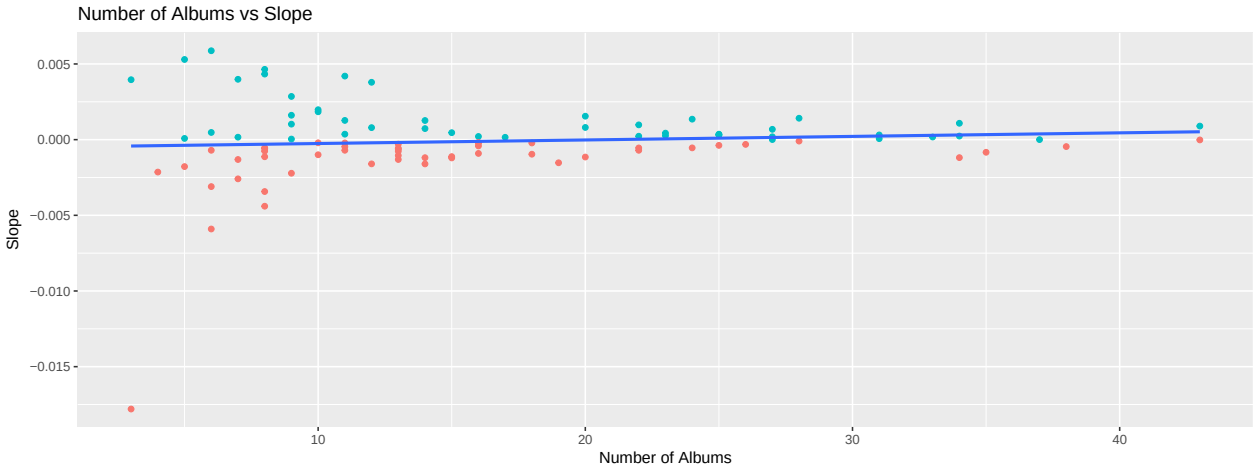
Now that we have a sense of the data lets look a little deeper into why some artists may have a negative or positive slope. First, we can look at the length of careers.



Interestingly enough artists with a length of ten or less and fifty Plus both have more positive slopes. This would make sense for artist's with less than ten years because to make the Billboard Top 100 they must have been extremely talented artists and likely became popular very quickly. As for 50 plus, this could account for loyal fans that will always listen to that one artist no matter the quality of the album. Yet, the vast majority of artists with careers from ten to fifty years have negative slopes.

Album Count

To explore the relationship between the number of albums an artist produces and the slope we can plot both features on a scatter plot.



As we can see, there is only a slight relationship between the number of albums and an artist's slope. This would make sense because the more albums an artist produces, the more likely they are to create more popular music. Additionally the more albums an artist has, the more leeway they have to put out a few less popular albums.

Conclusion

It is hard to say if an artist makes "better" music earlier or later in their career. There are so many aspects to what makes some music better than others, and even then one person's favorite song may be hated by another person. Although this analysis did not define or explore this, it did explore how the current popularity of the greatest artists differs from album to album.

Out of these 95 artists, 57% of them created their most popular music earlier in their career. The majority of the artist's slopes were relatively close to 0 meaning that there was only a slight overall change in popularity over their career. Artist's making music less than ten years are likely to have a positive slope meaning they are progressing with their music. While artists who have been making music for 10 to 50 years a much more likely to have a negative slope. Lastly, the more albums an artist produces, the more likely that artist's has made more popular music later in their career.

Moving Forward

There is so much more that can be done to answer if an artist makes "better" music earlier or later in their career. This analysis biased because the Billboard Top 100 does not acuratly represent all artists. So The immediate next steps would be to analyze more artists and compare them across different genres. Also, the Spotify API provides feature analysis of individual songs so I can see how an artist's sound has changed over their career to see if there is a correlation between a change in sound and a change in popularity. Most importantly there are more variables that need to be considered to determine better music not just the current popularity of the music. These variables may include listen count, album sales, news around album and more.

I hope you had as much fun reading my report as I did making it. I will continue to work on this project as more data becomes available. But I would love to hear any feedback or suggestions!

Thank you,

David Gibson

Contact Me

Github Repo