



# EDD WEBSTER

## Football Data Analytics Portfolio

---

February 2024



# PORTFOLIO OVERVIEW

01

This portfolio aims to demonstrate the professional data science experience that I have gained over the last five years, developed from working in an elite, first team football environment, as well as several best-in-class organisations.

02

The decks demonstrates a broad range of data insights skills, ranging from data processing, modelling, visualisation and results communication.

03

A strategic approach to dissecting and addressing examples of difficult industry issues are discussed in detail. For each challenge, an approach and a summarised solution is provided with some real-life examples.

04

Finally, this portfolio aims to showcase the measurable impacts of the challenges tackled, as well as provide opportunity to demonstrate further work and skills.

# Contents

- **Introduction** (slides 4-8)
- **Areas of Impact** (slides 9-10)
  - 1) **Football Intelligence Reporting** (slides 11-12)
  - 2) **Insight Delivery to Football Stakeholders** (slides 13-16)
  - 3) **Internal Scouting and Recruitment Player Rating System** (slides 17-21)
  - 4) **Data and Infrastructure Engineering** (slides 22-32)
  - 5) **Open Source Contributions** (slides 33-34)
- **Summary** (slides 35-37)

All publicly available code: [github.com/eddwebster/football\\_analytics](https://github.com/eddwebster/football_analytics) and Tableau visualisations: [public.tableau.com/profile/edd.webster](https://public.tableau.com/profile/edd.webster).



# INTRODUCTION

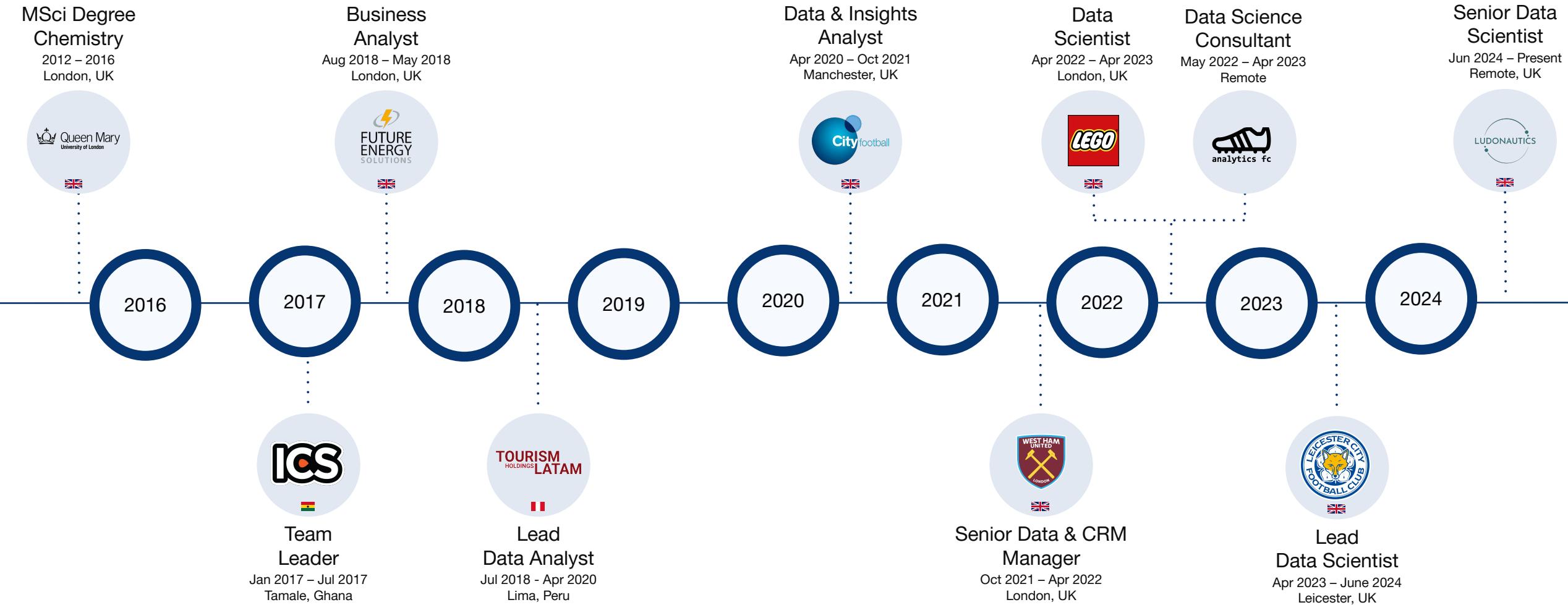


# The reason for producing this football data analytics portfolio?

To demonstrate my broad range of data science and data engineering skills, as well as my experience and knowledge of football analytics concepts and analytical techniques.



# Timeline of My Background and Education Over the Last Ten Years



For more information, see: [linkedin.com/in/eddwebster](https://linkedin.com/in/eddwebster) and my CV: [eddwebster.com/downloads/EddWebsterCV.pdf](http://eddwebster.com/downloads/EddWebsterCV.pdf)

# Objectives of the Data Insights Department of a Football Club

- 1 To provide access to data for the club, to enable every decision made by football operations to be objective and informed by data.
- 2 Educate decision makers and empower them to want to use data.
- 3 To maximise performance and subsequently, **win more games.**

# Areas of a Football Club that Data Insights Can Have an Impact

1

**Recruitment and scouting** – using models to bring a positive impact on player trading outcomes in the transfer market;

