



The Most Popular Songs from 2010-2019

Ben Cassell & Rob Coven



Background

We got the dataset from Kaggle

The dataset was created using information from the Spotify API

The creator accessed the API through “Organize Your Music Playlist.com”

Data includes:

- 603 Total Entries
- 50 Unique Genres
- 184 Unique Artists



Overview - Columns

15 Columns

Row ID

Title of the song

Artist

Genre of the song

Year it was popular

10 Characteristics

Overview - Characteristics



BPM	Beats Per Minute - The tempo of the song
Energy	The energy of a song - the higher the value, the more energetic the song is
Danceability	The higher the value, the easier it is to dance to this song
Loudness (dB)	Volume of the song measured in Decibels
Liveness	The higher the value, the more likely the song is a live recording
Valence	The higher the value, the more positive mood for the song
Duration	The duration of the track in seconds
Acousticness	The higher the value the more acoustic the track is - less electric instrumentation
Speechiness	The higher the value the more spoken word the song contains
Popularity	How popular the song was, higher value means more popular



Changes to the dataset

While exploring the dataset we found many genres only had 1 or 2 songs associated with them

To make analyzing the data easier, whenever we grouped by genre we decided to drop any genre with less than 3 songs

30 Genres met this criteria and were removed from the dataset whenever we calculated averages for the genres



Original Questions

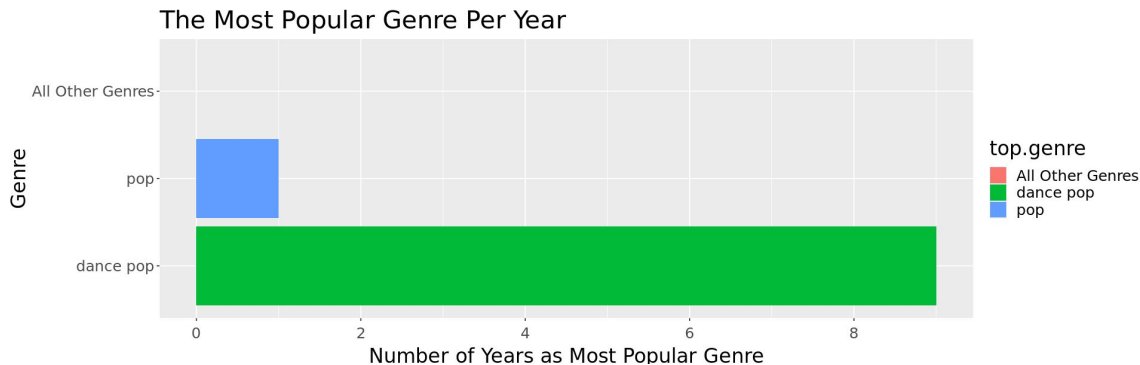
We went into the data with several questions we were hoping to answer

- What is the most popular genre of each year?
- Over the course of the years in the range is there a shift in popularity towards a specific genre?
- What is the average length (in minutes) of a popular song?
- What trait differs the most between the top songs of each year?
- What is the highest common trait among the top songs each year?

What is the most popular genres of each year?

