

## Machine Learning - Ex - Data Preparation

### Ex1

Ex1\_data1.csv & Ex1\_Data2.csv are files containing data about 40 basketball players.

You need to prepare the data for the next ML phase for predicting statistical data about basketball players:

- Gather the data
- Drop duplications & Irrelevant data
- Check for missing data
- Handle missing data (if there any)
- Get a “quick” statistics overview with “describe” function

### Ex2

There is a **red\_wine\_quality** DataSet:

- The last column is the target variable – The Quality of a red wine
- The rest of the columns are the features

Prepare the data:

- Check if there are empty values and fill them if there.
- Select the most influence features. Select the features with a correlation to the target that is higher than 0.2 (Create a correlation matrix)