Data Literacy In Sports

Seminar

✓ *Sports Science Bachelor*

Please use the following subject for any email-communication „*Lecture Data Literacy| [YOUR REQUEST]*"
at sportsanalytics@uni-saarland.de

Who Should Join the Course

- 🤔📈🤾🏃 You want to **understand how data is used to analyze performance** and decision-making in sports.

- 🧠🤓 You want to **understand the limits, risks, and biases of data-driven decision-making**.

- 👩🏽‍🏫📈 You enjoy **problem solving** and are motivated to **learn modern analytical approaches** used in high-performance sport environments.

- 💡 You want to critically **interpret analytical outputs** (metrics, models, dashboards) rather than treat them as black boxes.

- 📈📊 You want to strengthen your ability to **communicate data-based insights** clearly and responsibly.

Course Summary

Data Literacy Education in Sports (Abbr. DataLiter.)			
Study Semester 1	**Cycle** Summer Semester	**Weekly Hrs.** 2	**ECTS Pts**. 4
Module Instructor	Ing. Abdelrahman H. M. Abdelsamad		
Curriculum	Elective Course Sport Science for (M.Sc. And B.Sc. program)		
Admission Prerequisites	None		
Assessments	Cognitive Competence Test–Final Examination		
Semester Workload	Total: 120 hrs • In-Person Attendance: 18 hrs • Preparation & Assignments/ Exercises: 102 hrs		

Course Overview Structure—Learning Objectives

Foundations of Data Literacy → **Data Interpretation and Evaluation** → **Applications and Transfer of Concepts**

Learning Objectives:

Foundations	Interpretation	Applications
• Understand fundamental data types and structures. • Explain the role of data, evidence, and uncertainty in data-driven decision making. • Clearly distinguish core concepts, including data science, machine learning and big data.	• Critically evaluate data sources, data quality, limitations, and potential sources of bias. • Identify common pitfalls in data interpretation. • Understand and interpret validity and reliability of analytical results at a conceptual level.	• Develop a conceptual understanding of how outputs from basic analytical and machine-learning models should be interpreted. • Transfer general data literacy principles to applied sports contexts. • Read, discuss, and critically reflect on relevant literature in sports data science.

Conceptual Course Framework

Foundations and Motivation → Importance of data literacy for modern, evidence-based decision making

Core Concepts and Models → Data Literacy, Data Science, Machine Learning, and Artificial Intelligence

