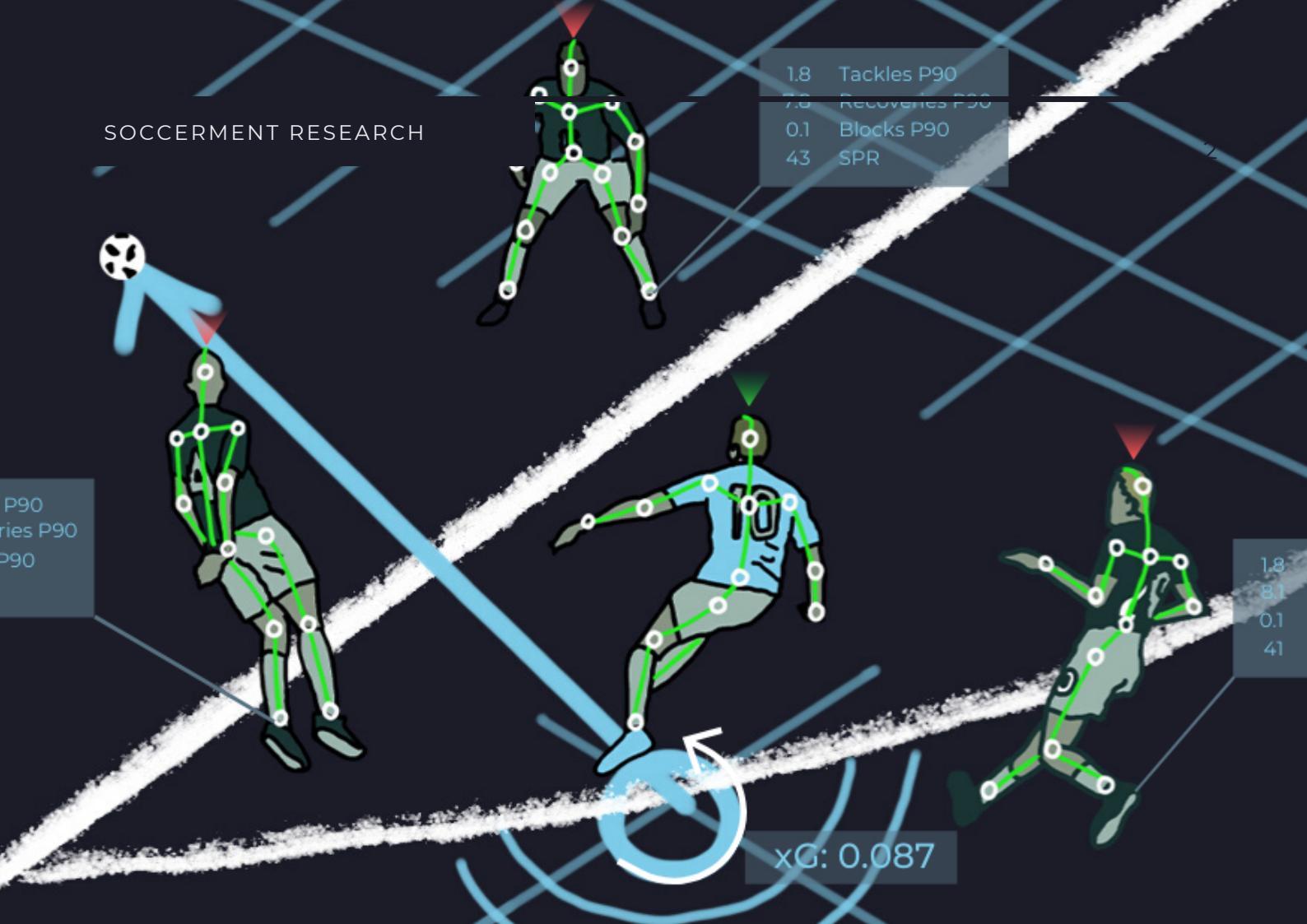


SOCCERMINT RESEARCH



# THE FOOTBALL ANALYTICS HANDBOOK

---

CLUBS THAT AREN'T PLANNING TO JUMP ON THE ANALYTICS BANDWAGON RUN THE RISK OF BEING LEFT BEHIND.





Over the past two decades, the influence of data analytics has been growing in every aspect of our lives: in businesses of every kind, but also in healthcare, media and sports. Until a few years ago, football was thought to be immune from this trend. Now, the early adopters in the major football leagues are thriving thanks to the competitive advantage that investments in data analytics are beginning to provide them.

- 
- |           |                                       |           |                    |
|-----------|---------------------------------------|-----------|--------------------|
| <b>3</b>  | A Brief History of Football Analytics | <b>19</b> | Player Development |
| <b>5</b>  | Applications of Football Analytics    | <b>23</b> | Injury Recovery    |
| <b>6</b>  | Smart Scouting                        | <b>26</b> | Books and Articles |
| <b>12</b> | Performance Analysis                  |           |                    |

---

Handbook  
**INDEX**



## A BRIEF HISTORY OF FOOTBALL ANALYTICS

Football analytics is not as young a discipline as we often think. The first proponent was an accountant in the British Royal Air Force named Charles Reep, who, after World War II, started collecting and analysing data about football matches, using a pencil and paper. Reep concluded that most goals were scored from fewer than three passes, and therefore that it was important to get the ball forward as soon as possible. His theory became known as the long ball, and would have great influence on English football for many years, especially in the Eighties. What Charles Reep was doing right was data collection. What he did wrong were the conclusions he reached after analysing the database he put together with so much passion.

