

# Project submission 2

Niki

2023-10-27

**(1) What is the topic that you have finalized? (Answer in 1 or 2 sentences),**

The topic I've selected is regarding songs on Spotify. I'm interested in investigating what makes a song popular on the platform.

**(2) What are the data sources that you have curated so far? (Answer 1 or 2 sentences).**

Primarily, I found a data set on Kaggle that directly fits my problem statement! Link: <https://www.kaggle.com/datasets/nelgiriyeewithana/top-spotify-songs-2023>

**(3) What is the question that you are going to answer? (Answer: One sentence that ends with a question mark that could act like the title of your data story),**

What affects music popularity on the Spotify music streaming platform?

**(4) Why is this an important question? (Answer: 3 sentences, each of which has some evidence, e.g., "According to the United Nations..." to justify why the question you have chosen is important),**

First, according to a report by the International Federation of the Phonographic Industry (IFPI) in 2021, streaming now accounts for over 62% of global recorded music revenues, with Spotify being one of the major players in this industry. Understanding the factors that influence popularity on this platform can have a direct impact on the livelihoods of musicians and the overall music industry.

**(5) Which rows and columns of the dataset will be used to answer this question? (Answer: Actual names of the variables in the dataset that you plan to use). Include the challenges and errors that you faced and how you overcame them.** Mode, energy, instrumentality, and bpm! There were some missing data, hence called for a data cleaning exercise to ensure that the data set is ready to be analysed.