2

**Performance evaluation and player development** – measure the performance of players through models;

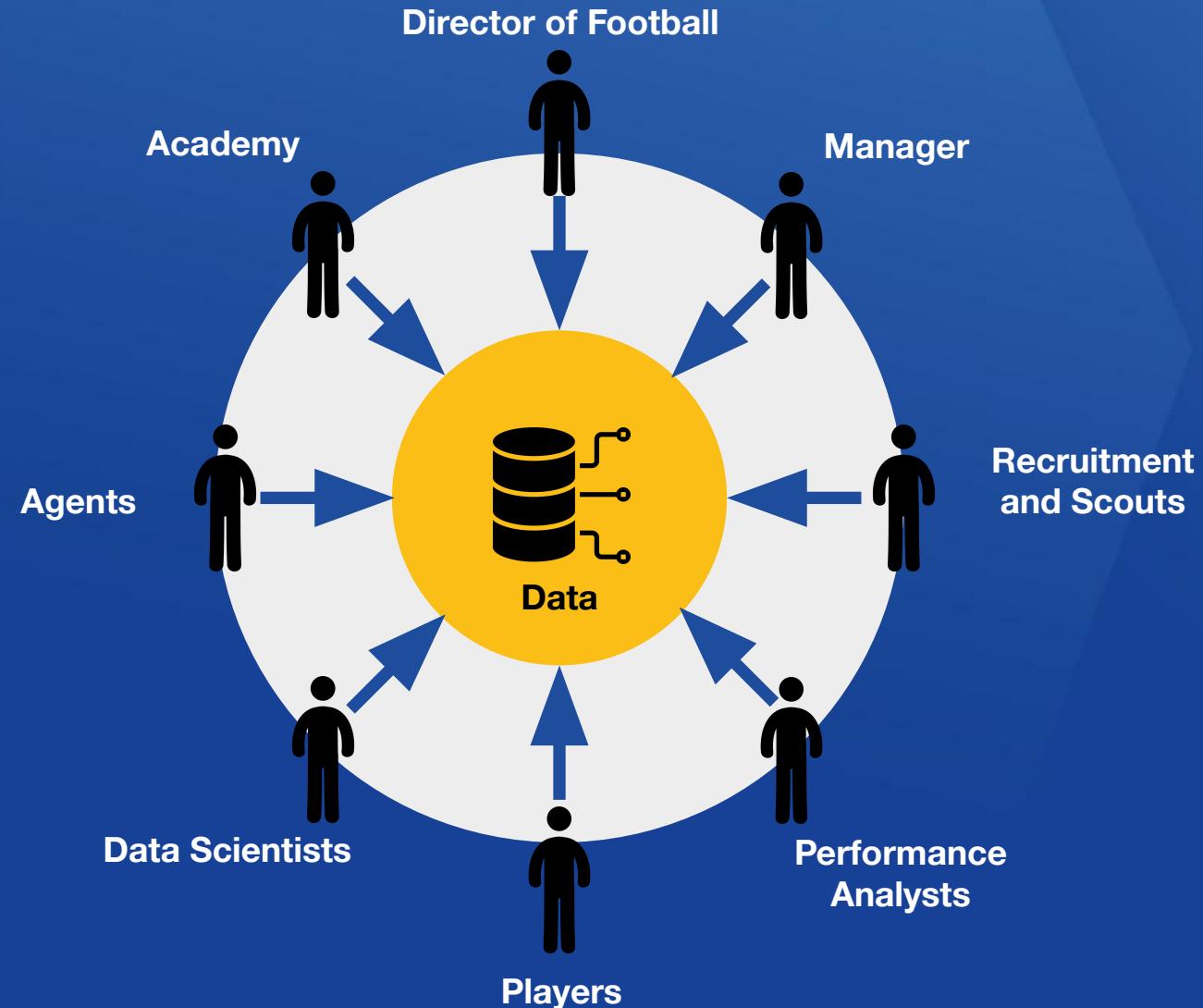
3

**Sport science and athlete monitoring** – provide reviews, audits, and recommendations for interventions; and

4

**Club strategy** – discovery of new insights through physics-based modelling that can be implemented by coaches in team strategy and affect performances on the pitch.

# Data-Driven Decision Taking Place Across a Football Club Club



# Areas of Expertise Required in a Functioning Data Analytics Team

Data  
Engineer



Data  
Scientist



Data Insights  
Analyst



Performance  
Analyst



Technical  
Expert



Infrastructure

Organise

Support

Connect

Enrich

Explore

Discover

Predict

Investigate

Inform

Visualise

Advise

Bridge

Connect

Align

Insights

Observe

Interpret

Context

Implement

Slide recreated from a slide by Tom Garner in his presentation: "Delivering Football Performance Insights" at the StatsBomb Conference 2023: [https://www.youtube.com/watch?v=ApWqVlz-ztQ&ab\\_channel=StatsBomb](https://www.youtube.com/watch?v=ApWqVlz-ztQ&ab_channel=StatsBomb). Slide can be found at around 4m30s

# AREAS OF IMPACT



# Application of Data Across Different Workstreams

## Football Intelligence Reporting



Automated reporting for opposition and post-match analysis, implementing bespoke KPIs and modelled metrics, curated to the football club's game model. Custom reports for the purposes of recruitment, performance analysis and sports science.

## Insight Delivery to Football Stakeholders



The application of data to make evidence-based decisions, that are delivered efficiently and effectively to seniors stakeholders and football practitioners.

## Internal Scouting Player Rating System



Adding context to the scouting and recruitment process with data via a bespoke, player rating system, based on a system of defined and weighted metrics, custom-made to the needs of the First Team and the game model.

## Data and Infrastructure Engineering



Creation of a back-end infrastructure, to enable automated reporting, player mapping, and predictive modelling, including the creation of the 'Player 360' reference tables.

## Open Source Contributions



Open source projects using publicly available event and tracking data. Projects include working with tracking data including Pitch Control modelling, Chance Quality modelling of shots from Event data, and determining ball-playing centre backs through clustering, to determine the "the next Gerard Piqué



# FOOTBALL INTELLIGENCE REPORTING



# Turning Big Data into Meaningful Insights Using Tableau



1

**Simplicity** - deliver clear and concise messages that can be comprehended quickly and concisely.

2

**Enable Control** - dashboard and visualisations need to include filters, tool tips, and actions, that can allow the user (often coaches) to find the insights themselves without the need for programming or dashboard building skills.