THIS IS A CLEAR EXAMPLE OF HOW DATA ALONE ISN'T ENOUGH TO GAIN A COMPETITIVE EDGE, WHAT'S EVEN MORE IMPORTANT IS THE ABILITY TO INTERPRET IT.

---

A brief history of  
**FOOTBALL  
ANALYTICS**





The publication in 2003 of 'Moneyball: The Art of Winning an Unfair Game' was sports analytics' real introduction to a wider audience. According to the book, before sabermetrics were introduced to baseball, teams were dependent on the skills of their scouts to find and evaluate players. Football managers, initially, remained sceptical. Most people thought that, unlike American sports, football would be impossible to analyse using data. But a few stubborn people disagreed. And they went on to start a data revolution in the Beautiful Game. One of them was Billy Beane himself, who is deeply interested in football.

The most recent developments, the successful case studies of Liverpool, AZ Alkmaar, FC Midtjylland and Brentford, show that ...

**...NOT ONLY IS IT POSSIBLE TO ANALYSE FOOTBALL, BUT THOSE WHO DO IT PROPERLY GAIN AN IMPORTANT COMPETITIVE EDGE.**

---

A brief history of  
**FOOTBALL  
ANALYTICS**





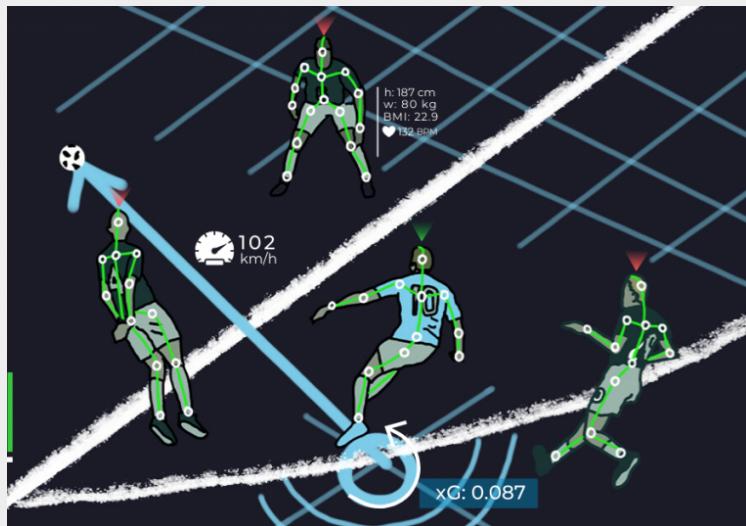
## APPLICATIONS OF FOOTBALL ANALYTICS

In the past few years, the exponential speed of improvement in the technologies supporting the collection, storage and analysis of data has gone hand-in-hand with an exponential increase in the human capital invested in sports analytics.

As the datasets have grown and improved, the number of potential applications of data analytics to the game has multiplied, making “football analytics” a quite generic concept. We try to summarise the main fields of application.

We will go through:

- SMART SCOUTING
- PERFORMANCE ANALYSIS
- PLAYER DEVELOPMENT
- INJURY PREVENTION AND REHABILITATION



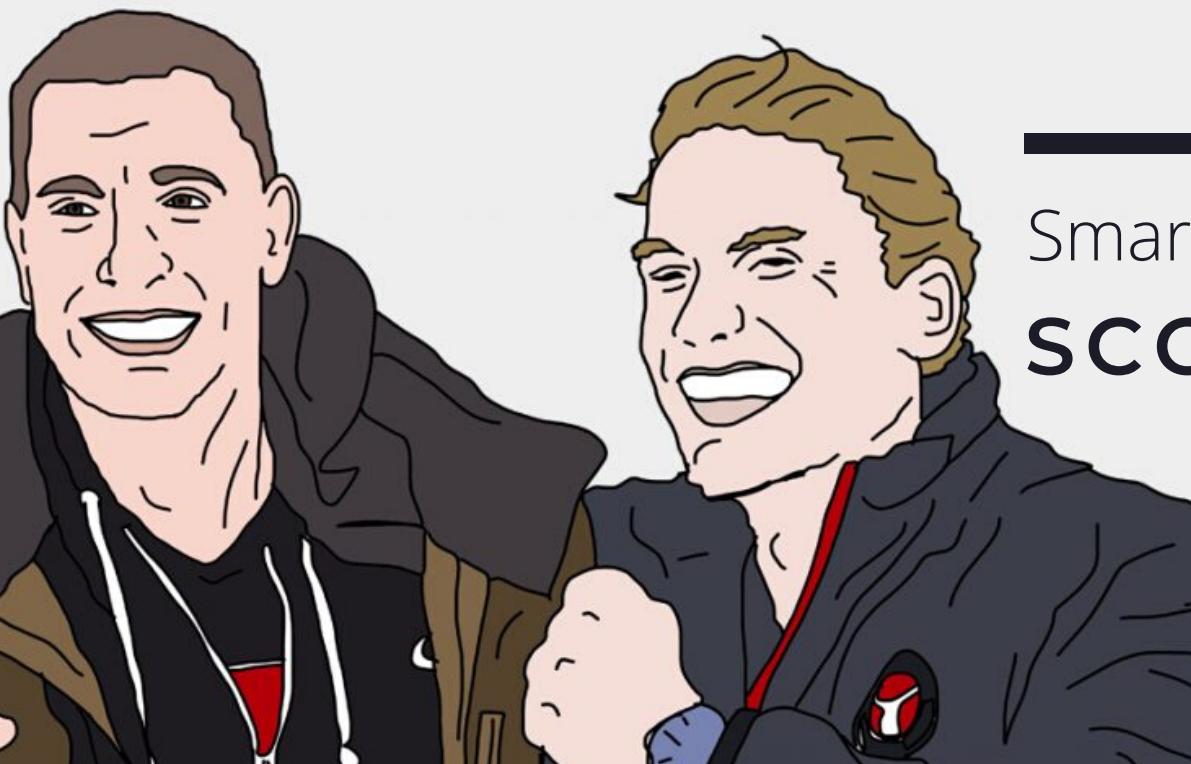
Applications of  
**FOOTBALL  
ANALYTICS**



