MACQUARIE
University
SYDNEY·AUSTRALIA

Presented to Macquarie Business School
Macquarie University
Session 2, 2023

In partial fulfillment
of the unit
In BUSA8090

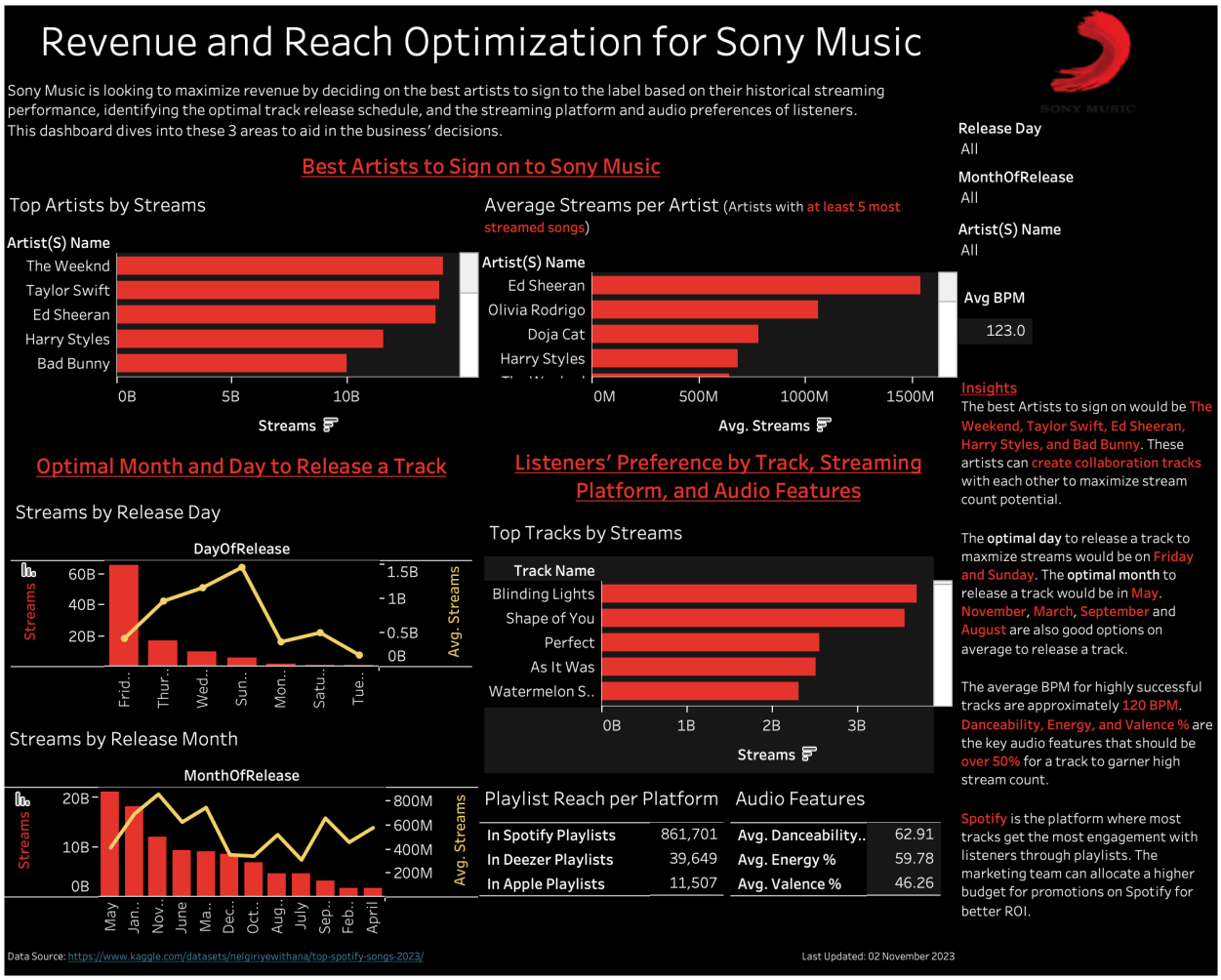
Assignment 2

Submitted by:
Beatrice Sophia Pamittan
47558865

Submitted to:
Dr. Lin Yue
03 November 2023

Part A

Please see Tableau File in the folder to interact with filters. If sort error occurs, please untick Data Interpreter.