3

**Flexibility** - filters, tooltips and actions can allow the user to find the insights themselves.



# A Selection of Football Insight Analysis Tools 2/2

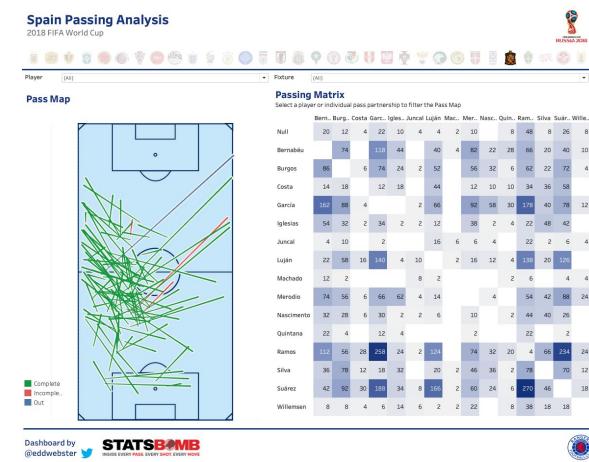
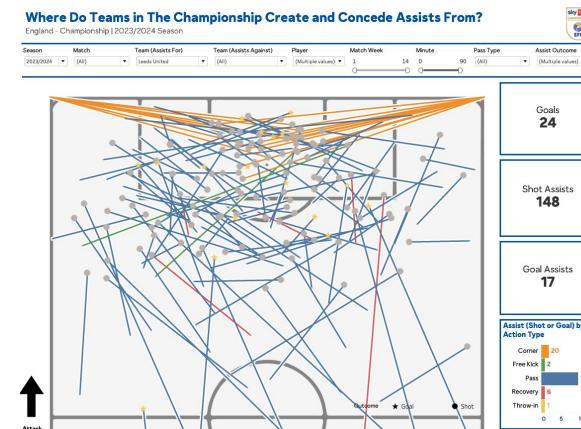
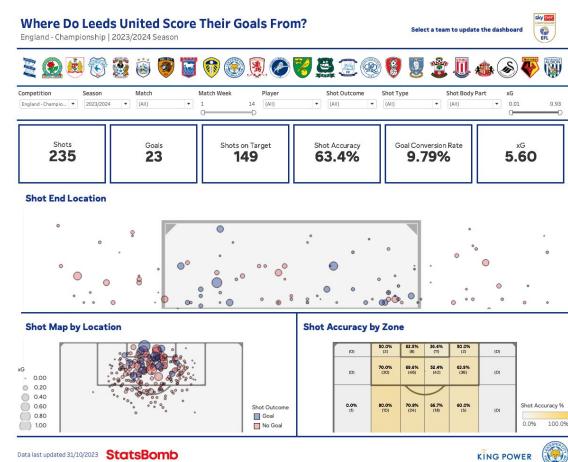
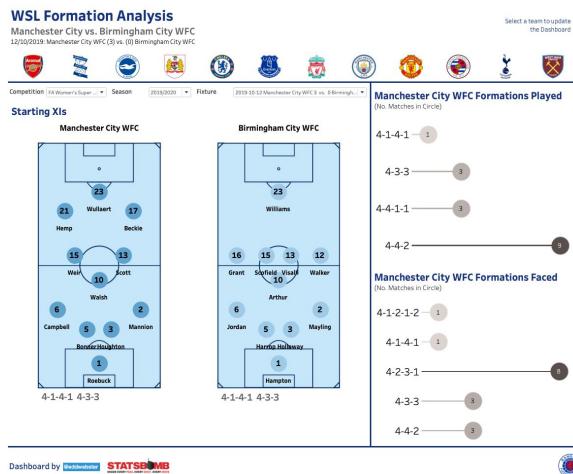
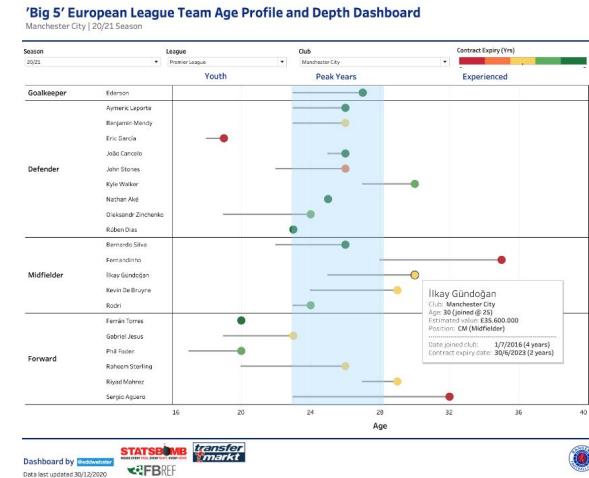
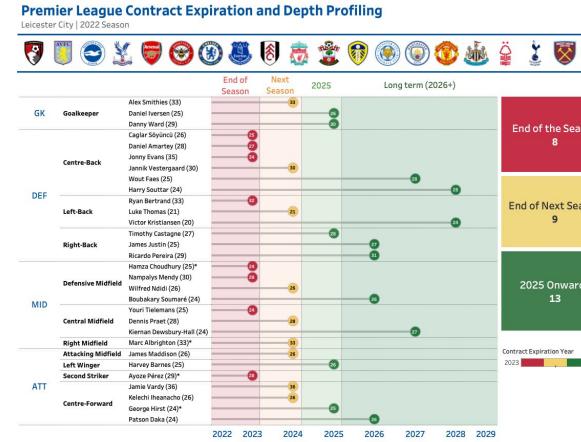
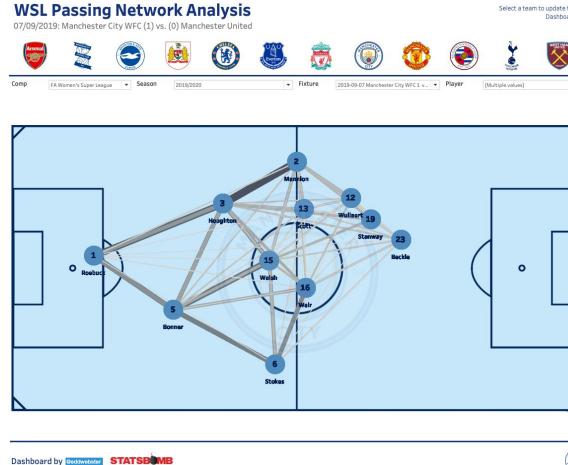
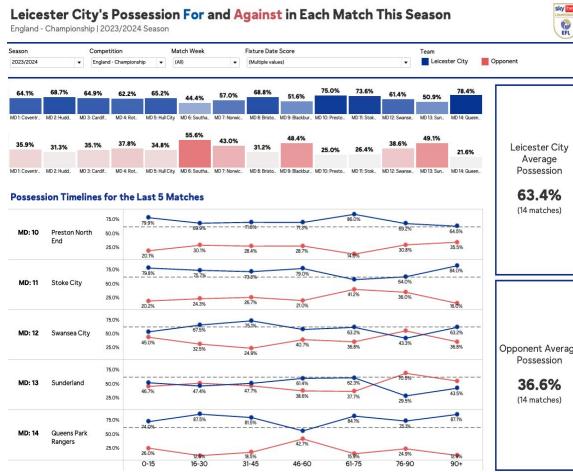