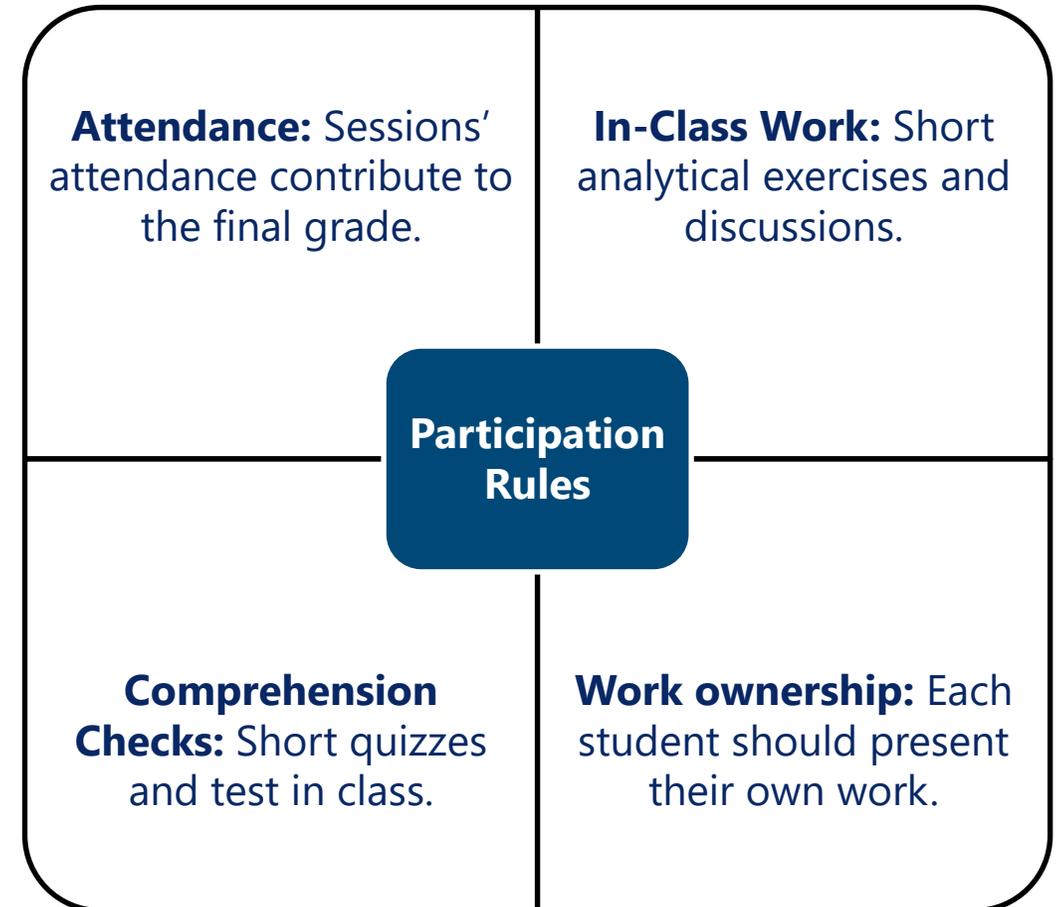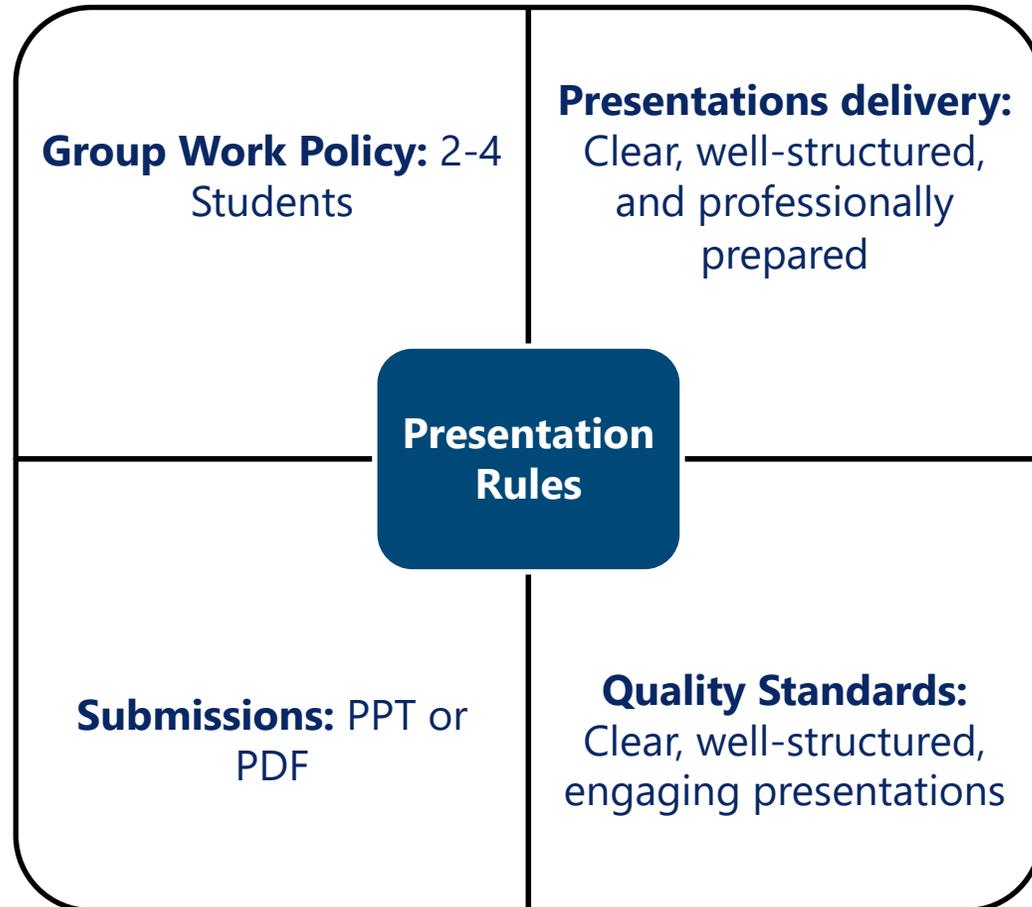**Fundamental Analytical Methods** → Regression, classification, clustering and outlier analysis.