Dance Pop was the most popular genre 9 years out of 10

It was only beaten once in 2019 by Pop



Methodology

-We determined how many times a genre appeared each year

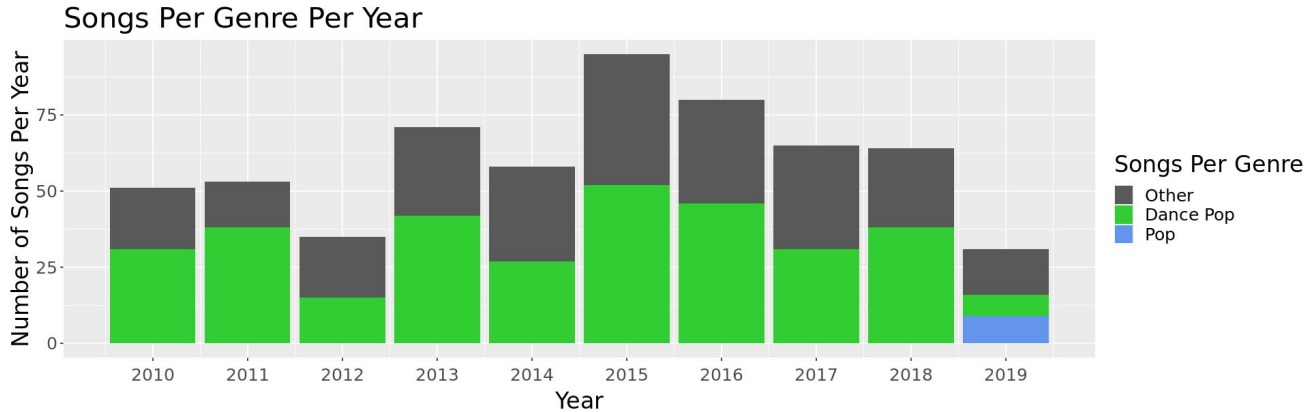
-Using that data we determined the genre that appeared the most each year

-Then we just counted the number of times each genre was the most popular genre a year

Over the course of the years, is there a shift in popularity towards a specific genre?

No; Dance Pop was consistently the most popular genre each year

Except in 2019 where it was beaten by Pop by just 2 songs



Methodology

-Counted how many times each genre appeared per year

-From this we determined what the the most popular genre was each year

3.74 Minutes

What is the average length (in minutes) of a popular song?



The Bad Questions

While analyzing the data we decided two of our original questions didn't make sense:

- What trait differs the most between the top songs of each year?
- What is the highest common trait among the top songs each year?

The issue with these questions is they assume the values of each column can be compared; we later decided they couldn't be

There were two problems: Scale and Incompatibility



New Questions

As we progressed we found new questions and concepts we wanted to explore

Do the characteristics of the songs change over the years? If so, how?

What was the distribution of artists over the genres?

Who was the most positive and negative artist based on a sentiment analysis run on song titles? How has sentiment changed over time?



Do the characteristics of the songs change over the years? If so, how?

The average of each characteristic does change significantly throughout the years

In particular:

The BPM

The Duration

The Danceability

The changes these characteristics went through were the most interesting to us

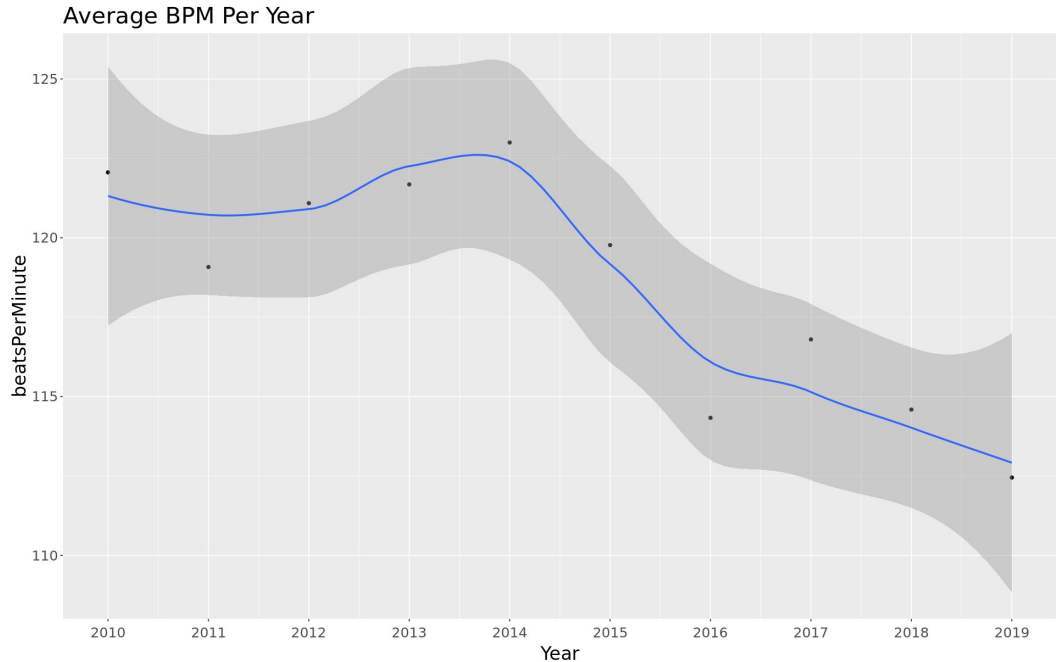


How has the average BPM changed over the years?

BPM reached a peak around 2013-2014 after it continued on a downward trend

Methodology

Taking an average BPM value each year we found that the speed of popular music has changed dramatically over the years

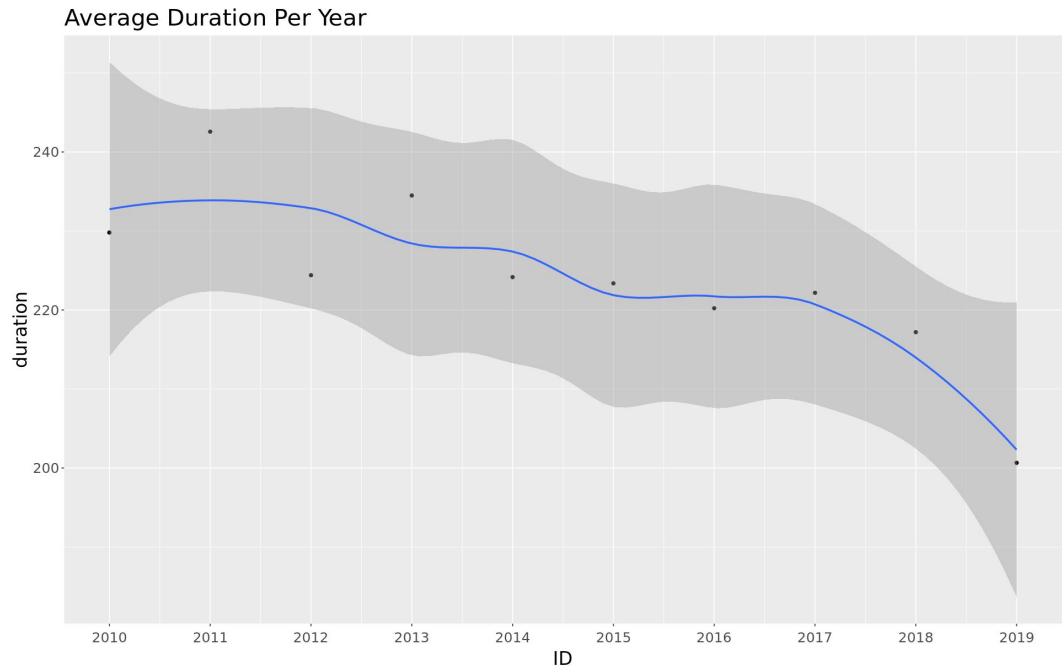


How has the average duration changed over the years?

The average duration hit a peak in 2011 and has continued in a downward trend since

Methodology

- Grouped data by year
- Calculated the average duration for each year
- Graphed with a smoothing line