## SMART SCOUTING

Data analytics is a very powerful tool for scouting. Here's why:

- **SAVINGS.** Searching for players in large and detailed databases allows clubs to save incredible amounts of time and money. We believe that databases cannot replace scouts, but rather that they can complement their talent identification skills.
- **SILICON HAS MORE MEMORY THAN YOU AND I.** Do you remember all the actions, all the shots, all the crosses and dribbles by your favourite team in the past season? Of course not. Luckily, computer doesn't have a favourite team or player, it can remember everything that has occurred in past seasons.
- **LIMITING PRE-CONCEPTIONS.** Skimming through databases often gives rise to counter-intuitive results. In some cases, these might be false positives. In other cases, they deliver proper "eureka moments".



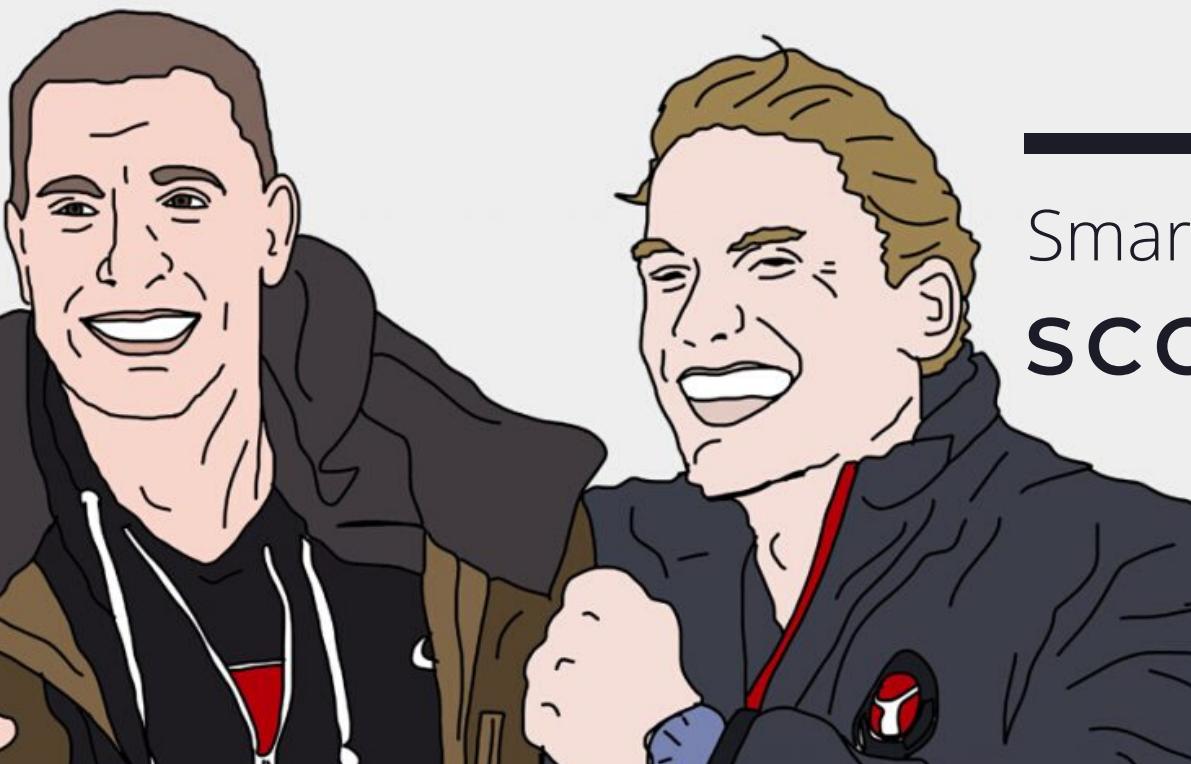
Smart  
**SCOUTING**



Matthew Benham is an ex-City trader who founded the betting-consultancy, Smartodds, a privately-owned company that provides statistical research and sports modelling services to customers like professional gamblers. In June 2012 Benham became the majority owner of Brentford FC, which was then fighting for promotion from League One. In July 2014 he became the owner of Danish club FC Midtjylland.

The English club Brentford and the Danish FC Midtjylland have one thing in common, apart from their ownership: they are able to punch above their weight, i.e. they have figured out how to compete with much smaller budgets than their peers.

**THE ADVANCEMENTS THAT WE HAVE SEEN IN THE INDUSTRY IN THE PAST FIVE YEARS WILL BE DWARFED BY WHAT WILL HAPPEN IN THE NEXT FIVE.**



Smart  
**SCOUTING**



The way they do this is a relatively simple concept, but one that it is very hard to actually pull off: maximise returns in the transfer market. Their success in the transfer market allows Brentford to successfully compete in the English Championship, despite having a salary budget of less than £15m, 60% lower than the league average (£39m).