Data Description, Evaluation, and Interpretation → Data quality, interpreting results and analyzing potential sources of error.

Application and Transfer → Data literacy principles in sports contexts and meaningful insights.

Semester Overview

Date	Topic	Presenter
10.04.2026	Introduction—Why Data Literacy? Technological Developments and Data-Driven Decision Making	Abdelrahman Hesham
17.04.2026	Introduction—Evidence-Based Decision Making and the Role of Data Across Domains	Abdelrahman Hesham
24.04.2026	Data vs. information vs. knowledge—Measurement validity and reliability	Student Groups
01.05.2026	Public Holiday (Labour Day)	—
08.05.2026	Core Definitions Data, Data Literacy, Data Science, Machine Learning, Artificial Intelligence & 4 V's	Abdelrahman Hesham
15.05.2026	Data science methods—Exploratory Data Analysis (EDA) & Handling Missing Data	Student Groups
22.05.2026	Datasets manipulations (Train–Test Splits)—Model Generalization	Student Groups
29.05.2026	Model evaluation methods (Accuracy , Precision , Recall) and F1 Score	Student Groups
05.06.2026	Key pitfalls in data science—Bias–Variance Trade-Off & Overfitting.	Abdelrahman Hesham
12.06.2026	Data Governance, Ethics and Responsible Data Usage	Student Groups
19.06.2026	A critical eye on data: selection bias and sampling bias	Student Groups
26.06.2026	Causation vs. Correlation—Interpreting Results and Avoiding False Conclusions	Student Groups
03.07.2026	Applications of Data Literacy: Case Studies with a Focus on Sport Analytics	Abdelrahman Hesham
10.07.2026	Communicating Data Insights: Reporting, Visualization, and Storytelling	Student Groups
17.07.2026	Semester Summary, Examination Information, and Course Evaluation	Abdelrahman Hesham

Course Rules

Group Work Policy: 2-4 Students

Presentations delivery: Clear, well-structured, and professionally prepared

Presentation Rules

Submissions: PPT or PDF

Quality Standards: Clear, well-structured, engaging presentations

Attendance: Sessions' attendance contribute to the final grade.

In-Class Work: Short analytical exercises and discussions.

Participation Rules

Comprehension Checks: Short quizzes and test in class.

Work ownership: Each student should present their own work.

Grading

Multiple Choice Questions Tests (20%)

- Short MCQ tests will check understanding of **core data concepts.**
- **Lowest test results will be dropped**.
- Each test graded on a **0–100-point scale**.

Attendance & Participation (20%)

- Students may not miss more than **two sessions**.

Presentation Tasks (60%)

- Evaluation based on **objective criteria**, such as:
  - ✓ Clarity, structure, and communication quality of the presentation.
  - ✓ Accuracy and depth of content.
  - ✓ Critical interpretation of data, concepts, or literature.
  - ✓ Visual quality and professionalism of slides.
- Graded on a **0–100-point scale** for each presentation.

Score between 1 and 100 =
0,2 * MCQ-Quizzes
+ 0,2 * Attendance
+ 0,6 * Presentations

~

Translation to Grades 1,0 to 5,0

General References

1. Baba, V. V., & Hakem Zadeh, F. (2012). Toward a theory of evidence-based decision making. Management Decision, 50(5), 832–867.

2. Belkin, M., Hsu, D., Ma, S., & Mandal, S. (2019). Reconciling modern machine-learning practice and the classical bias–variance trade-off. Proceedings of the National Academy of Sciences, 116(32), 15849–15854.

3. Efron, B. (1994). Missing Data, Imputation, and the Bootstrap. Journal of the American Statistical Association, 89(426), 463–475.

4. Powers, D. M. W. (2011). Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation. Journal of Machine Learning Technologies, 2(1), 37–63.

5. Price, P. C., Jhangiani, R. S., & Chiang, I.-C. A. (2015). Reliability and Validity of Measurement. Research Methods in Psychology.

