

# DATA SCIENCE WORKSHOP

## ASSIGNMENT - DAY 2



Kaggle Competition: [PUBG Finish Placement Prediction](#)

**Problem Statement:** You are provided with a large number of anonymized **PUBG game stats**.

Each row contains one player's post-game stats. Create a model which **predicts players' finishing placement** based on their final stats, on a scale from 1 (first place) to 0 (last place).

Perform the PUBG data analysis and submit a .ipynb file that answers the following:

- Does healing improve the chance of winning the game?

**Hint:** Use Correlation

- What is the effect of Knocking (DBNO), Assisting or Reviving on Winning Percentage?

**Hint:** Build a joint plot

- What is the effect on the MAE if we are not removing any attribute from the data? Was it a good decision?
- Use feature engineering to create new features and check the correlation with target variable
- Use highly correlated new features and calculate the accuracy of the model
- Can we improve the accuracy of the model? If yes How

**Hint:** You can use [boosting algorithm](#) and submit a .ipynb file