[CLICK HERE](#)

## N. MAUPAY

Last season at Brentford  
2018/2019

**43**

Appearances

**25**

Goals

**8**

Assists

**1.8**

Purchasing fee (£ m)

**20**

Selling Fee (£ m)

**18.2**

Profit (£m)

[CLICK HERE](#)

## A. GRAY

Last season at Brentford  
2014/2015

**45**

Appearances

**17**

Goals

**6**

Assists

**0.5**

Purchasing fee (£ m)

**12**

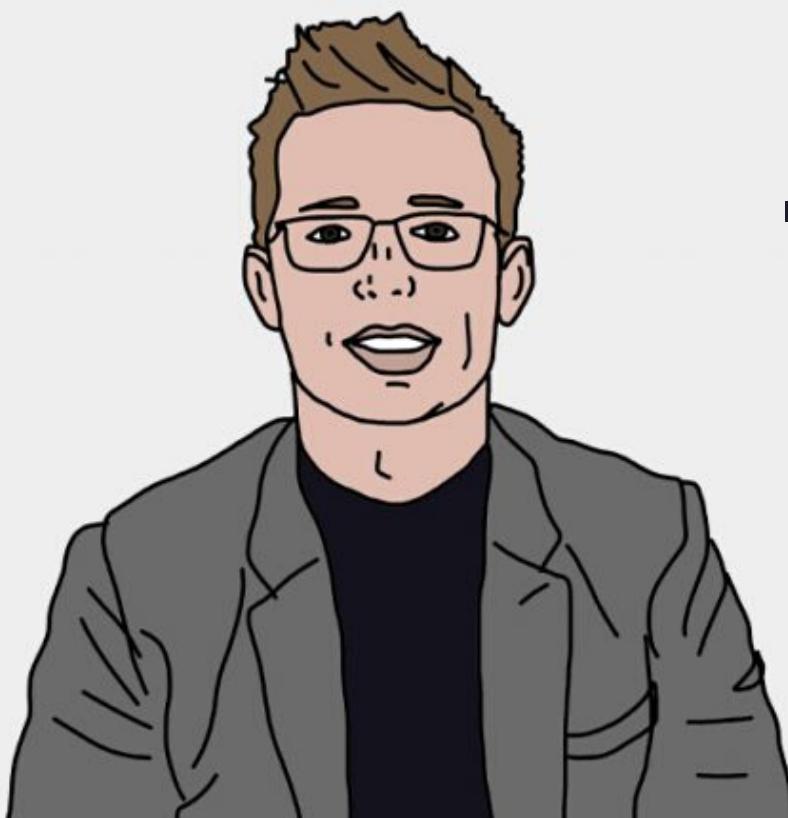
Selling Fee (£ m)

**11.5**

Profit (£m)



In 2010 Liverpool was acquired by American company Fenway Sports Group, owners of the baseball team Boston Red Sox since 2002. At Liverpool, Fenway Sports Group decided to invest money in data analytics. They soon hired Damien Comolli as Director of Football Strategy. Before leaving Liverpool though, he did make one important signing, a man who went on to become pivotal to the club's recent successes. For once, we are not talking about a playing star, but a data analyst: Michael Edwards. He works along with the head of recruitment Dave Fallows, the chief scout Barry Hunter, and a four-man research team headed up by Cambridge graduate Ian Graham (PhD in Physics) and including William Spearman (PhD in Physics and ex-CERN), Tim Waskett (PhD in Astronomy) and Dafydd Steele(Statistical researcher). This group really is football analytics's dream team. Football clubs, when keen on data analysis, normally have at most one person with such a background. Having four of them is more than unprecedented, it is revolutionary.



