

Chapter 1

Dataset

The following analysis will use data from the [Tidy Tuesday Spotify Songs dataset](#). Here is the R code to import the dataset:

```
spotify_songs <- readr::read_csv('https://raw.githubusercontent.com/rfordatascience/tidytu
```

The dataset contains 32,883 songs in the genres of EDM, Latin, Pop, R&B, Rap, & Rock and 23 variables describing characteristics of the songs and the playlists they were found in. The data dictionary can be found [here](#).

The data was accessed 4/5/2023.

Research Questions

Our project will investigate the following research questions:

- Are shorter songs more likely to be popular?
- Are more intense songs more likely to be popular?

Research Hypotheses

We hypothesize that shorter songs are more likely to be popular.

We hypothesize that more intense songs are more likely to be popular.