Tableau visualisations created with publicly available data can be found at my Tableau Public profile: [public.tableau.com/profile/edd.webster](http://public.tableau.com/profile/edd.webster).



# INSIGHT DELIVERY TO FOOTBALL STAKEHOLDERS



# Providing Post-Match and Opposition Analysis to Key Stakeholders



**Opposition Analysis Pre-Match Report**  
English Championship 2023/24

Plymouth Argyle  
09 December 2023

Leicester City v Plymouth Argyle

**Team Attacking Metrics Summary**

Team	Goals	Shots on Target	xG	PPDA
Brentford	12	12	1.2	1.2
Sheffield United	11	11	1.1	1.1
Southampton	10	10	1.0	1.0
Leicester City	9	9	0.9	0.9
Wolverhampton Wanderers	8	8	0.8	0.8
Nottingham Forest	7	7	0.7	0.7
West Bromwich Albion	6	6	0.6	0.6
Cardiff City	5	5	0.5	0.5
Sheffield Wednesday	4	4	0.4	0.4
Millwall	3	3	0.3	0.3
Derby County	2	2	0.2	0.2
Plymouth Argyle	1	1	0.1	0.1
Stoke City	0	0	0.0	0.0

**Plymouth Argyle Will Look To Create Chances on the Counter Attack**

**Opponent Minutes Played**

**xG Trends**

**Team Form**

**Post-Match Report**  
28 October 2023

Queens Park Rangers 1 - 2 Leicester City

English Championship - MATRADE Loftus Road


**Stats Summary**

Match Stats	Opponent
Goals	Shrewsbury Town
Shots on Target	15
Expected Goals (xG)	2.84
Non-penalty Expected Goals	2.84
Penalty Area Touches	30
Passes into Penalty Area	8
Possession Percentage	44.38

**Match Summary**

**Pressing**

**Balls Lost and Won by Zone**

**Opposition Analysis Pre-Match Report**  
English Championship 2023/24 Season

Plymouth Argyle (H) - Pre-Match Report | 09 December 2023

**Where Do Millwall Score Their Goals From?**

Location	Season	Home	Away
Opposition Box	2022/23	242	149
Opposition Box	2023/24	242	149
Opposition Box	Total	484	398
Opposition Box	Opposition Box	61.6%	61.6%
Opposition Box	Opposition Box	21	21
Opposition Box	Opposition Box	8.7%	8.7%
Opposition Box	Opposition Box	20.45	20.45

**Shot End Location**

**Shot Map by Location**

**Shot Accuracy by Zone**

**Opposition Minutes Played**

**xG Timeline**

**Match xG**

**Opponent Possession For and Against in Each Match This Season**

**Match Summary**

**Attacking League Ranking**

**Wout Fees**

# Delivering Insights Directly to the Club's Senior Stakeholders



Rotherham Play Better Against the Big Teams

GW	Opponent	Position at time of match	Date	Score	Result
2	Blackburn Rovers	7th	12th August 2023	2-2	Draw
5	Norwich City	5th	2nd September 2023	2-1	Win
8	Preston North End	3rd	23rd September	1-1	Draw
11	Southampton	10th	7th October 2023	1-1	Draw
12	Ipswich Town	2nd	7th November 2023	2-2	Draw
17	Leeds United	3rd	24th November 2023	1-1	Draw



Championship Team Metrics Comparison (Home Matches MD 1-19)



At Home, Plymouth Argyle's are much more effective in front of goal, featuring in the **top 20%** for Shots On Target p90 (10.7) and Goals p90 (2.26). They still however conceded a lot of chances, reflecting in the high values for Shots, Shots On Target, xG, Touches in the Box, Passes in the Box and OBV Conceded.