---

Smart  
**SCOUTING**

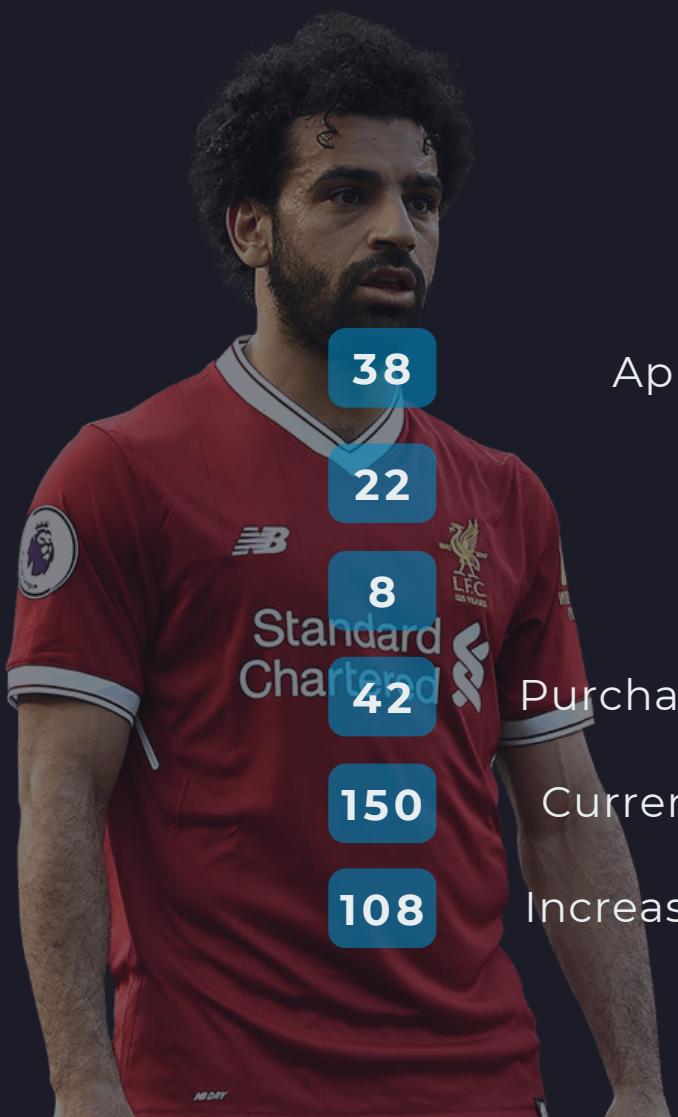


The purchases of Mané, Salah, Alisson and Virgil van Dijk all seemed too expensive at the time. Instead, the players' valuations have significantly increased since Liverpool bought them. Below we compare the purchasing fee with the current valuation (post-COVID) of two of Liverpool's most important signings since 2016/17 (the first full season with Edwards as Sporting Director).

[CLICK HERE](#)

## M. SALAH

Last season at Liverpool  
EPL 2018/2109

[CLICK HERE](#)

## S.MANE'

Last season at Liverpool  
EPL 2018/2109





So, what is the secret formula that Brentford and Liverpool are using? We believe that it comes down to two main factors.

- **DATA ANALYTICS.** Brentford and FC Midtjylland use data and algorithms from Smartodds which are based on the concept of Expected Goals; FC Liverpool have formed a dream team of data scientists to better gather and analyze data.
- **SELLING THE OVER-PERFOMERS.** Not being afraid of selling a player is key to both Brentford's and Liverpool's financial approach. In the transfer market, football clubs often forget long-term, rational considerations and make their fans' short-term emotions the main driver of their decisions. Like Matthew Benham's Brentford, Liverpool aren't afraid of selling important players, because the proceedings give them the opportunity to reinvest the money in either undervalued talents or players who better suit the club's playing style.



---

Smart  
**SCOUTING**



## PERFORMANCE ANALYSIS

Football clubs, along with the media and even fans, can, with the support of data, easily understand how a team or player has performed over the course of a match or season. Even the most basic stats, like the number of shots, shots on target, and ball possession, can offer a guide, albeit an incomplete one, of how two teams have performed and if their performance was in line with the result of the match. If we then improve the granularity of the stats and include selected performance indicators, the dataset gradually begins to give a more defined and complete picture of what has happened on the pitch, which enables a detailed analysis of the game. Such granularity can offer a coach more information on how the team has performed and even whether or not it complied with the pre-match instructions.

**BIG DATA ANALYSIS HELPS US TO DISCOVER COUNTER-INTUITIVE FACTS ABOUT FOOTBALL, UPON WHICH WE CAN BUILD AN INNOVATIVE AND WINNING STRATEGY.**



---

Performance  
**ANALYSIS**



## AN EXPECTED GOAL REPRESENTS THE EXPECTED VALUE FROM A SHOT, I.E. THE PROBABILITY OF THAT SHOT BECOMING A GOAL.

Calculating such probability depends on a number of factors, including distance from goal, the angle of the shot, the body part the shot was taken with (head, strong or weak foot), the playing situation (open play, counterattack), and others (whether there were any opponents blocking the space ahead, for example).

Expected Goals are normally calculated using a “frequentist” approach: a dataset of several thousand shots is categorised according to the aforementioned variables (distance from goal, angle, body part, etc). If one type of shot, which for example occurred 10,000 times in the dataset, is scored 200 times, then that kind of shot is “expected” have an xG of 0.02 (2% probability). Using this methodology, we can determine that a penalty is worth 0.76 Expected Goals, i.e. a penalty has an implied probability of 76% of being scored.



---

## Performance ANALYSIS



How can Expected Goals revolutionise football? In our view, there are two main areas that are clearly open for disruption:

- A deeper and wider analysis of Expected Goals will “force” football coaches to rethink their tactical strategies, especially in one specific area: shot location. In the next 10-15 years we expect to see in football something similar to what has happened in the NBA: a complete overhaul of shooting locations. By looking at the most common shot locations in the NBA 2001-02 compared to the ones of 2016-17 it is noticeable how risk-reward considerations based on analytics have dramatically increased the amount of three-pointers and limited the number of two-pointers outside “the paint”.
- We would expect (and encourage) football clubs to increasingly rely on objective measures such as Expected Goals when it comes to gauging teams’ performances and evaluating how well or badly coaches are doing. It happens all too often that coaches are sacked after poor results despite these not being caused by poor team performance, but rather by bad luck.



---

## Performance ANALYSIS



Expected Assists are derived from Expected Goals and measure the probability of a created chance (or key pass) being converted into a goal. Very skilled creators whose teammates haven't been too good at finishing can, with the help of Expected Assists, be recognised for their efforts. Take for example, Kylian Mbappé: in the French Ligue 1 he managed 5 assists this season, or 0.3 per 90 minutes. Not a bad return, but far from the highest-ranking players on this metric. However, he has created 10.6 Expected Assists, 0.625 per 90 minutes, one of the highest in Europe.

