# Project timeline

## Files

2 excel files and one power bi were used for this project. The 1st excel file (League points across top 5 leagues) was used to import league table data across the top 5 leagues from the 2015/2016 season till the 2024/2025 season with the data obtained used to find the clubs that remained in their respective league for 10 years, the points obtained each season and the position they finished on each table.

The 2nd excel file (Non-relegated teams cost of success) was used to tabulate across the 10 seasons the wage bill, transfer fee, points and position finished across 10 years in each league (The sheets are hidden but they are named after each league if you want to view them). The respective wage bill, transfer fee, points and position where then all appended into 4 respective tables which forms the base of the data used in power bi.

Lastly the power bi file was used to adjust the data for analysis, create measures for calculations, adjust relationship models for effective filtering, plot visuals , add respective slicers and edit their interaction with the visuals.

## Detailed steps

### First excel file

* Used “Get data “ > “Other sources” > “From web” to input league tables for each season across the top 5 leagues placing them in their respective sheets. Source used was transfermarket
* While inputting the table data, the data needed to be transformed due to empty columns, and missing headers hence power query was used.
* After inputting all the tables across the top 5 leagues a list of all the clubs that remained across the 10 years of their respective leagues needed to be obtained. This was accomplished by first appending all tables across the years for each respective leagues and using the COUNTIF function to find out how many times the club appeared in the appended table and If they appeared 10 times then they were present across the 10 years they fit the criteria. The FILTER function was then used to extract those club names using that criteria and the UNIQUE function to remove the duplicates and return a distinct list.

### Second excel file

* After getting the distinct list of the clubs in the 1st excel file, the list was copied into the 2nd excel document and used to create 4 tables for each league which consisted of the wage bill, transfer fee, points obtained and position finished across the 10 seasons
* Both the wage bill(data obtained from capalogy) and transfer fee data was obtained manually while the position and points were obtained from the league tables in the 1st excel file using XLOOKUP function
* After completing each tables across the leagues they were all appended using power query based on wage bill, transfer fee, position and points. A new column stating the league of the clubs was added to the original tables and reflected on the newly appended tables.
* 2 new tables consists of the clubs and leagues respectively where created for relationship purposes whenused in power bi (They are in the league and club tabloe sheet)

### Power bi file

* Used “Get data” > “Excel workbook” to import the 4 appended table, league table and club table data from the 2nd excel file.
* The 4 appended table were unpivoted in power query to take it easier for analysis
* Created a query table consisting of the seasons for filtering
* Created a club and league table for relationship purposes (Linked to the league table in a many-to one relationship) and the 4 appended tables in a one-to-many relationship
* Created measures to calculate the total cost and cost per point (can be found in the club and league table)
* Added league and season slicers using the league and season table respectively
* Added pie chart showing total cost across league (Editted interaction so it does not work with league slicer)
* Created table consisting of clubs, sum of wage bill, sum of transfer fee, sum of points, total cost, cost per point and average position (works with both slicers)
* Added line chart showing average position for clubs.