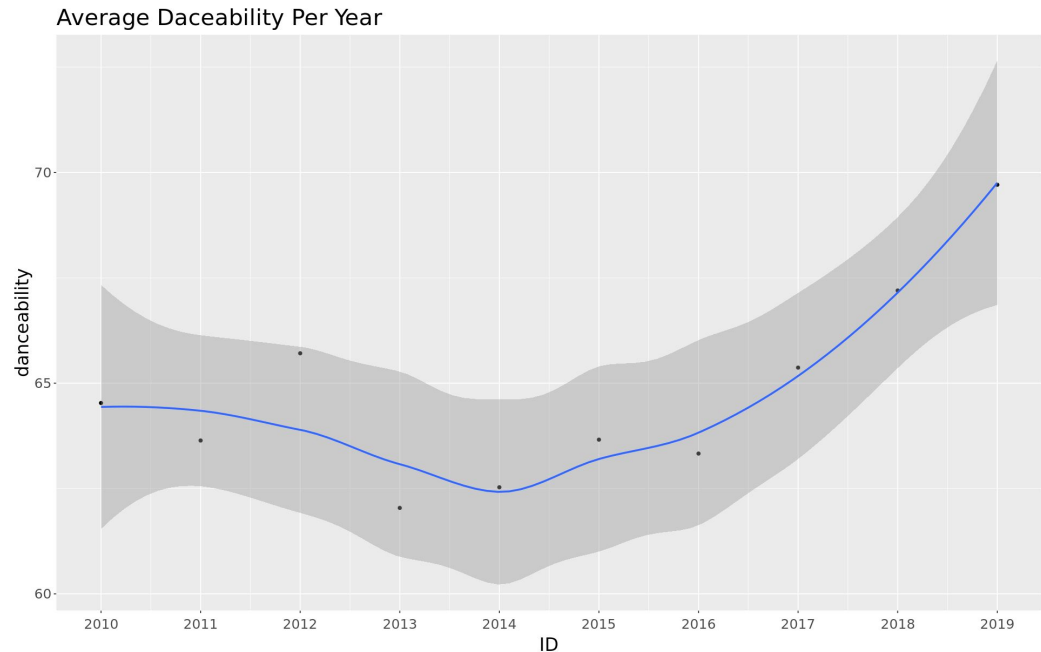


How has the average danceability changed over the years?

Danceability was at its lowest in 2013, and entered a strong upward trend after that

Methodology

- Grouped data by year
- Calculated the average danceability for each year
- Graphed with a smoothing line

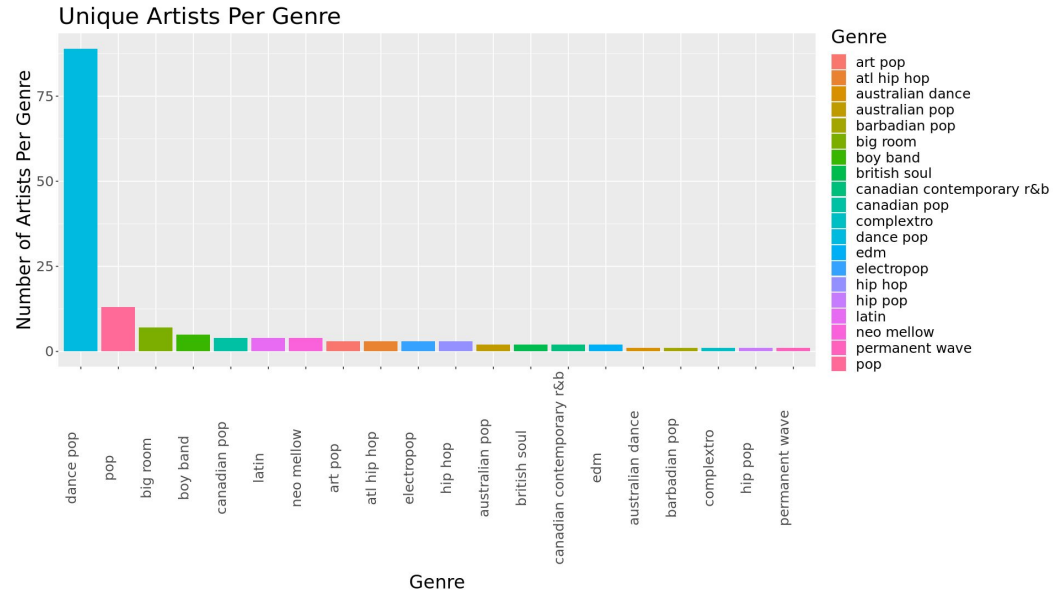


What was the distribution of artists over the genres?