[CLICK HERE](#)

## K. MBAPPE'

Paris Saint-Germain  
Ligue 1 2019/2020

20

Appearances

18

Goals

5

Assists

0.3

Assists P90

0.6

xA P90

**WE HAVE DEVELOPED OUR OWN PERFORMANCE INDICATOR, WHICH WE APPLY TO ALL FOOTBALLERS AND TEAMS IN OUR DATABASE, CALLED SOCCERMENT PERFORMANCE RATING ('SPR')**

The SPR is a synthetic measure with which we can value a player's overall contribution to the team's performance. All events on the pitch are taken into consideration and weighed using ad-hoc coefficients, which also depend on the players' roles. The algorithms split the players' contribution into three phases – defence, buildup and attack – in order to give us a clearer view of the players' overall approach to the game.



---

## Performance ANALYSIS

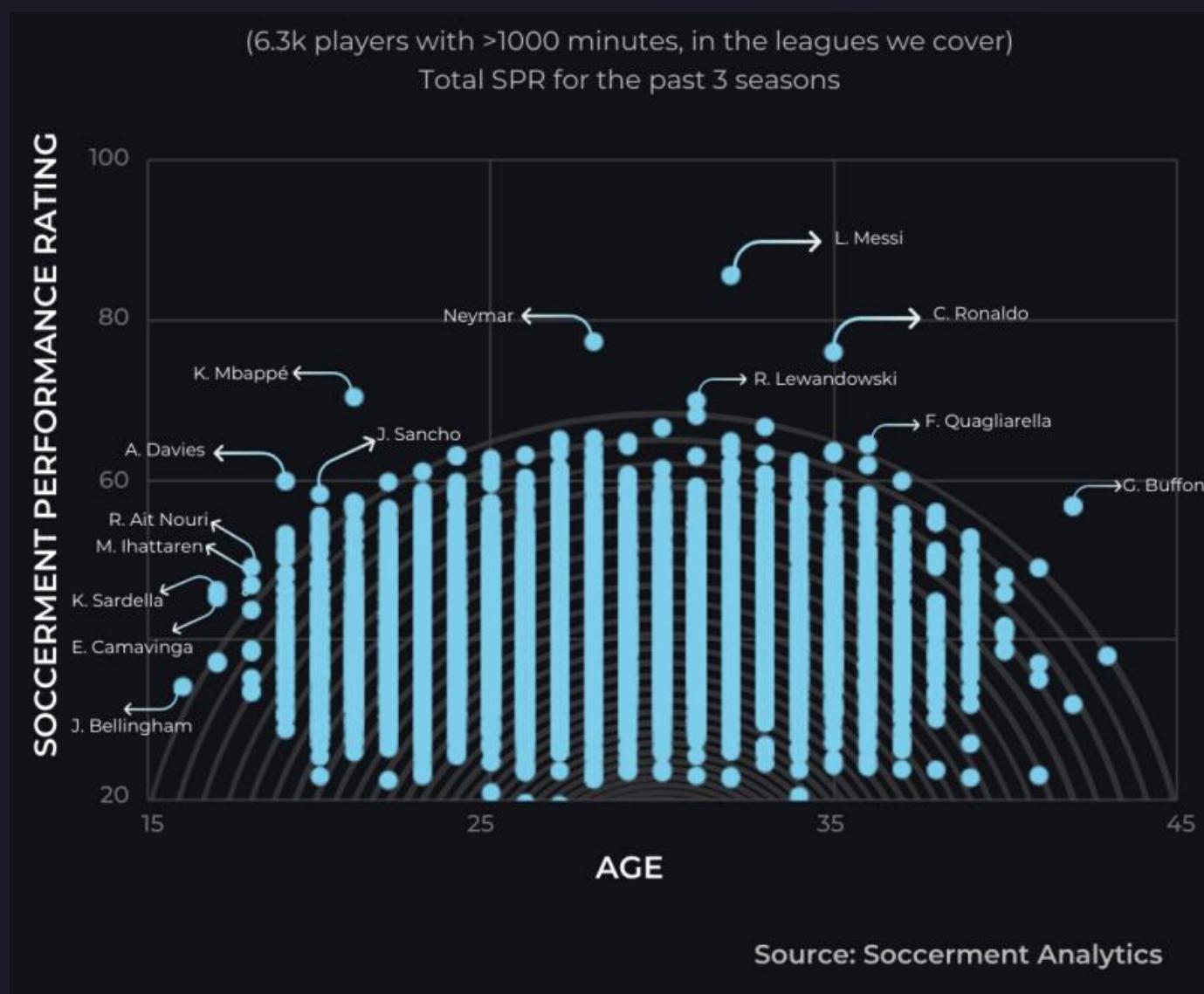


The Performance Indexes, reproduced in our spider charts, offer a synthetic measure of a very specific aspect of player performances. Each performance index takes into account n metrics belonging to a specific technical event. For example, for the “Vision” index, we take into account, among other things, the quantity and quality of chances created, through balls and long balls; while the “Passing” index applies different weights to the players’ passes, depending on the area of the pitch from which they originate, their direction, their accuracy and their length.