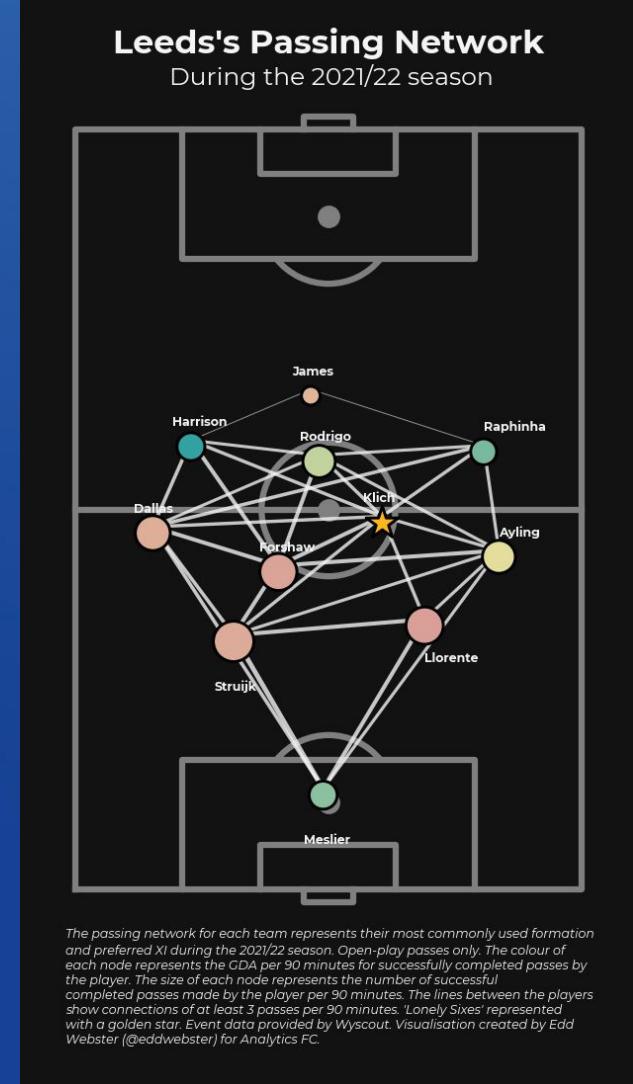
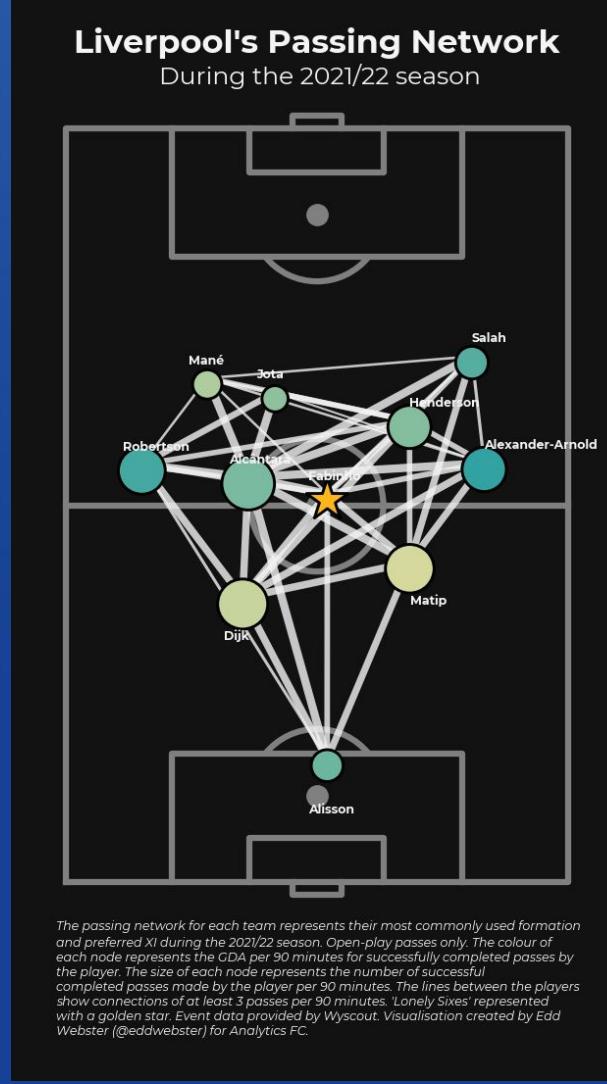
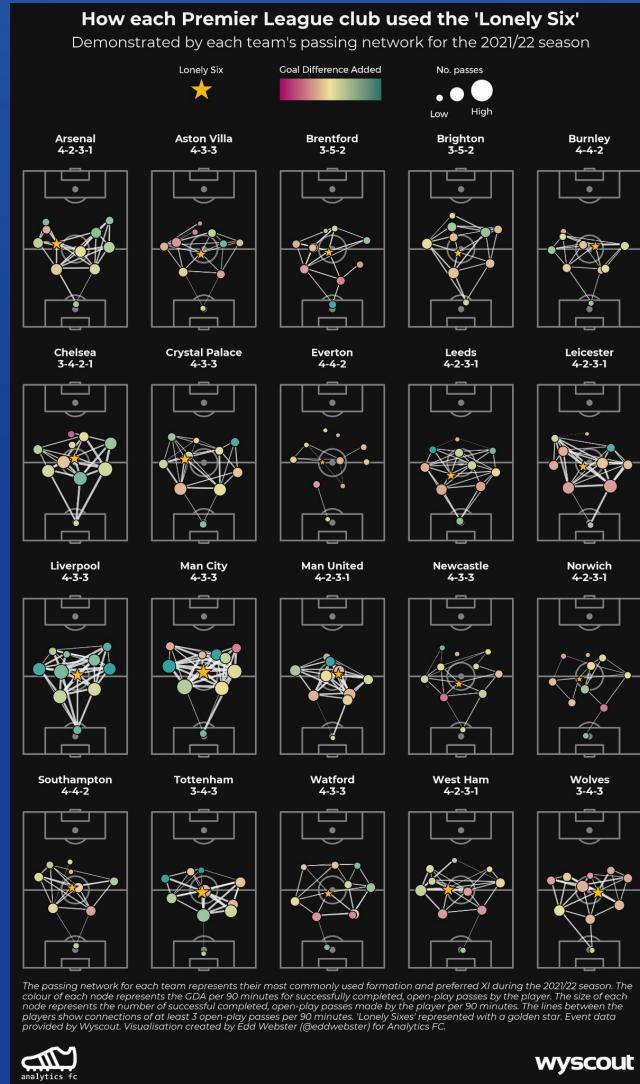
StatsBomb

# Delivering Insights as Part of Analytics FC Coach ID Service





# Tailor-made Analysis for Online Publication





# PLAYER RATINGS FRAMEWORK





# The Purpose of Data Analytics for Recruitment and Scouting

- 1 To narrow a whole cohort of players to a manageable number of targets for further investigation by domain experts, to optimise the resources available to the department.
- 2 Provide an objective assessment of player's in-game performance for both their output and playing style.
- 3 Empower decision makers that want to use the data to further inform their decision making and increase the likelihood of success for the purposes of player recruitment.