Part B

1. **<u>Describe the four data analytics processes for data governance. List the features of R/tableau that can help you provide visualization according to data governance principles.</u>**

   The four data analytics processes for data governance are:

   1. **Establishing the "Why"**
      a. We need to know the purpose or mission for commencing a data analytics project. In this project, the purpose of the Business Analyst was to report to the Chief Marketing Officer whose main concern is to put into place great marketing strategies that would contribute to the overall business of

Sony Music. The goal in this case is to maximize revenue through streams as well as the reach for tracks released and produced by Sony Music.

2. **Defining the Data**
    a. Since we have already identified the problem, we can now identify the area of focus and determine what data you need to collect. For this project, the areas of focus identified are the artist's streaming count, the optimal release schedule, and the listeners' preferences in track, streaming platform, and audio features. With this, we can pinpoint which attributes in the data are going to be used in addressing the overall concern.
3. **Defining When to Collect the Data**
    a. As Business Analysts, we should be aware of when to collect the data as it should be done in a timely manner while still maintaining integrity between you and the stakeholder. This can be done by ensuring that the data obtained is relevant and up to date. This was ensured by identifying that the tracks located in the data set are recent and up to date in order for the CMO to make an informed decision.
4. **Proper Stakeholder Communication**
    a. Next, we should know and understand who are the stakeholders we will be communicating with regarding the data. In this case, it would be the CMO. Hence, the context of the data visualizations and insights should be relevant to their concerns. We also try to minimize jargon and try to communicate as concise as possible in order to provide instant insights to the stakeholder.

In Tableau, there is so much room for flexibility in customization of dashboards in order to comply with data governance requirements. However, filters must be done objectively and not subjectively to skew or introduce biased conclusions to the visualizations. Parameters should also be chosen and implemented in an objective manner. It is also important to ensure that the axis of the graphs presented all start at the origin to give a clear view of the data and not to manipulate what is being depicted. Another feature would be Data Profiling where you can assess the features, statistics, and distributions of the data.

2. **<u>Ethics is an important aspect of Data visualization, discuss and apply any three ethics values that you have maintained.</u>**

The three ethics values that I have maintained are (1) choosing the proper visualization to represent the data, (2) fairness, and (3) transparency. First of all, properly representing the data is a very simple, yet important way in upholding ethics in data visualization. Choosing the right plot, graph, or visualization to represent the data will give the audience a straightforward and objective look at what is trying to be conveyed. Second, fairness was applied by ensuring that everything was done from an objective perspective so that there was no bias or

discrimination in data selection or analysis. This would also ensure that there would be no information that would skew the perception or influence the decision of the audience in a certain way. It was always ensured that the origins of the axis always started at 0 for uniformity and to clearly show the scale of the data provided. Third, transparency was also an important ethics value that was upheld in this project. With transparency placed as a priority, the reliability of the information presented is upheld. This was done by clearly communicating the data source as well as brief but proper explanations for any data manipulation and methods done.

3. **Discuss Task A visualization and provide sufficient justification on the prediction.**

The dashboard created is for Sony Music's Chief Marketing Officer (CMO) and they are concerned with maximizing revenue and reach of tracks produced by Sony Music. Revenue can be maximized by maximizing the number of streams. Reach of the track, on the other hand, can be maximized by taking a look at the behavior of listeners and their preference of music streaming platforms. Hence, they need to analyze 3 different areas in order to collectively address this problem. These 3 are (1) the artist's streaming count, (2) the optimal release schedule, and (3) listeners' preferences in track, streaming platform, and audio features.

