Resources for Shaping Data & Visualisation (Reactor Event Sep 2021)

Here are the resources for attendees for the Reactor event in September 2021 and a brief explanation of the dataset. This tutorial case study uses Power BI's Query Editor to transform data about football English Premier League (EPL).

The source data is the EPL match record for the last (20-21) season. This is available at [CourseDatasets/EPL Data 20-21.csv at main · MarkWilcock/CourseDatasets · GitHub](https://github.com)

To import the data directly into Power BI, use the URL that returns the raw data at [https://raw.githubusercontent.com/MarkWilcock/CourseDatasets/main/Football/EPL Data 20-21.csv](https://raw.githubusercontent.com/MarkWilcock/CourseDatasets/main/Football/EPL%20Data%2020-21.csv)

The original data is from [England Football Results Betting Odds | Premiership Results & Betting Odds (football-data.co.uk)](https://football-data.co.uk)

You will need to download Power BI Desktop (free) from the Microsoft Store or from [Download Microsoft Power BI Desktop from Official Microsoft Download Center](https://microsoft.com)

There is a Power BI Desktop file, EPL Analysis.PBIX, that contains a worked example of the session.

In this session we transform the data into the league table results. This is a summary and ranking of each team's performance. You can use this to reconcile your results -

[Premier League Table, Form Guide & Season Archives](https://example.com) and look at the 20-21 season

Background to the case study and explanation of terms

The English Premier League Table is a summary report of the relative performance of about 20 football teams during a football season. This is what the columns mean

- Team – there are typically about 20 teams in a league
- Played – the number of matches played by the team
- Won – the number of matches won by the team
- Drawn – the number of matches drawn by the team
- Lost – the number of matches lost by the team
- For - the number of goals scored by the team during the season
- Against – the number of goals scored against the team scored during the season
- GD – the goal difference, equal to *For* less *Against*. Note that this can be negative.
- Points - the number of points accumulated by the team. A team gains 3 points if they win a match, 1 if they draw and none if they lose.

The source data for the report, the detailed operational data, are the match results. During a season, each team in the league will play every other team twice, once at their home ground and once at their opponent's home ground.

The essential columns in the Match table are the

- home team, the team playing at its own stadium

- away team, the visiting team
- full time home goals (FTHG), those scored by the home team during the match
- full time away goals (FTAG), those scored by the away team during the match

Using FTHG and FTAG, we can determine the winning team and the points awarded to each team for the match.

Note that each match may generate points for both the home team and the away team, and we must build a table of match results based on the match table. For example, let's assume that Chelsea have already played Arsenal both home and away with these results

Date	Home Team	Away Team	FTHG	FTAG
1-Sep	Chelsea	Arsenal	3	2
1-Oct	Arsenal	Chelsea	4	4

The Match Results table would look like

Date	Team	For	Against	Points
1-Sep	Chelsea	3	2	3
1-Sep	Arsenal	2	3	0
1-Oct	Chelsea	4	4	1
1-Oct	Arsenal	4	4	1

Here is a description of the columns in the source data - we only need the column names in bold.

Field Name	Description
Date	Match Date
HomeTeam	Home Team
AwayTeam	Away Team
FTHG	Full Time Home Team Goals
FTAG	Full Time Away Team Goals
FTR	Full Time Result (H=Home Win, D=Draw, A=Away Win)
HTHG	Half Time Home Team Goals
HTAG	Half Time Away Team Goals
HTR	Half Time Result (H=Home Win, D=Draw, A=Away Win)
Referee	Match Referee
HS	Home Team Shots
AS	Away Team Shots
HST	Home Team Shots on Target
AST	Away Team Shots on Target
HF	Home Team Fouls Committed
AF	Away Team Fouls Committed
HC	Home Team Corners
AC	Away Team Corners
HY	Home Team Yellow Cards
AY	Away Team Yellow Cards
HR	Home Team Red Cards
AR	Away Team Red Cards