# Squad - Azúcar Moreno –
# Eurovision Song Contest 2023

**Introduction**:

The purpose of this project is to determine which countries will be in the top five of the 2023 Eurovision Song Contest.

Eurovision is an annual competition where performers represent their countries by presenting different styles of songs. Given the competition's popularity and the amount of existing data, we chose this topic to apply as many of the methodologies we have learned during the masters.

**Methodology**:

*Data Acquisition:* We extracted a series of data using Python Beautiful Soup and the Google extension Table Capture. We also downloaded a dataset from Kaggle.

*Data Cleaning and Transformation:* We created specific functions and saved them in .py files to clean and transform the data, making it ready for storage.

*Storage*: We chose to store the data in a relational database like MySQL.

*Visualization*: We used Power BI to display the information interactively.

*Machine Learning*: We used three different techniques to predict the top five countries:

> Logistic Regression: We chose this classification algorithm because of its effectiveness in predicting binary outcomes.
>
> Random Forest: It's an ensemble learning algorithm that combines multiple decision trees to create a more accurate and robust model, ideal for our problem because it can handle non-linear relationships between the features and the target variable.
>
> Neural Network: It's also an ensemble learning algorithm, which can handle high-dimensional data and non-linear relationships.

**Conclusions**:

The models were able to predict the countries that will be in the top five with varying degrees of accuracy. Sweden, Malta, Denmark, Israel and Belgium were the most frequently predicted countries. Although the models were generally able to predict the top five with a high degree of precision and recall, the F1 score was relatively low. This indicates that they had difficulty balancing both precision and recall for some countries. Interestingly, Sweden appeared multiple times in the results and the model that made this prediction had a high F1 score of 0.9, suggesting that Sweden may be a strong contender for this year's Eurovision contest.

**Future Work**:

This project can be used to predict the winner in future years, including improvements to the datasets since more data will be available regarding voting patterns. Additionally, the project can be used to analyze the political influence of different participating countries at the time of voting. Finally, the data can be used to analyze the features of songs separately.

By: Jessica López, Fran Ávila, Rubén Tenreiro Assembler Institute of Technology