

# Module 4. Organizing Information

## 1. What is Big Data?

The term "big data" refers to large volumes of data generated at a speed and variety that exceed the capacity of traditional data processing systems to capture, manage, and process them efficiently (Mayer-Schönberger & Cukier, 2013). Big data involves the use of advanced technologies to analyze this data and extract valuable information that can be used for decision-making in various fields, including football (McAfee & Brynjolfsson, 2012).

According to Bai and Bai (2021), big data in sports is defined as the determination, storage, management, and analysis of data. When talking about big data in the world of football scouting, it refers to the use of large volumes of data to identify, assess, and recruit players more efficiently and accurately. Understanding the concept of big data in football involves recognizing that this data ranges from basic statistics such as goals and assists to more detailed information such as the distance covered by players, speed, ball possession, and pass accuracy.

## 2. The Application of Big Data in Scouting

In modern football, the use of big data has completely revolutionized how teams analyze and improve their performance. This innovative approach takes advantage of the vast amount of data generated during matches, training, and player analysis to provide deep and strategic insights. From evaluating individual performance to optimizing tactics and preventing injuries, big data has transformed all aspects of football. Below, we present the different big data tools:

### GPS

GPS tracking devices in football are technological tools used to monitor and collect precise data on players' physical performance during training and matches. These devices are part of the players' equipment and work by continuously capturing information such as the distance covered, speed, acceleration, coverage areas on the field, active pauses, and movement intensity.

As for the functionality that this tool provides for scouts, these devices allow tracking and analyzing the physical performance of players during matches and training. Scouts can assess metrics such as speed, distance covered, and endurance, providing insights into the athletic ability and physical performance of the players.

### Tracking

### Advanced

There are platforms that use computer vision algorithms and machine learning to analyze

### Video

### Analysis



match and training videos, extracting data on tactics, player movements, ball possession, and more. The functionality for scouts of these platforms is that they allow scouts to identify specific technical and tactical skills, evaluate decision-making under pressure, and study the playing style of individual players to determine their suitability and potential for the team.

**Sports Analytics Platforms**

These are software designed specifically for managing and analyzing data in sports. They allow teams and analysts to study trends, playing patterns, and individual and collective performance.

These software platforms help scouts analyze complex data and visualize patterns of play and player performance over time. This helps identify trends, compare players in different technical and tactical aspects, and make informed decisions about talent selection and development.

**Sensors and Wearable Devices**

These are technologies integrated into players' uniforms or worn devices to collect biometric and performance data in real time, such as heart rate, body temperature, and sleep quality.

These tools are useful for scouts as they provide real-time biometric and performance data during evaluations and physical tests. Scouts can use this information to assess endurance, recovery, and the physical capacity of prospective players, improving the accuracy of their evaluation of athletic potential.

**Data Management Platforms**

These are systems for storing, organizing, and managing large volumes of data generated by multiple sources both on and off the field, ensuring accessibility and security of the information.

For scouts, data management platforms centralize and organize data from multiple sources (GPS tracking, video analysis, biometric data, etc.) to facilitate the comparison and tracking of prospective players over time. This allows scouts to maintain a detailed and accessible record to make data-driven strategic and tactical decisions for talent recruitment.

**Table 1. Examples of Systems by Tool Type in Big Data**



Tool type	Systems
<b>Tracking GPS</b>	<ul style="list-style-type: none"> <li>• Catapult Sports.</li> <li>• SATSports.</li> <li>• Polar.</li> <li>• Zephyr Technology.</li> <li>• Fitogether.</li> <li>• Player Maker.</li> </ul>
<b>Advanced video analysis</b>	<ul style="list-style-type: none"> <li>• Sportscode de Hudl.</li> <li>• Nacsport.</li> <li>• Wyscout.</li> <li>• InStat.</li> <li>• Metrica Sports.</li> <li>• Spiideo.</li> <li>• Coach Logic.</li> <li>• ChyronHego.</li> </ul>
<b>Sports analytics platforms</b>	<ul style="list-style-type: none"> <li>• Tableau.</li> <li>• Qlik Sense.</li> <li>• Power BI.</li> <li>• Sportlyzer.</li> <li>• TeamSnap.</li> <li>• SportsEngine.</li> <li>• Sports Performance Tracking.</li> <li>• Rapsodo.</li> <li>• Kinexon.</li> <li>• Prozone.</li> </ul>
<b>Sensors and wearable devices</b>	<ul style="list-style-type: none"> <li>• Polar.</li> <li>• Fitbit.</li> <li>• Garmin.</li> <li>• Whoop.</li> <li>• Zephyr Performance Systems.</li> <li>• Viper by STATSports.</li> <li>• Hexoskin.</li> </ul>



Tool type	Systems
<b>Data management platforms</b>	<ul style="list-style-type: none"> <li>• SportVU.</li> <li>• SAP Sports One.</li> <li>• EDGE10.</li> <li>• Kitman Labs.</li> <li>• Fusion Sport.</li> <li>• CoachMePlus.</li> <li>• AMS (Athlete Management System).</li> <li>• Zone7.</li> <li>• SportsEngine HQ.</li> <li>• MongoDB.</li> <li>• PostgreSQL.</li> <li>• Microsoft SQL Server.</li> </ul>

Source: own creation.