Artist's Performance based on Streaming Count

First, the CMO should be concerned with choosing the best artist to sign on to the Sony Music label. With this, we take a look at the first 2 visualizations on top. They are entitled Top Artists by Streams and Average Streams per Artist. Top Artists by Streams shows the cumulative number of streams an artist has received with their tracks that have made it to the list of Most Streamed Songs on Spotify. The artists to consider signing based on streams would be The Weeknd, Taylor Swift, Ed Sheeran, Harry Styles, Olivia Rodrigo, and Doja Cat. It can be seen that the Top Artists based on cumulative streams are not identical to the Top artists based on average streams. This implies that artists such as The Weeknd and Taylor Swift have more songs on the most streamed songs on Spotify that will

On the other hand, Average Streams per Artist shows the average number of streams an artist gains with their best performing songs. Since each artist has a different number of tracks within the dataset, getting the average would show how many streams they are able to obtain per track. It is also mentioned that the data for this visualization is filtered to contain artists only that have produced at least 5 most streamed tracks. This is mentioned for transparency as well as for the Marketing team to identify the artists who have a proven track record of those who are able to churn out tracks that garner a high number of streams. By filtering the data as such, artists who have only produced, for example, 1 very successful track or so-called "one hit wonders" throughout their career are not considered as Sony Music would be more inclined to sign artists who are reputable in releasing consistently high streaming tracks over a period of time. This is beneficial to look at especially when it comes to artists releasing albums as it would be more likely that more songs on their album would gain lots of traction and streams if they historically do well on average as compared to artists who have only 1 very successful song throughout their career.

Optimal Release Schedule

Second, we take a look at the optimal release schedule by day of the week and by month. The bar graphs show the cumulative number of streams per day and per month. While the line chart shows the average streams per day and month. The trend seen is that tracks released on Friday generally have a high number of streams. However on average, Sundays are also a good day to release as it shows the highest number of streams in this regard. For the best month to release a track to maximize streams, it would be May, November, March, August, and September based on the cumulative streams per month as well as the average number of streams per month.

Listener Behavior and Preferences

Third, we dive into the behavior and preferences of the listeners. We see a bar chart of the top streamed songs. When filtering the data on the dashboard, the user is able to get a glimpse which songs are included in the filter as well as their accumulated number of streams each. Next, we take a look at the Playlist Reach per Platform. This information is indicative of the streaming platform preference of the listeners. It implies that the more users there are on the platform, the more playlists the track is included in per platform. In addition to that, playlists are crucial in attaining exposure and reach for a track. For example, if a track is in a popular playlist, the track can potentially reach a much wider audience. Multiple playlists can also increase the track's stream count as this brings about long term exposure to other listeners as compared to charts where the rankings are very short term. In addition, the more playlists a track can be found in constitutes to more engagement between the track and the listener. Through this, there is a better chance for other listeners to discover the track and increase its reach. With this, Spotify is the streaming platform that dominates the number of playlists a track is included in. We can infer that the majority of listeners stream music on Spotify compared to other streaming platforms. Last but not least, we take a look at audio features. The top 3 features with the largest values were only selected as others presented low values. It is very likely that if the Danceability, Energy, and Valence % is above 50, the track would gain a lot of streams as this is what most people prefer.

In conclusion, in order to maximize revenue, high performing artists based on streams such as The Weeknd, Taylor Swift, Ed Sheeran, Harry Styles, Olivia Rodrigo, and Doja Cat should be considered. These artists can create collaboration tracks with each other to maximize the stream count potential of a track upon release. The optimal schedule to release a track should also be carefully chosen such as the day and the month based on historical data of streams. In this case, Next, a significant portion of the marketing budget for a release campaign can be aimed at Spotify listeners as that is where the majority of listeners and audience are in order to have a higher chance to boost streams. Lastly, danceability, energy, and valence are prominent audio features that are preferred by listeners and can very likely attract more streams.