# How do you measure...



Shooting?  
Scoring?  
End Product?  
Finishing?

- Goals
- Shots
- Shots on Target
- Shot Conversion
- Shot on Target Conversion
- Pressured Shot
- Range
- xG



# PROSPECTS Rating System

- **PROSPECTS** ratings: Performance Ratings for Objective Scouting of Players to Evaluate Contribution and Technical Skill.
- A rating system designed with Leicester City's game model in mind, bespoke to the qualities of players that football practitioners at the clubs such as the manager, coaching staff, and scouting and recruitment are looking for.
- The PROSPECTS Data-Driven Player Rating System aims to provide accurate measurements of performance for a player in a match based on their in-game event data (passing, shooting, duels, pressing, etc.). These ratings can then be aggregated and determined for a specific period or whole season across a competition.
- The ratings framework takes into account not just player output such as Value Framework outputs (xT, OBV, etc.) , xG, or success percentages (goal conversion, pass completion), it is also used to evaluated playing style.
- The PROSPECTS Player Ratings takes into account over 200 different metrics derived from event data, with the final score being devised from six Categories, corresponding to different Technical Qualities for a player's performance specific to the Leicester City game model. The six sub-scores are:
  - Scoring
  - Passing
  - Possession
  - Defending
  - Creating / Attacking
  - Goalkeeping
- These six categories are then broken down further into Technical Qualities that are discussed further on the following slides.
- Using each of the six categories, the Overall Ratings is calculated according to the player's position in the match. This ensures that ratings are distributed fairly across each of the categories regardless of position. For example, a defender would not be expected to perform as well in the attacking category as a centre-forward, and therefore the weighting applied to defenders and centre-forwards for those actions are different.
- These raw scores are then smoothed out to measure how close they are to the average (Z-scores) and then converted into an intuitive rating of 1-100 to provide a final rating.



# PROSPECTS Rating Proposed Categories and Respective Qualities



## Scoring



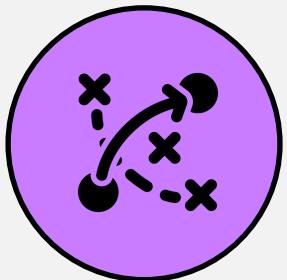
Goals  
Shots  
Conversion  
Position

## Creating



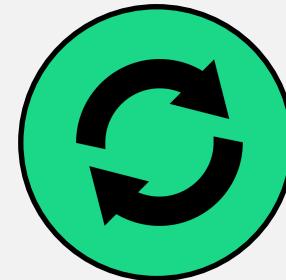
Crossing  
Assists  
Dribbles  
Carries

## Passing



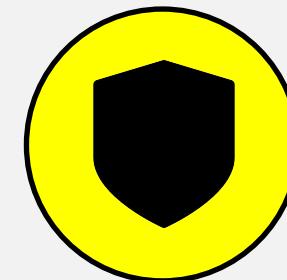
Involvement  
Accuracy  
Direction  
Intent

