src - FTBL - Project 3 Passing Analysis

src-ftbl is on a mission to elevate data analysis in football, led by a team from around the world of sector leaders and industry experts.

This report presents an analysis of a **player's (“Jude Bellingham " or “JB”)** passing ability primarily using Statsbomb 360 data and from the several sources FBref, WhoSocred and Understat.

The goal of this analysis is to assess the player's performance in passing and provide insights into their impact on their team. The analysis utilises Python programming language and Jupyter Notebook for data processing and visualisation.

1. What was your approach to the problem and why did you choose that methodology?

1 Defining the Question and Intention

The evolution of football has been parabolic, the advancement of stats, wages transfer fees, broadcast deals and stadia, with so much occurring in the game it is hard to define the signal in the noise, where do you start?

Passing is a pillar of the game of football, but is also intrinsically linked with positioning which are core tenets and foundational blocks of the ecosystem of the game, and determinants of a clubs’ success.

Taking a 10,000 ft view of the problem, the approach taken here has been a strategic view to produce the assessment. Taking an “if so what” approach allowed the analysis to become a value add rather than a weather report. This philosophy informed the basis of the visualisations used in the report to tell the story of how passing can affect other facets of the game, build up team structure etc.

The second step in defining the problem and the overall approach to the assessment, was to isolate emergent elements in the game, blockbuster transfers fees, highly rated players and sustainability. From there the emergence of Financial Fair Play and new money into the game further creates the need for clubs to be more methodical and strategic in their analysis, defining the opportunity cost the net present value to build a sustainable pathway for their club to grow when making strategic acquisitions.

Based on the above scope, players were selected on the following criteria:

1) who are being linked with big money moves (£100m+) this summer (2023);
2) and would be joining a new league

From this filter **Jude Bellingham** was the choice of player to analyse for this assessment.

Data Processing and Visualisation

The following methodology was employed to analyse the player's passing ability:

2.1 Data Collection
The first step involved obtaining the Statsbomb 360 data for the player under consideration. The dataset includes information such as pass coordinates, outcomes, contextual data (match, team, opposition), and additional relevant attributes.

In addition Data was sourced and corroborated from other sources:

- FBref
- WhoScored
- Understat

2.2 Data Preprocessing
After data collection, the dataset was preprocessed to ensure its suitability for analysis. This included cleaning the data by removing irrelevant columns, handling missing values, and converting data types if necessary.

2.3 Exploratory Data Analysis (EDA)
An initial exploration of the data was conducted to gain insights into the depth of the player's performance (passing) and content of the data. Basic statistics such as pass completion rate, total number of passes, and average pass length were retrieved. In addition specific modules and packages (such as mplsoccer) to delve further into the data to support the development of visualisations.

2.4 Visual Representation:
Visualisations such as heatmaps and pass network diagrams were utilised to visualise passing patterns. These visual representations offer intuitive insights into the player's passing patterns and spatial awareness.

Reasons for the approach
Overall, this approach was chosen to leverage available data, provide a comprehensive evaluation, visualise passing patterns, offer contextual comparison, ensure objectivity, identify strengths and weaknesses, and allow for potential improvements. It aims to deliver a robust analysis of the player's passing ability to support decision-making in football analysis.

Note see section 3 on additional items which would have been conducted if there was additional time and data available.

2. What are the pros and cons of this approach?

Pros:

- The approach looks beyond the face of the problem and probes to delve further by asking questions about positioning and other contributing factors around passing.
- The chosen approach allows for a comprehensive assessment of the player's passing ability, incorporating both quantitative and visual insights.
- Utilisation of advanced passing metrics and network maps provides a deeper understanding of the player's performance.

Cons:

- The analysis heavily relies on the quality and completeness of the Statsbomb 360 data.
- Subjective aspects of passing ability, such as decision-making and vision, are not fully captured.
- Lacking comparisons and benchmarks of other players does not provide full context to evaluate the player's passing ability.

3. Given more time or data, how would you improve your submission?

Given more time or data, the following improvements could be made to enhance the analysis:

Additional Approach items
1.4 Base Level Passing Metrics
Rather than pulling from third party sources - calculating several advanced passing metrics were calculated. These metrics include:

- Pass completion rate: The percentage of successful passes out of the total attempted passes.
- Pass accuracy: The average distance between the intended target of the pass and the actual outcome.
- Progressive passes: The number of passes that significantly move the ball forward towards the opposition goal.
- Key passes: The passes that lead to a shot attempt by a teammate.
- Assists: The passes that directly result in a goal by a teammate.

1.5 Comparative Analysis

To provide context to the player's passing ability, a comparative analysis would be conducted. This would involve comparing the player's passing metrics with relevant benchmarks or other players in the same position. Comparisons would be made to league averages, team averages, or other players with similar roles, and crucially transfer market value.

The comparative analysis would allow for deeper insight on a player's integration into a new team and the impact on a club's 5 year transfer strategy and sustainability criteria (see 1.7 below).

1.6 Advanced Passing Metric Analysis

Incorporate additional passing metrics, such as pass intensity, pass variety, or pass combinations, to provide a more comprehensive evaluation of the player's passing ability.

Analyse the player's passing performance in different game situations, including attacking phases, counter-attacks, or set pieces.

Considering the integration of machine learning techniques to predict the success of passes based on contextual features, enabling a more predictive analysis of the player's passing ability.

1.7 Value Analysis

With transfer spending increasing year on year, money helps but does not guarantee success, therefore, the question and analysis of this report intends to highlight areas of value creation for the players potential move.

Blending these passing metrics along with other forms of data such as wellness/biomarker data could be leveraged to create a predictive model on how the player may integrate in certain conditions, times of the season/matches, and build an expected performance rating.

Going forward including transfer market value, trophies won, matches won etc, chemistry with teammates style of play to create a **net rating** would be a transformational step forward in bringing the impact of the analysis further forward especially for evolving recruitment strategies.

4. APPENDIX

Please click on the **link** or see **README.md** to review source code and associated links to the 5 slide presentation and a self-evaluation report.

README.md

This report contains the following Deliverables:

- Source code for the analysis
    - JB Passing map
    - JB Heatmap
    - England France Network Passmap

- A 5 slide maximum presentation to deliver to the coaching staff about some insight of a player's passing ability

- An additional report answering the following questions:
    - The approach to the problem and why did you choose that methodology?
    - What are the pros and cons of this approach?
    - Given more time or data, how would I improve the submission?