**Slide 1:**

For my dataset, I chose to look at top songs on Spotify from 2010 to 2019 according to Billboard. I thought it would be a fun dataset to analyze and might show some interesting information.

Some of the questions that I was looking to answer include:

* What were popular genres over the years and how have they evolved?
* Repeat artists over the years and their general progression
* The average length of popular songs per year

**Overview slide:**

With this dataset, some of the elements I was able to analyze per song included:

* Title
* Artist
* Top Genre
* Year
* Beats Per Minute
* Energy
* Danceability
* Loudness
* Liveness
* Valence
* Duration (length in sec)
* Acousticness
* Speechiness
* Popularity

**Cleaning:**

There weren’t too many items to clean with this set, however, as you can see, the column names were not appropriate descriptors of what was being measured so I updated those.

Another big thing I did was look for missing data. In the second image, you can see there each column had a 0% so there wasn’t any missing data.

Finally, I also looked for duplicates, specifically in the Title column. I was curious to see if there were multiples over the years and if this was an accident or true data to be analyzed. A couple of the songs appeared twice in the dataset. Ultimately, I don’t think those were the result of an error.

**Top Genres:**

So I started looking at the data from a big picture. The first thing I looked at Top Genres over the entire 9 year period. It looks like Dance Pop was the most prevalent, followed by Pop and Canadian Pop.

If you’re curious like about what Canadian Pop is, rest assured it’s artists like Justin Bieber and Shawn Mendes.

With this analysis I came across another burning question – What on earth is British Soul? It’s Adele. She has her own category, naturally.

**Frequency of Artists:**

The next thing I was curious about was artists – who had the most hits over the nine year period? It was hands down Katy Perry with 17 hits, followed by Justin Bieber who had 16 hits.

**Artist Frequency Per Year:**

For my next visualization, I dug a little bit deeper into the frequency of artists and look at how many hits they had per year. The ones that are highlighted green show if an artist had four or more hits in on year.

As you can see 2015 was a really good year for our Canadian Pop friend, Justin Bieber who had a total of 9 hits that year.

**Average Length of a song and BPM:**

For this one, I was curious about the average length of a song and if it changed from 2010 to 2019. One thing I noticed was that in 2010, the average top song length was 229.8 seconds, whereas, in 2019 it was 200.65. Did our attention spans decrease? It’s a curious thought. You can see a similar trend with BPM.

**Speechiness vs. Popularity:**

According to the dataset, speechiness is calculated as “The higher the value the more spoken word the song contains.” So the more words, the higher the score.

In this instance, I compared the total speechiness score to popularity to see if there was a correlation. Specifically, I looked at if a song was above average in popularity which is shown in green to songs that were below average in popularity in yellow. What I found was that there isn’t quite a correlation. The second visualization at the bottom is more direct and shows that songs with a below average popularity had an average speechiness of 8.375 and songs with an above average popularity had an average speechiness of 8.35.

**Popularity vs. BPM**

The final visualization looks at song popularity and Beats per Minute. You can see that songs with a higher popularity have a BPM cluster around 90 and 130. To be clear, I’m not saying that creating a song within this cluster of BPM will cause you to have a more popular song.

**End**

That’s all I have for you. You can take a look at my portfolio links here or browse the dataset on Kaggle.

Thank you!