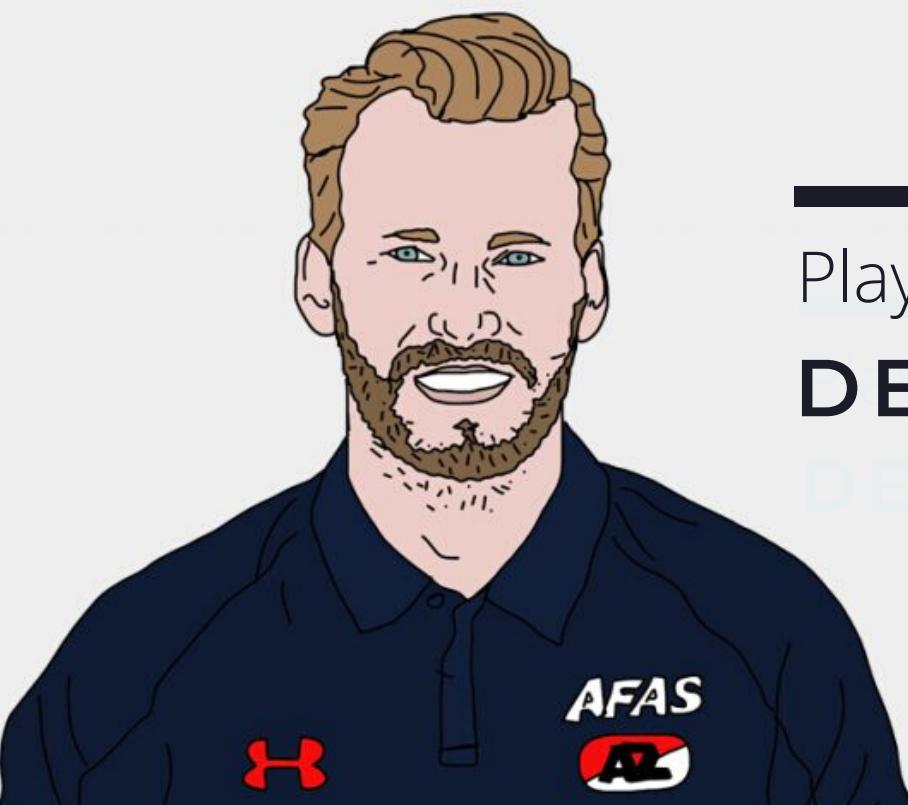
The Performance Indexes and the SPR are powerful scouting tools. The former can clearly identify the best players in one specific, technical aspect of the game; while the latter offers an immediate view on players' performance level during the season, or even throughout their entire career. Using a performance index like ours it is easier to spot outliers. In the case below, we used it to find the outliers for different age groups. For more information, you could read our post on [LinkedIn](#).





## PLAYER DEVELOPMENT

Data analytics is also growing in importance when it comes to youth development. The reason is simple: having objective and measurable feedback can help both coaches and players speed up the learning processes and create virtuous development cycles. Basically, data analytics becomes a tool to help predict and cultivate players' potential. The best example of this is probably the Dutch club AZ Alkmaar. We have written extensively about how AZ structure their development process through data in the analysis linked here below. In the analysis we also explain how long-term planning requires clear and efficient governance, which is still not too common at football clubs, at any level. Above is pictured Marijn Beuker, AZ Alkmaar's Director of Sports Development.

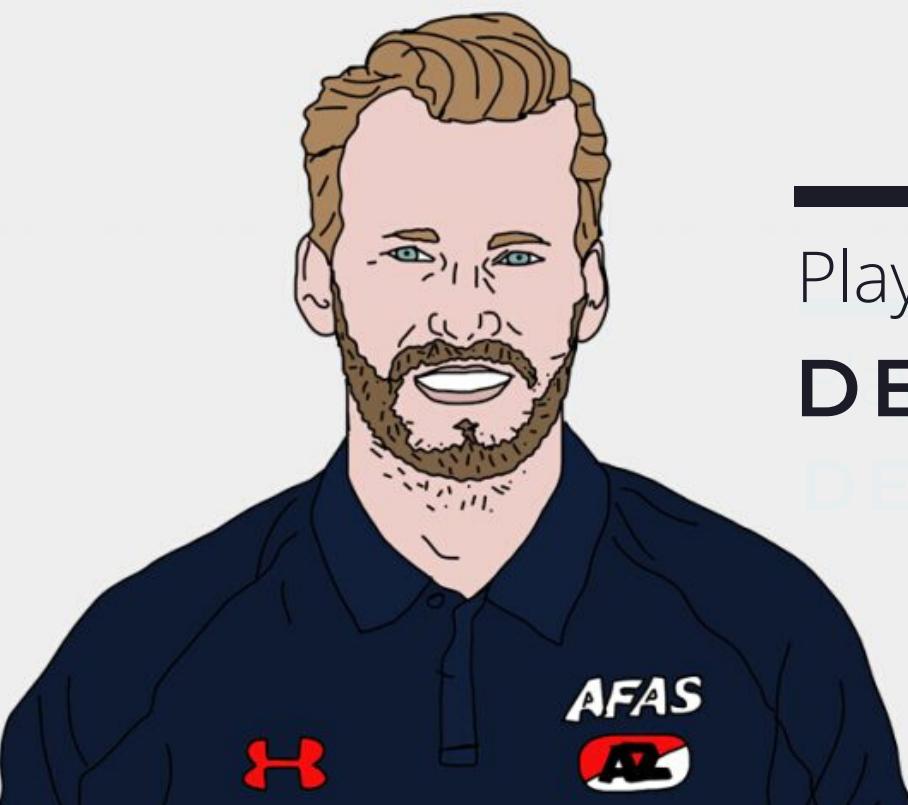


Player  
**DEVELOPMENT**



## 'WE DON'T BUY SUCCESS, WE CREATE IT' (MARIJN BEUKER)

A packed audience is patiently waiting for the speaker to start. The title of the presentation is intriguing: "How AZ Alkmaar Became One of Europe's Best Talent Factories Using Data". We are there, very curious, to listen to Marijn Beuker, Head of Performance & Development at AZ Alkmaar. Our curiosity stems from the development angle of the presentation. While data analytics is usually presented as a tool for scouting players and/or for match analysis, Mr Beuker's speech focuses on the importance of data analytics in player development, i.e. as a tool to help predict and cultivate players' potential. Compared to many other teams' practices, this really is next level thinking. Won over by the speech, by AZ's history and by their recent performances, we began to gather and analyse information about the club and its most valuable assets – the many talented young players currently on its roster.



