This project is an analysis of my streaming history from October 21, *2021* - October, 19th *2022*. To obtain this streaming history, you may request it from Spotify under account > privacy settings, it takes about 1-5 days. I chose to do this project because music is one of my most personal hobbies and I always look forward to Spotify's yearly "Spotify Wrapped" which displays your top 5 tracks and artists of the year. You can say I grew impatient, why not create my own wrapped using *BigQuery SQL*, *Google Spreadsheets*, and *Tableau* :)

Furthermore, I also specifically focus on Kendrick Lamar's playtime from May - September. Why? His long-awaited album 'Mr. Morale & The Big Steppers', was released on May 13th, 2022. This was a big deal for me as I am a big Hip-Hop fan. So much so that I attended his concert on September 9th, 2022. For these reasons, I have specified my playtime for both May - September but also only May and September, to display how much I listened on release day and also in preparation for the concert.

In this project I accomplished

- Top 20 most streamed tracks
- Top 10 streamed artists
- Kendrick Lamar's minutes and tracks streamed for May and September specifically
- Kendrick Lamar minutes streamed from May through September
- Tableau data visualizations for each dataset

Steps I took to accomplish these tasks

1) Spotify sends your data in .json format. I converted these files (4) to .csv format to import and query more easier

2) These .csv files were imported into BigQuery for analysis. I chose BigQuery for my analysis because most of my files consisted of over 10,000 rows and BigQuery did not limit my rows as MySql Workbench did
3) As seen in my uploaded “1&2Combined_Queries.sql” file, I combined my 4 StreamingHistory files together. Furthermore, I combine my “msPlayed”, milliseconds, for each track to find my most streamed tracks.
4) BigQuery allows you to save your queried data directly to a google spreadsheet. By doing this I was able to eliminate tracks that contained no playtime, shortening my datasets.
5) As seen in my uploaded file, “top20tracks.sql”, I then import these shortened datasets to obtain my top 20 most streamed tracks. I also converted milliseconds to hours and rounded that result so that I can easier visualize this data.
6) As seen in my uploaded file, “top20artists.sql”, I repeat steps 3-5 to obtain my top 20 most streamed artists instead of tracks. In addition, I chose to only visualize my top 10 instead of 20 as it proved a more appropriate amount of artists to visualize.
7) As seen in my uploaded file, “Kendrickcombination.sql”, I specifically pull Kendrick Lamar streamed tracks that exceed 30 seconds and were streamed between May 13th, 2022, and September 9th, 2022. I also converted milliseconds to minutes to visualize this data by day.
8) Finally, I imported my queried datasets to Tableau for visualization.

Why did I choose the graphs that I did?

- For “**Top 20 tracks.png**”, I chose a Side By Side Bar Graph. This was an appropriate graph for this dataset as I had multiple tracks in my top 20 for the same artist. For example, Frank Ocean alone

had 8 tracks in my top 20. Thus, making this graph the most appropriate

- For “**Top 10 Artists.png**”, I chose a horizontal bar graph to properly display the large number of hours I streamed for each artist.
- For “**kendrick.png**”, a highlighted table proved to be the best method for illustrating my minutes streamed in the days of May (when the album was released) and September (the 9th was the concert day). I also filtered the tracks to only display tracks from his newest album and tracks in his concert setlist.
- For “**kendrickpie.png**”, I used a pie chart to display my months and minutes listened for those months. It proved appropriate as I only was visualizing 5 months. Furthermore, these months alone consisted of over 60 hours of my overall listening time for Kendrick Lamar for the year.