***ANALYSIS STEPS :***

1- I tried to familiarise myself with the different performance variables in the database and to understand the links between them. To help me, I created an excel document to sort the metrics and contextualise them.

2- I chose data and data groupings to analyse performance on different axes. I encountered difficulties because I don't know the sport and some variables defined as the sum of two variables, were not the case. Ex: "disposal" > 0 with 1 "badPass". I deduced that the eliminations did not quantify the unsuccessful events. Therefore I created a new variable "corrected gain" which follows the basic definition. I used the regression method to show that these two variables were different, which justified the inclusion of this new variable in my analysis.

3- I organised the data in tables (pivot), more readable for my analysis and my graphs.

4- I made a scatter plot and a density plot to see how my data were organised in order to find an analysis axis. I observed different subgroups on the variables that I related. I wanted to use the Kmean clustering method to organise my data.

5- I looked for different packages and options to optimise my graphs to improve the visual aspect.

You will can find more details inside my code.