Part 1.2 Summary Report

1. Unique Athletes

The dataset contains 1,287 unique athletes, providing a large and diverse sample suitable for performance analysis and comparisons.

2. Unique Teams / Sports

There are 92 distinct teams or sports represented, supporting multi-sport analytics and cross-team benchmarking.

3. Date Range of Available Data

The data spans from October 15, 2018 to October 21, 2025, giving a seven-year period ideal for long-term trend and development analysis.

4. Records by Data Source

Approx. 6.6 million total records:

- Kinexon: 4,073,754 (largest source)
- Hawkins: 2,492,372
- Vald: 51,300

Kinexon dominates the dataset, while Vald contributes far fewer entries.

5. Missing or Invalid Player Names

There are 0 missing or invalid player names, indicating strong data integrity for athlete identifiers.

6. Athletes With Data From Multiple Systems

A total of 541 athletes have data from two or more systems, allowing cross-device comparisons but also requiring attention to potential measurement differences.

Conclusion

Overall, the dataset is large, diverse, and high quality. Its long time span and multi-source composition make it well-suited for metric analysis, longitudinal modeling, and advanced performance insights in later project stages.