Hello everybody, what we worked on during this semester was analysing the top 2000 songs of Spotify from 1999 to 2019

-------------------------------------------------------

In this case we were both client, meaning the ones asking the task providing the data and requesting some questions to be answered, and data scientists team, meaning those that indeed analysed and answered the research questions. The questions in this case are the following: SAY THE QUESTIONS

After having those questions answered, we needed also to propose different approaches to:

Predict a songs popularity

Predicts its explicit label

-------------------------------------------------------

The dataset was taken from a famous data collection website called Kaggle. It contains 2000 rows, each representing a hit song, and 18 different features. In the table you can see the features that, for our specific analysis, were the most used and relevant. SAY THE FEATURES

-------------------------------------------------------

The dataset was taken from a famous data collection website called Kaggle. It contains 2000 rows, each representing a hit song, and 18 different features. In the table you can see the features that, for our specific analysis, were the most used and relevant. SAY THE FEATURES

Summary statistics show the most important quartiles (1, median, 3), mean, std and min-max values for every numerical / binary column. It shows also the count of the categorical features.