6. Wolff, A., Gooch, D., Montaner, J. J. C., Rashid, U., & Kortuem, G. (2016). Creating an Understanding of Data Literacy for a Data-driven Society. The Journal of Community Informatics, 12(3).

7. Zhu, M. (2004). Recall, Precision and Average Precision. Department of Statistics & Actuarial Science University of Waterloo.

Classification (Applications in Sports)

1. Barnett, T., Reid, M., O'Shaughnessy, D., & McMurtrie, D. (2012). Game Theoretic Solutions to Tennis Serving Strategies. *ITF Coaching & Sport Science Review*, *20*(56), 22–25.

2. Bauer, P., & Anzer, G. (2021). Data-driven detection of counter-pressing in professional football. *Data Mining and Knowledge Discovery*, *35*(5), 2009–2049.

3. Bauer, P., Anzer, G., & Shaw, L. (2023). Putting team formations in association football into context. *Journal of Sports Analytics*, *9*(1), 39–59.

4. Fassmeyer, D., Anzer, G., Bauer, P., & Brefeld, U. (2021). Toward Automatically Labeling Situations in Soccer. *Frontiers in Sports and Active Living*, *3*.

5. Zhang, G., Kempe, M., McRobert, A., Folgado, H., & Olthof, S. B. (2025). Navigating team tactical analysis in football: An analytical pipeline leveraging player tracking technology. *Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology*.

Regression (Applications in Sports)

1. Wang, D., Zhang, X., Xie, Y., & Zhu, Q. (2025). Quantifying momentum and influencing factors of tennis players using the XGBoost model. Scientific Reports, 15(1), 17297.

2. Williams, L. V., Liu, C., Dixon, L., & Gerrard, H. (2021). How well do Elo-based ratings predict professional tennis matches? Journal of Quantitative Analysis in Sports, 17(2), 91–105.

3. Anzer, G., & Bauer, P. (2021). A Goal Scoring Probability Model for Shots Based on Synchronized Positional and Event Data in Football (Soccer). Frontiers in Sports and Active Living, 3.

4. Robberechts, P., & Davis, J. (2020). How Data Availability Affects the Ability to Learn Good xG Models. In U. Brefeld, J. Davis, J. Van Haaren, & A. Zimmermann (Eds.), Machine Learning and Data Mining for Sports Analytics (pp. 17–27). Springer International Publishing.

5. Scholtes, A., & Karakuş, O. (2024). Bayes-xG: Player and position correction on expected goals (xG) using Bayesian hierarchical approach. Frontiers in Sports and Active Living, 6.

Clustering (Applications in Sports)

1. Bialkowski, A., Lucey, P., Carr, P., Yue, Y., Sridharan, S., & Matthews, I. (2015). Identifying Team Style in Soccer Using Formations Learned from Spatiotemporal Tracking Data. *IEEE International Conference on Data Mining Workshops, ICDMW*, *2015*, 9–14.

2. Shelly, Z., Burch, R. F., Tian, W., Strawderman, L., Piroli, A., & Bichey, C. (2020). Using K-means Clustering to Create Training Groups for Elite American Football Student-athletes Based on Game Demands. *International Journal of Kinesiology and Sports Science*, *8*(2),

3. Soto-Valero, C. (2017). A Gaussian mixture clustering model for characterizing football players using the EA Sports' FIFA video game system. *RICYDE. Revista Internacional de Ciencias Del Deporte. Doi:10.5232/Ricyde*, *13*(49), 244–259.

4. Zhang, S., Lorenzo, A., Gómez, M.-A., Mateus, N., Gonçalves, B., & Sampaio, J. (2018). Clustering performances in the NBA according to players' anthropometric attributes and playing experience. *Journal of Sports Sciences*, *36*(22), 2511–2520.

5. Anzer, G., Bauer, P., & Brefeld, U. (2021). The origins of goals in the German Bundesliga. Journal of Sports Sciences, 39(22), 2525–2544.