Player  
**DEVELOPMENT**



There are several young players at AZ Alkmaar worth focusing on. Calvin Stengs is a left-footed attacking midfielder, he usually starts off as a right winger in a 4-3-3 or 4-2-3-1 formation. However, he normally cuts inside to finish attacking phases in the centre, between the lines, as a 10. From that position, close to the penalty area, he likes to perform accurate through balls to his teammates. Very creative, technically skilled and with remarkable vision, he is the archetype of the chance creator. He recently made his debut with the Dutch national team (with 2 assists, against Estonia).

[CLICK HERE](#)

## C. STENG'S

WAM | AZ Alkmaar  
Eredivise 2019/2020

**25**

Appearances

**5**

Goals

**7**

Assists

**2.4**

Successful dribbles  
P90

**0.7**

Through balls P90

**2.5**

Chance created P90



Boadu is the archetypal modern striker. Physically strong (although not imposing), he is very good at running behind opposing defences and finding space to shoot. He recently made his debut with the Dutch national team, scoring a goal against Estonia. This makes him the perfect partner for Calvin Stengs, who is very good at calibrating accurate through balls.

[CLICK HERE](#)

## M. BOADU

FW | AZ Alkmaar  
Eredivise 2019/2020

**24**

Appearances

**14**

Goals

**6**

Assists

**2.4**

Shots P90

**52%**

Shooting Accuracy

**7**

Touches in the box  
P90



## INJURY PREVENTION

The origins at Milan Lab. It was founded in 2002 to reduce the risk of AC Milan players getting injured, help injured players recover faster, and improve training methodologies by personalising sessions. It is said that AC Milan's chairman, Silvio Berlusconi, decided to form Milan Lab and tasked Belgian doctor Jean Pierre Meersseman with the job after Fernando Redondo, recently signed from Real Madrid, ruptured his anterior cruciate ligament during pre-season training. Milan Lab's fame grew over the years and began to be seen as one of the drivers of AC Milan's domination of European football. When the Rossoneri won the Champions League in 2007, for instance, the team's average age was well above 30 and captain Paolo Maldini was approaching 40 years of age. Milan Lab seemed to be dramatically extending players' careers. Although the use of technology and data was pivotal to Milan Lab, its detractors say that at the very core of the project was the chiropractic philosophy, which is seen as unorthodox by most of the scientific world.



---

Injury  
**PREVENTION**



## TODAY: DATA & TECH TO HELP ON ATHLETIC PERFORMANCE AND INJURY PREVENTION

Nowadays, any high-level club consistently monitors players' position, while velocity tracking with GPS allows them to objectively measure the "external load", i.e. the amount of work performed on the pitch. External load tracking through GPS can be also used to predict the risk of injury professional players. Researchers discovered that the greatest injury risk occurred when players accumulated a very high number of short bursts of speed during training over a three-week period. Players recorded significantly higher meters per minute in the weeks preceding an injury, compared with their seasonal averages (+9.6 and +7.4% for 1- and 4-week blocks, respectively), indicating an increase in training and gameplay intensity in the lead up to injuries. Furthermore, "injury blocks" showed significantly lower average new body load compared with seasonal averages (-15.4 and -9.0% for 1- and 4-week blocks, respectively).



Injury  
**PREVENTION**



## PERFORMANCE MONITORING CAN BRING A CLEAR VISION OF THE FRAMEWORK REQUIRED TO DEVELOP AND OPTIMISE THE MODELS USED TO ANALYSE TRAINING LOADS.

Current research in sports biomechanics supports this trend, showing that athletes' movement patterns deteriorate in a potentially harmful fashion when repeating intense running and changes of direction. Here is a [link to a research paper](#) on this subject, authored by Soccermint's Head of R&D Matteo Zago and published in the European Journal of Sports Science. In sum, performance monitoring can bring a clear vision of the framework required to develop and optimise the models used to analyse training loads. This will help these models gain better insight into players' fitness, readiness to perform and fatigue, and improve the quality and efficiency of their assistance to coaching staff.



## Injury PREVENTION



## MASTER FOOTBALL ANALYTICS: BOOKS AND ARTICLES TO READ

### Books:

- “Moneyball: The Art Of Winning An Unfair Game” (by Michael Lewis)
- “The Numbers Game” (by Chris Anderson and David Sally)
- “Soccernomics” (by Simon Kuper and Stefan Szymanski)
- “Soccermatics” (by David Sumpter)
- “The Expected Goals Philosophy” (by James Tippett)
- “Football Hackers” (by Christoph Biermann)

On YouTube: 'The Best Books About Football Analytics' (Italian)

### Articles by Soccermet Research:

- “The growing importance of football analytics”
- “AZ Alkmaar: Creating Success Through Data Analytics”
- “AZ Alkmaar's Talent Factory”
- All the articles



Master  
**FOOTBALL  
ANALYTICS**

FIND OUT  
MORE

