

# Portfolio Task

You will receive player datasets for the Top 5 European Leagues

You will receive team dataset for Premier League and a club

Data cleaning, manipulation and transformation

Clean the data - drop any duplicate or empty rows. It is good practice to do, even if the dataset is clean or trusted!

Convert any stats to percentiles

Concatenate the datasets together into a main file

**Bonus -**  
Could you loop through each row and change the positions to DEF, MID, WING, ST (e.g. change a position beginning with RCB to DEF)

Data visualisation, football knowledge

Perform exploratory data analysis - try to find statistically significant stats (think R squared, P squared values, compare against xG or xG against - depending on whether it's in possession or out of possession) that the club are underperforming in

Using your team insights - identify players that could fix the problems (e.g. if a team is struggling to progress to final 3rd, find a midfielder with high progressive passing)

Visualise team insights in Tableau and add to portfolio

Visualise your chosen players using pizza or radar plots (Python generated) - selecting relevant position specific stats, add to portfolio

**Bonus -**  
Could you improve the dataset with calculated fields e.g. expected points, % of recoveries in certain zones etc, final third entries