### 3. Data Collection: Where You Get the Information for the Report

To create a report like the one you've seen throughout this course, websites play a crucial role by providing access to detailed data necessary for creating strategic and accurate reports. These specialized websites gather a wide range of statistics, from individual player performance to comprehensive team and league analysis. They make it easier for analysts and scouts to obtain key information about goals, assists, accurate passes, among other indicators that are essential when evaluating player performance and adaptability in different contexts. It is important to note that obtaining information from grassroots football is a bit more complicated since fewer resources are available. Sometimes, it is necessary to reach out to contacts in order to obtain information. To obtain data from grassroots football categories, it is useful to use websites from different regional federations, although this information is often less specific and requires looking at each federation individually.

Here are some of the most recognized websites for obtaining the data needed for a football report:

WhoScored is a platform known for offering a wide range of detailed data and statistics on the performance of players and teams in football worldwide. Its strengths include:



- Extensive data: It provides comprehensive statistics such as goals, assists, accurate passes, successful dribbles, and more, both at the individual and team levels.
- Global coverage: It includes a wide variety of football leagues and competitions, from major European leagues such as the Premier League and La Liga to leagues from other continents like the MLS and the Chinese Super League.
- Detailed analysis: It offers advanced tools to analyze tactical performance, compare players and teams, and evaluate trends throughout the season.
- Ease of use: Its interface is intuitive and easy to navigate, allowing users to quickly access relevant information and generate personalized analysis.

However, WhoScored also has some potential downsides:

- Limitations in lower leagues: While it covers a wide range of leagues, the depth and quality of data may vary for lesser-known leagues or regional competitions, limiting its usefulness for scouts and analysts interested in less familiar contexts.
- Data accuracy: Data and statistics generated by automated algorithms may not always fully reflect a player's real performance in specific situations, which could lead to biased or inaccurate interpretations.
- Dependence on available data: The quality and accuracy of the analysis depend on the availability and reliability of the original data, which can be subject to errors or delays in updating the information.

BD Fútbol is a website known for providing a wide variety of data and detailed statistics about football. It focuses especially on European leagues and international competitions. Its strengths include:

- Detailed data: It offers exhaustive statistics such as goals, assists, accurate passes, successful dribbles, and more, both at the individual and team levels.
- Wide coverage: It includes major European leagues such as the Premier League, La Liga, Serie A, and Bundesliga, as well as international competitions like the UEFA Champions League and the Europa League.
- Advanced analysis: It provides tools for analyzing tactical performance, team effectiveness in different match situations, and comparing the performance of players and teams across seasons.
- Easy-to-use interface: It is easy to navigate, making it simple to search for and interpret data for analysts, coaches, and scouts.

However, BD Fútbol also has some potential downsides:

- Cost of access: While it provides valuable data, full access to all features and detailed analysis may require a premium subscription, which could limit its availability to users with tight budgets.



- Limitations in lower leagues: While it covers major leagues, the coverage and depth of data may be less for smaller leagues or lesser-known competitions, limiting its usefulness for scouts and analysts interested in less popular contexts.
- Dependence on available data: The quality and accuracy of the analysis depend on the accuracy and updating of the data available, which may be affected by the availability and reliability of the original data source.

Resultados Fútbol is a platform known for providing updated information and real-time football match results globally. Its strengths include:

- Real-time updates: It offers live results from football matches in various leagues and competitions around the world, allowing users to follow sports events in real time.
- Extensive coverage: It includes a wide variety of leagues and tournaments, from major European leagues like the Premier League and La Liga to international competitions like the UEFA Champions League and Copa Libertadores.
- Calendar and schedules: It provides a detailed calendar with upcoming match schedules, making it easier to plan and track sports events.
- User-friendly interface: It is easy to use and navigate, allowing users to quickly find the results and information they are looking for.

However, Resultados Fútbol also has some potential downsides:

- Limitations in statistical detail: Unlike specialized platforms such as WhoScored or BD Fútbol, the depth and detail of the statistics and performance analysis may be limited, which might not be ideal for analysts and scouts who require more detailed data.
- Lack of deep analysis: The platform mainly focuses on providing results and scores, and may lack advanced tools for tactical analysis, detailed player or team comparisons, and trends throughout the season.
- Dependence on data accuracy: The quality and accuracy of the results depend on the original data source and timely updates, which could affect the reliability of the platform in situations involving rapid changes or unexpected events.