## Possession



Composure  
On-Ball Decisions  
Progressive Passing  
Interceptions

## Defending



Aerial Ability  
On Ball  
Off Ball  
Fouls

## Goalkeeping



Shot Stopping  
One-on-Ones  
Distribution  
Decision Making



# Rating Players Based on a Specific Game Model

**Striker**



**HIGH PRESSING**

**LINK UP PLAY**

**ATTACKING FIRST  
POST FOR CROSSES**

**RUNS IN BEHIND**

**STRETCH LINE OR  
FALSE 9**



**3-2-2-3  
In Possession**

**OFFENSIVE**

**GOOD FOOTBALL**

**FROM THE BACK**

**ORGANISED,  
DYNAMIC**

**NUMBERS REPRESENT  
PLAYER PROFILES**



# DATA ENGINEERING AND INFRASTRUCTURE



# Licensed and Public Data Providers Accessible to the Club



## Event Data

**StatsBomb**

 **Opta**  
BY STATS PERFORM

**wyscout**

## Tracking Data

 Second Spectrum

**SKILLcorner**

## Sports Science and Physical Data

  
**CATAPULT**

  
**FIRSTBEAT**

  
**VALD**  
**SKILLcorner**

## Scouting and Financial Data

  
**transfer  
markt**

  
**CAPOLOGY**

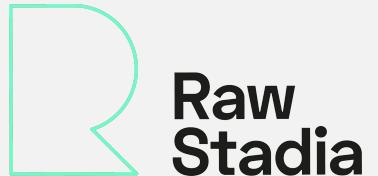
  
**CIES** FOOTBALL OBSERVATORY

  
**SCISPORTS**

  
**SCOUTASTIC**

## Additional Data

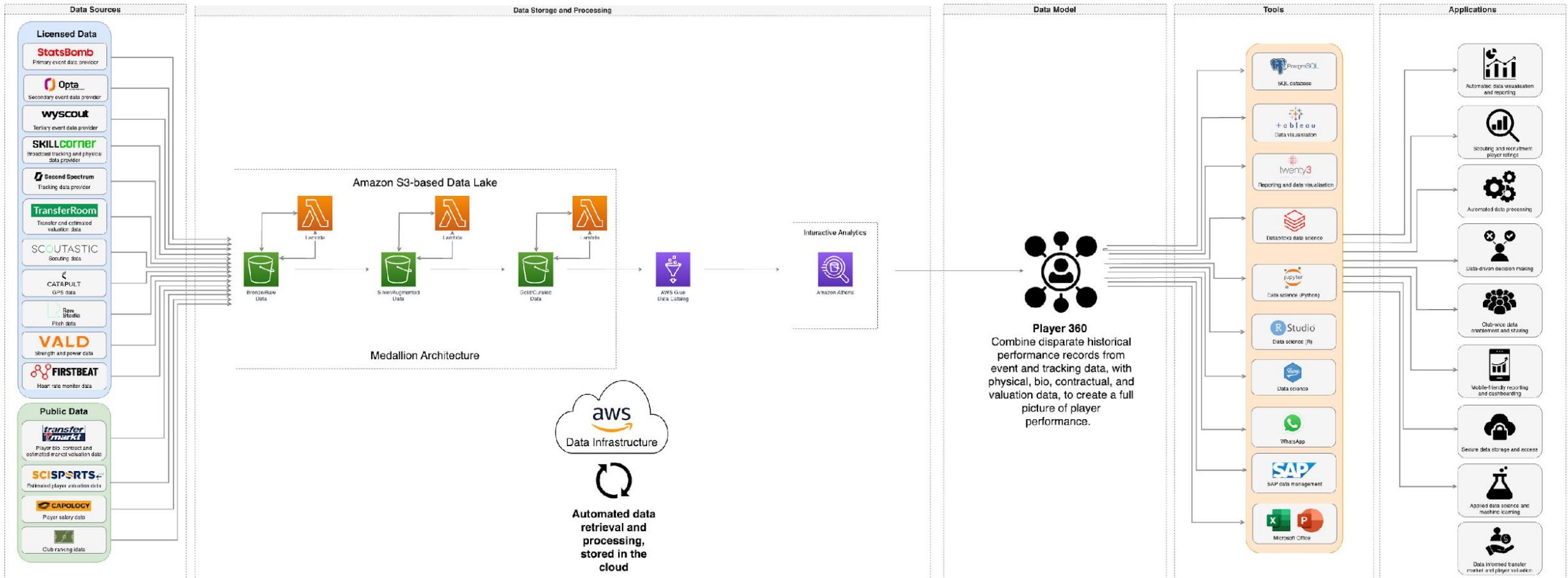
  
**FBREF**

  
**Raw  
Stadia**





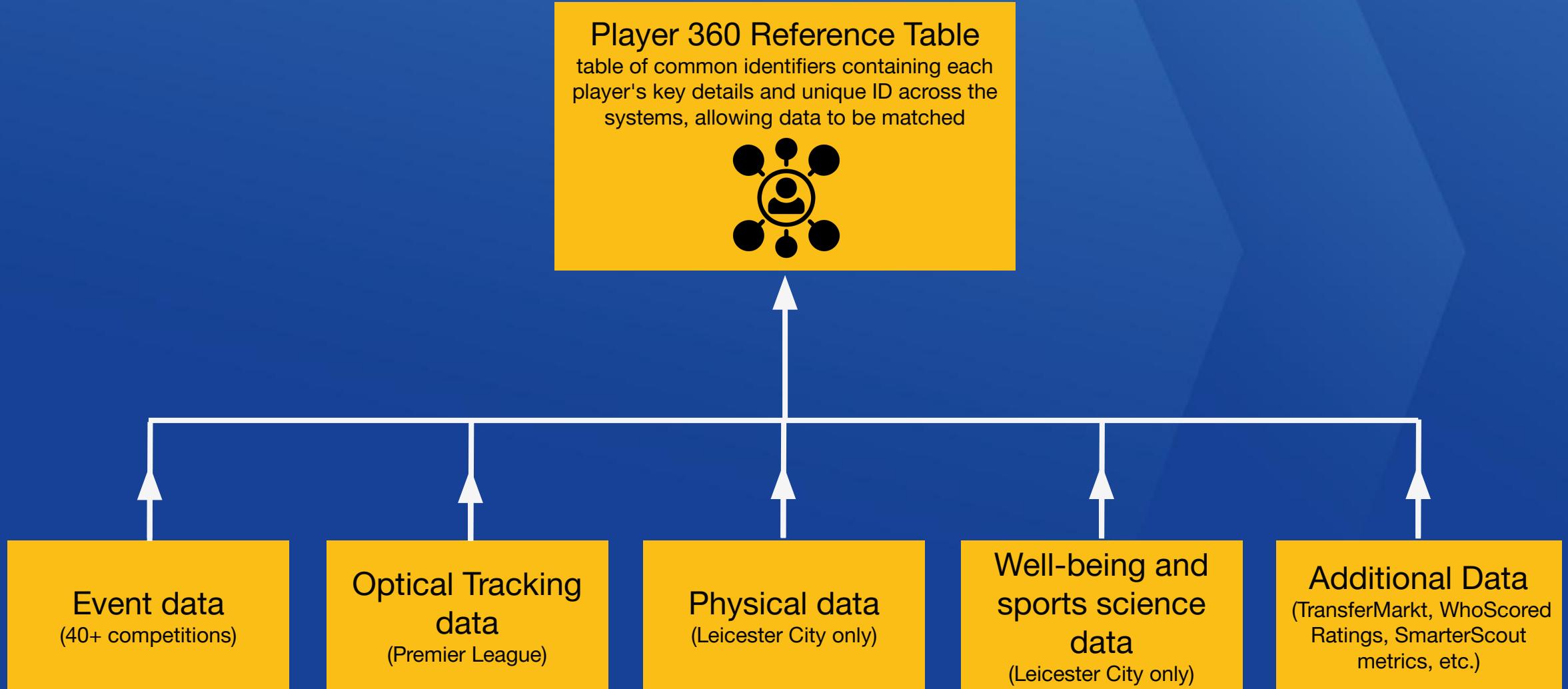
# AWS Cloud-Based Data Infrastructure and Tooling



Direction of flow



# Building the Player 360 Database





# OPEN SOURCE CONTRIBUTIONS





# Overview of Open Source Projects

## 1 Recruitment analysis of 'ball-playing' centre backs

Scenario for a data-driven recruitment analysis, to determine "the next Gerard Piqué", for a hypothetical, newly-promoted club in the 'Big 5' European leagues. This analytical piece features the use of Machine Learning techniques as part of the player selection process, including K-Means clustering and Principal Component Analysis.

## 2 Opposition analysis of a Premier League team through the application of Event and Tracking data

Application of Opta Event data and Second Spectrum Tracking data for two of Crystal Palace's matches in the 21/22 season against Brighton & Hove Albion and Leicester City, as part of an opposition analysis report for a Premier League club.

