

PRACTICAL CASE PROPOSAL

Authors:

LAST NAMES, NAMES	GROUP
Alcorcón Navarro, Eduardo	89
Pineda Fernández, Alfonso	89

Brief description of the study you propose to do:

We propose a study of different parameters of songs in spotify. We want to analyze and see the differences between the most listened and liked songs of spotify.

Data file description

Number of items in the sample: 100 used of 646+ available

Number of variables and their description: 13 variables (not everyone is going to be used)

- Main variable (principal study object)
 - o Y: weeks_on_chart (quantitative, continuous variable, represented in % over how many weeks, if the song has been on the charts for more than a year, the percentage would be higher than 100%)
- Secondary variables (3 at least)
 - o X₁: danceability : This variable represents how spotify reflect how good a song can be danced (range from 300 to 100)
 - o X₂: peak rank (quantitative, integer variable)
 - o X₃: duration (quantitative, integer variable)
 - o X₄: tempo (quantitative, integer variable)
 - o X₅: loudness (quantitative, continuous variable, from -12,5 db to -2,5)
 - o X₆: gender / band (categorical variable: Male, Female, Band, Colab: More than one artist or band)

Data source:

We have acquired this data from the web kaggle, specifically from this [link](#), plus the gender of each of the artists, or if they are a band, we have combined them into one variable, as mentioned in X_6 . We plan on using only the first 100 songs