The overwhelming majority of artists were in the Dance Pop category

Methodology

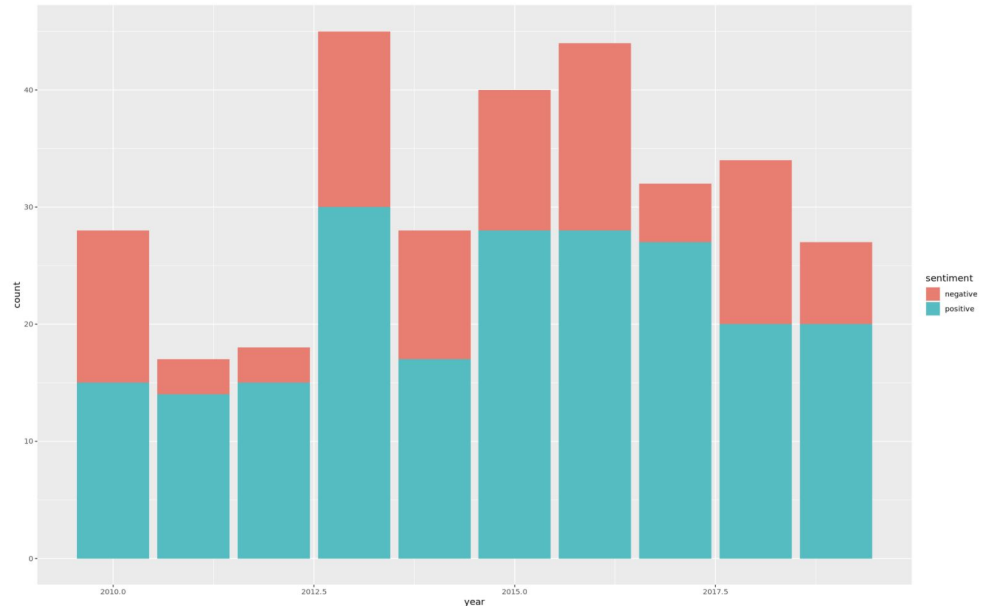
- Grouped data by genre
- Eliminated all non-unique artist values
- Counted all unique artists in each genre





Sentiment Analysis: Change Over Time

By running a sentiment analysis on song titles using the bing lexicon, we found the titles to be positive more often. Peak positivity can be found around 2013 while negativity hit a high in 2016.





Sentiment Analysis: 10 Most Positive and Negative Artists

Taking the sum sentiment value using the affin lexicon we were able to find the 10 most positive and negative artists over the past decade

artist	sum_value
<chr>	<dbl>
Kesha	-7
Fergie	-6
Britney Spears	-5
Alicia Keys	-4
Avril Lavigne	-4
David Guetta	-4
Missy Elliott	-4
Nick Jonas	-4
Christina Aguilera	-3
Justin Timberlake	-3

artist	sum_value
<chr>	<dbl>
Ed Sheeran	10
OneRepublic	10
Beyonc	9
One Direction	9
Selena Gomez & The Scene	8
Rihanna	7
Sigala	7
Bruno Mars	6
Justin Bieber	6
Lady Gaga	6



Sentiment Analysis: Sentiment Based on Genre

Grouping the sentiment counts by genre we were able to find that both the most positive and negative genre was Dance Pop. This is due to Dance Pop making up a large percentage of the data set.

top.genre	sentiment	n
<chr>	<chr>	<int>
dance pop	positive	118
dance pop	negative	57
pop	positive	25
canadian pop	positive	7
pop	negative	7
boy band	positive	6
atl hip hop	positive	5
british soul	positive	5
canadian pop	negative	5
art pop	negative	4



Conclusion

Lower BPM, Sentiment, High danceability
Why were these trends appearing?

Do current world events influence the sentiments or BPM of a song?

Could social media have influenced the average danceability of songs?

Why was Dance Pop so overwhelmingly popular?

Can we predict the popularity of a song based on its characteristics?