## 3 Physical data analysis of player GPS training data

Working applications of player training data for the purposes of performance analysis in a club setting, with an output of player reports to measure speed, time, distance covered, acceleration, and number of sprints that players conducted in each of the training sessions they participated in.

## 4 Post-tournament analysis of the England men's team at EURO 2020 through the application of StatsBomb Event and OBV Data

Short analytical piece that applies a data-driven approach in an attempt to answer the question 'England at EURO 2020: How Good Were They Actually?'.

## 5 Tableau Football Intelligence reporting tools

Creation of analysis tools for analysing player performance, pre- and post-match tactics, and recruitment, enabling the democratisation of data to all levels of a footballing organisation or club. Data used features StatsBomb Event data for the FA WSL for the 19/20 and 20/21 seasons.

## 6 Applications of Tracking data

Engineering and analysis of Signality, Metrica Sports, and Stats Perform Tracking data. Analysis includes analysing passages of play, determining physical performance, the implementation a basic Pitch Control models, and applying Expected Possession Value (EPV) frameworks.

## 7 Expected Goals (xG) modeling

Training an Expected Goals (xG) model using Event data, subsequently applied to a separate match dataset of Event and Tracking data, to analyse the performance of the teams in question. Models created using Logistic Regression, Random Forests, and Gradient Boost Decision Trees (GBDT) algorithms including XGBoost and CatBoost.

These projects are all discussed in further detail in my Data Science Portfolio: [docs.google.com/presentation/d/1EQ4meB5Y2TGuln3K31Odkvol-8QIPW2MhZE8UEhZ38c](https://docs.google.com/presentation/d/1EQ4meB5Y2TGuln3K31Odkvol-8QIPW2MhZE8UEhZ38c).



**eddwebster/football\_analytics**

 A collection of football analytics projects, data, and analysis by Edd Webster (@eddwebster), including a curated list of publicly available resources published by the football analytics community.

 Jupyter Notebook  1.5k



**Edd Webster**  
**Football Analytics**

# SUMMARY

# Insights this Football Data Portfolio Aims to Provide

1

**Data insights experience** – this portfolio aims to have provided cross-section of my five years of data insights experience, developed from elite football environments and the commercial departments of top-flight football club, as well as best-in-class industries.

2

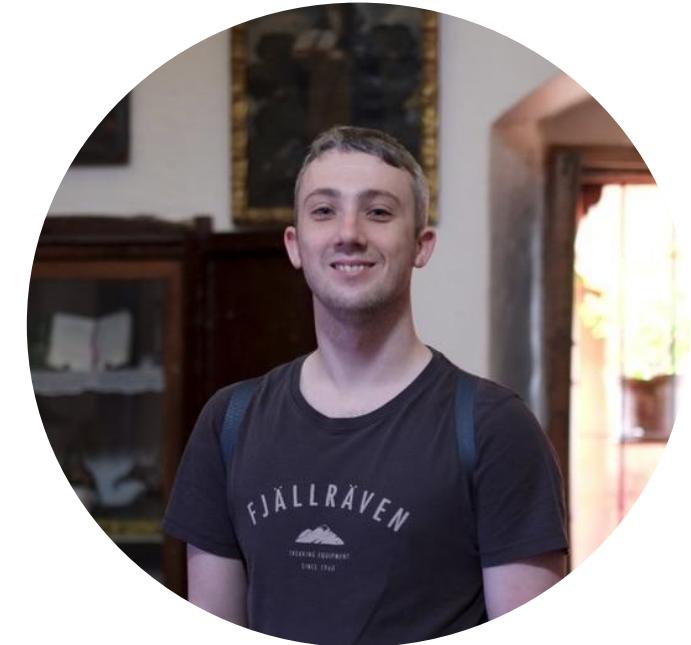
**Diverse skill-set and problem-solving approach** – breakdowns are provided for tackling common industry challenges with working examples, covering the concepts of data processing, modelling, management, visualisation, and strategic planning.

3

**Measurable impact** – demonstrating the quantifiable outcomes for football analytics data and insights.

# For More Information...

- CV: [eddwebster.com/downloads/EddWebsterCV.pdf](http://eddwebster.com/downloads/EddWebsterCV.pdf)
- Football Analytics GitHub (code): [github.com/eddwebster/football\\_analytics](https://github.com/eddwebster/football_analytics)
- Google Drive (code + data):  
[drive.google.com/drive/folders/1yofRp9536WnRL7NKJa5XYjv1vJvDYVbi](https://drive.google.com/drive/folders/1yofRp9536WnRL7NKJa5XYjv1vJvDYVbi)
- Football Data Analytics slide deck (this one you are viewing):  
[docs.google.com/presentation/d/1sWb5UDDp\\_3kIAdLF0E30HTOo0ZKNcK2KPPHAfmscI8](https://docs.google.com/presentation/d/1sWb5UDDp_3kIAdLF0E30HTOo0ZKNcK2KPPHAfmscI8)
- Football Data Science slide deck (of projects featured in the Open Source section):  
[docs.google.com/presentation/d/1EQ4meB5Y2TGuln3K31Odkvol-8QIPW2MhZE8UEhZ38c](https://docs.google.com/presentation/d/1EQ4meB5Y2TGuln3K31Odkvol-8QIPW2MhZE8UEhZ38c)
- TGG Article regarding my thoughts and experience of Data Science and Football Analytics:  
[trainingground.guru/articles/edd-webster-from-man-city-to-lego](https://trainingground.guru/articles/edd-webster-from-man-city-to-lego)
- Tableau Public Profile: [public.tableau.com/profile/edd.webster](https://public.tableau.com/profile/edd.webster).
- LinkedIn: [linkedin.com/in/eddwebster](https://linkedin.com/in/eddwebster)
- Twitter: [@eddwebster](https://twitter.com/eddwebster)
- Email: [edd.j.webster@gmail.com](mailto:edd.j.webster@gmail.com)





Edd Webster  
Football Analytics