Transfermarkt is a tool recognized for providing detailed information on player transfers, market values, and statistics in football. Its strengths include:

- Transfer data: It offers updated information on player transfers, including transfer fees, involved clubs, and contract details.
- Market valuations: It provides estimated market valuations for individual players, based on factors such as age, performance, and projection.



- Detailed statistics: It includes detailed statistics on matches played, goals scored, assists, yellow/red cards, and other relevant metrics to evaluate player and team performance.
- History and comparisons: It allows access to a player's transfer history and compare market valuations over time, making it easier to conduct historical analysis and future projections.
- User-friendly interface: The platform is easy to navigate and provides an intuitive user experience for searching players, clubs, and leagues.

However, Transfermarkt also has some potential downsides:

- Unofficial data: Although it provides estimates and valuations based on collected data, not all information may be 100% accurate or backed by official sources, which could lead to discrepancies in the data.
- Limitations in coverage of lower leagues: While it covers a wide range of leagues and competitions, the depth of data may be limited for smaller leagues or less popular regional competitions.
- Dependence on updated information: The quality and accuracy of the information depend on continuous updates from users and contributors, which can affect the reliability of real-time data.

LaPreferente is a website specialized in providing information about regional and non-professional football in Spain. Its strengths include:

- Coverage of regional football: It provides detailed coverage of non-professional football leagues and competitions in Spain, including lower categories, regional leagues, and local competitions.
- Real-time results: It offers live results from matches, scores, and updated standings, allowing users to follow sports events in real time.
- News and reports: It publishes news, reports, and interviews about teams, players, and relevant events in Spanish regional football.
- Easy-to-use interface: The platform is accessible and easy to navigate, helping users quickly find the results and information they are looking for.

However, LaPreferente also has some potential downsides:

- Focus on regional football: While it is an excellent tool for non-professional football in Spain, the coverage and depth of data may be limited for higher-level leagues or those outside Spain.
- Limitations in statistical detail: Unlike larger, specialized platforms, the depth and detail of statistics and performance analysis may be limited, which may not be ideal for analysts and scouts who require more detailed data.



- Update and data reliability: The quality and accuracy of results and information depend on constant updates from contributors and users, which could affect the platform's reliability in situations with rapid changes or unexpected events.

**Table 2. Summary Table of Different Platforms**

Web page	Strong points	Weak points
<b>WhoScored</b>	<ul style="list-style-type: none"> <li>• Extensive data.</li> <li>• Global coverage.</li> <li>• Detailed analysis.</li> <li>• Ease of use..</li> </ul>	<ul style="list-style-type: none"> <li>• Limitations in minor leagues.</li> <li>• Data accuracy.</li> <li>• Dependence on available data..</li> </ul>
<b>BDFutbol</b>	<ul style="list-style-type: none"> <li>• Detailed data.</li> <li>• Wide coverage.</li> <li>• Advanced analysis.</li> <li>• User-friendly interface.</li> </ul>	<ul style="list-style-type: none"> <li>• Access cost.</li> <li>• Limitations in minor leagues.</li> <li>• Dependence on available data.</li> </ul>
<b>Results of football</b>	<ul style="list-style-type: none"> <li>• Real-time updates.</li> <li>• Wide coverage.</li> <li>• Schedule and timings.</li> <li>• Pleasant interface.</li> </ul>	<ul style="list-style-type: none"> <li>• Limitations in statistical detail.</li> <li>• Lack of in-depth analysis.</li> <li>• Dependence on data accuracy.</li> </ul>
<b>Transfermarkt</b>	<ul style="list-style-type: none"> <li>• Transfer data.</li> <li>• Market valuations.</li> <li>• Detailed statistics.</li> <li>• History and comparisons.</li> <li>• User-friendly interface.</li> </ul>	<ul style="list-style-type: none"> <li>• Unofficial data.</li> <li>• Limitations in minor league coverage.</li> <li>• Dependence on up-to-date information.</li> </ul>
<b>La Preferente</b>	<ul style="list-style-type: none"> <li>• Regional football coverage.</li> <li>• Real-time results.</li> <li>• News and reports.</li> <li>• User-friendly interface.</li> </ul>	<ul style="list-style-type: none"> <li>• Focus on regional football.</li> <li>• Limitations in statistical detail.</li> <li>• Data updates and reliability.</li> </ul>



Source: own creation.



## References

**Bai, Z. y Bai, X.** (2021). Sports Big Data: Management, Analysis, Applications, and Challenges. *Complexity*, 2021, Article 6676297, 11 pages. <https://doi.org/10.1155/2021/6676297>

**Mayer-Schönberger, V. y Cukier, K.** (2013). *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Eamon Dolan/Mariner Books.

**McAfee, A. y Brynjolfsson, E.** (2012). Big Data: The Management Revolution. *Harvard Business Review*, 90(10), 60-68.

