**Song Placement on Spotify Charts**  
**Group 14- Jiya Dutta, Zach Davis, Saksham Gupta, Ethan Carter**  
**Note:** **Original group member Dylan Colon has been removed due to inactivity.**

**Project background introduction:** Top music professionals excel at analyzing data from platforms like Spotify, Shazam, and Deezer to understand the key elements behind chart-topping songs. Our research questions focus on identifying these elements (Research Questions 1-3) and exploring cross-chart rankings (Research Question 4).

**Project result highlight:** In this project, we addressed four main research questions that will be discussed in detail. The project’s crucial insights include the impact of release timing on chart performance, how audio attributes may predict chart positions, the impact of a song’s energy on chart performance and the predictive role of Deezer’s charts in comparison to Spotify. We conducted rigorous analysis with various model screening and selection procedures to test our hypotheses, eventually deriving meaningful conclusions to understand what makes songs popular.

**Project Data Introduction:** There are too many variables in the dataset to fit in one cover page, so we will only define the variables we use in the research questions here:

|  |  |
| --- | --- |
| **Feature Name** | **Feature Description** |
| **released\_year** | *Year when the song was released* |
| **released\_month** | *Month when the song was released* |
| **released\_day** | *Day of the month when the song was released* |
| **in\_spotify\_charts** | *Presence and rank of the song on Spotify charts* |
| **in\_apple\_charts** | *Presence and rank of the song on Apple Music charts* |
| **in\_deezer\_charts** | *Presence and rank of the song on Deezer charts* |
| **in\_shazam\_charts** | *Presence and rank of the song on Shazam charts* |
| **bpm** | *Beats per minute, a measure of song tempo* |
| **key** | *Key of the song* |
| **mode** | *Mode of the song (major or minor)* |
| **danceability\_%** | *Percentage indicating how suitable the song is for dancing* |
| **valence\_%** | *Positivity of the song's musical content* |
| **energy\_%** | *Perceived energy level of the song* |
| **acousticness\_%** | *Amount of acoustic sound in the song* |
| **instrumentalness\_%** | *Amount of instrumental content in the song* |

Data Source: <https://www.kaggle.com/datasets/nelgiriyewithana/top-spotify-songs-2023>   
Note: One row should be dropped before running the R code, as there is an error in the .csv file. We have uploaded the modified csv file with the final paper and code.

**Briefly describe what you learn from the project and what is the most challenging part:**  
**It is one thing to learn about a statistical method in class, and it is another thing entirely to apply what you have learned and follow the reasoning behind it.**

**In one sentence, what is your advice for the future student to deliver a high-quality project in the course:** Understand hypothesis test results for assumption checking to avoid misinterpreting the null hypothesis rejection, as some have had to correct this later.