Chair of Sports Analytics

Overview and Introduction

Website: https://www.uni-saarland.de/en/chair/sports-analytics.html

Who are we? Chair of Sports Analytics

High-Performance Sports

🏸 ⚽ 🚩 🏀 🎾 🏓 🤾

Objectives

🧩 **Solve Real-World Problems**

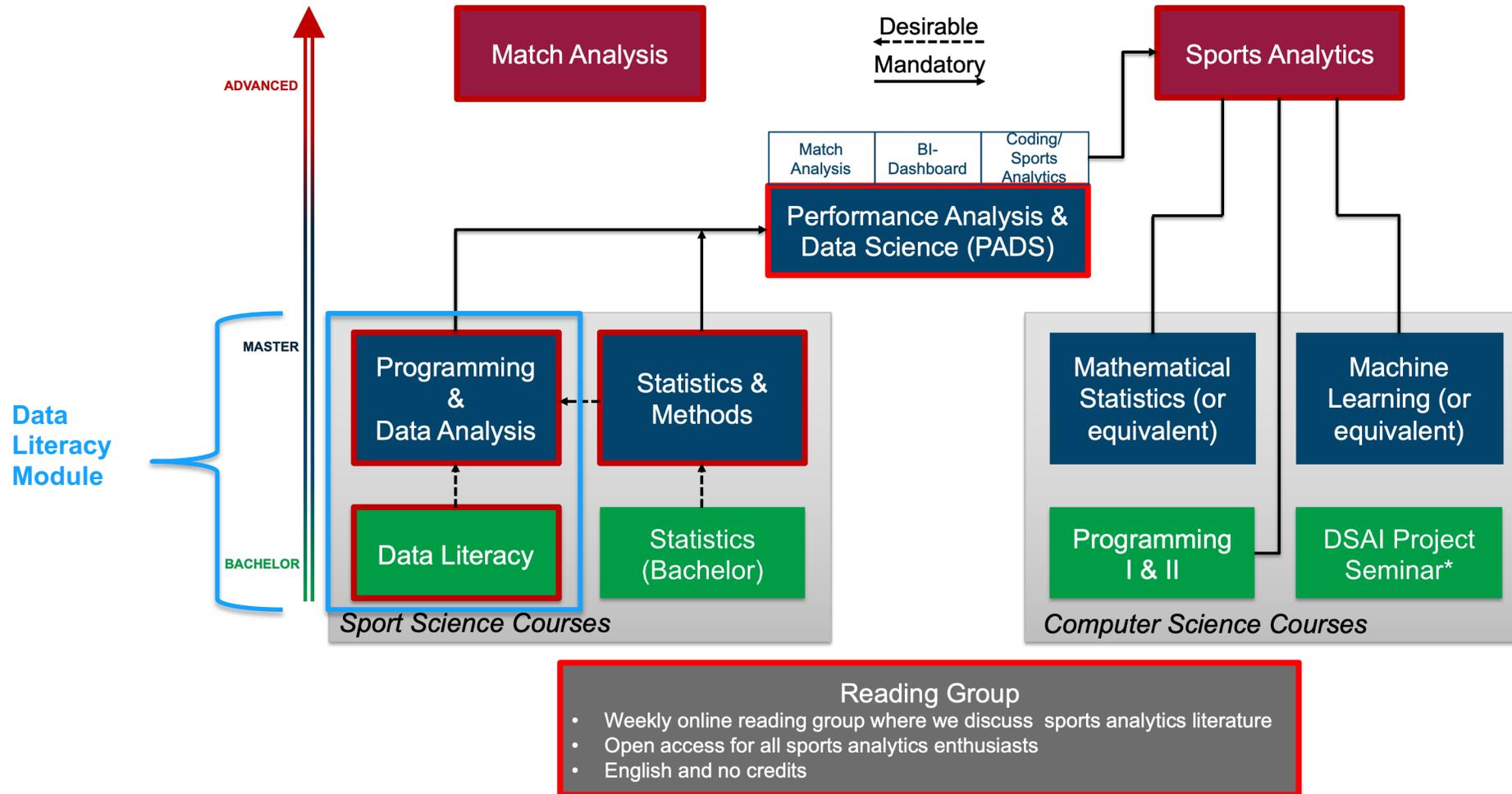💡 **Actionable Insights**

Tools & Methods

🔧 **Statistics, Data Science & Machine Learning**

📊 **Real-World Data**

Education

👩‍🏫 **Data-Literate Sport Scientists**

👨‍💻 **Data-Scientists in Sports**

Our Teaching Offers

Boxes with a red frame are teached by the chair of sports analytics

*Sports Analytics Task