DATA PREMIER LEAGUE 2026

INSMA

DATA PREMIER LEAGUE

Decode the Game. Build the Analytics.

PRELIMINARY ROUND

AUGUST 15 - DECEMBER 22, 2026

#DPL2026

Task Description

Participants must select **one of the two preliminary-round questions** and prepare a professional, evidence-based presentation supported by a recorded video explanation. The submission must be prepared using the standard presentation template provided by the organisers and accompanied by a recorded video explanation.

The presentation should contain 6-8 slides, excluding the title and reference slides where applicable. It may include the following elements:

- Context and analytical objective
- Data and methodology
- Performance overview
- Detailed analytical findings
- Strengths and weaknesses of chosen methodology
- Interpretation and practical significance
- Conclusion and key takeaways

The recorded video explanation accompanying the presentation should not be more than 6-8 minutes long.

Question 1

Assume the role of a cricket analyst working for an IPL franchise. Using the dataset provided, select one IPL team and prepare a comprehensive scouting, performance analysis and strategic review report.

All findings should be supported by appropriate statistical analysis, reproducible code and clear data visualisations.

The presentation should explain why these findings matter and what practical actions the franchise should take in response.

Question 2

Using the dataset provided, conduct a comparative analysis of the 2023 and 2024 IPL seasons to identify significant changes in team performance, player performance, tactical approaches and match patterns.

All findings should be supported by appropriate statistical analysis, reproducible code and clear data visualisations.

The presentation should explain why those changes matter and what practical lessons teams can draw from them.

Guidelines

Practical relevance: Ensure that your findings and recommendations are realistic within the context of an IPL franchise. Consider squad composition, player availability, overseas-player limits, role balance, likely recruitment constraints and the team's competitive ambitions.

Data-driven insight: Use the datasets provided to support every major conclusion. Demonstrate clearly how statistical analysis, coding and visualisation shaped your interpretation of team performance, player value, tactical patterns or trend analysis.

Reproducibility: Submit clearly organised and appropriately documented code that allows the judges to reproduce the main findings. Any custom metrics, models or additional datasets should be explained and referenced.

Communication: Present the analysis in a clear and professional manner. Visualisations should be easy to interpret, directly relevant to the argument and supported by a concise video explanation rather than simply being read from the slides.

Judging Criteria (1/4)

Criterion	70-100	50-69	40-49	0-39
Quality of Data Analysis (15%)	Demonstrates rigorous, accurate and well-justified statistical analysis. Methods are appropriate to the question and dataset, with effective comparisons, contextualisation and consideration of sample size, uncertainty and limitations. The analysis moves substantially beyond basic descriptive statistics and produces clear, reliable findings.	Demonstrates generally appropriate and accurate analysis, with suitable use of relevant metrics and comparisons. Some findings may remain descriptive, and the justification of methods, treatment of limitations or analytical depth may be inconsistent.	Uses a limited range of analytical methods, with analysis that is mainly descriptive or only partially appropriate. Comparisons may lack sufficient context, and there may be weaknesses in accuracy, methodological justification or interpretation of results.	Provides little or no meaningful analysis. Methods are inappropriate, incorrectly applied or insufficiently explained. Findings may be inaccurate, unsupported or based on a misunderstanding of the data.
Cricket Insight and Interpretation (15%)	Translates analytical results into sophisticated and meaningful cricket insights. Demonstrates a strong understanding of player roles, match phases, tactics, opposition, venue effects and game situations. Clearly explains what the findings mean beyond the headline statistics.	Provides relevant cricket interpretation and demonstrates a sound understanding of the game. Most findings are connected to tactical or performance contexts, although some interpretations may lack depth, specificity or consistency.	Demonstrates a basic understanding of cricket context, but interpretation is limited or overly dependent on conventional statistics. Connections between the results and cricket performance are weak, generalised or insufficiently developed.	Provides little or no meaningful cricket interpretation. Findings are presented without context or demonstrate significant misunderstandings of player roles, tactics, match situations or performance.

Judging Criteria (2/4)

Criterion	70-100	50-69	40-49	0-39
Practical Relevance and Recommendations (15%)	Presents clear, realistic and actionable recommendations that follow directly from the evidence. Recommendations demonstrate strong awareness of IPL squad rules, tactical requirements, player roles, recruitment constraints and franchise priorities. The likely benefits, risks and implementation considerations are clearly explained.	Provides relevant and broadly realistic recommendations supported by the analysis. The connection between evidence and action is generally clear, although some recommendations may lack detail, prioritisation or consideration of practical constraints.	Offers general or partially relevant recommendations, but their connection to the evidence is weak. Feasibility, tactical fit or practical implementation is considered only superficially.	Recommendations are absent, unrealistic, unsupported or unrelated to the analysis. The submission does not demonstrate how the findings could inform practical decisions.
Technical Execution and Reproducibility (15%)	Code is clearly documented and reproducible. Data cleaning, transformation, analysis and visualisation processes are transparent. Custom metrics, assumptions, packages and external data sources are clearly explained, enabling the principal results to be replicated.	Code is generally functional and organised, and most results can be reproduced. Documentation is adequate, although some steps, assumptions, dependencies or calculations may not be fully explained.	Code is included but is incomplete, poorly organised or difficult to reproduce. Important analytical steps are unclear, manually conducted or insufficiently documented.	Code is absent, substantially incomplete or non-functional. The analysis cannot be reproduced, and key calculations or data-processing steps are not explained.

Judging Criteria (3/4)

Criterion	70-100	50-69	40-49	0-39
Advanced Analytics and Machine Learning (10%)	Applies an appropriate and well-validated machine-learning or advanced analytical method to a clearly defined cricket problem. Model development is transparent, evaluation is rigorous, and results are interpreted effectively. The approach generates meaningful insights and supports realistic cricket decisions.	Applies a generally appropriate advanced method with reasonable model development and evaluation. The results are relevant, although validation, interpretation, comparison with baselines or practical application may lack depth.	Demonstrates a limited or partially appropriate use of machine learning. Model selection, validation or interpretation is weak, and the practical relevance of the output is not clearly established.	Machine learning is absent where required, incorrectly applied or poorly explained. The model contains major methodological weaknesses, cannot be reproduced or produces outputs with little relevance to the selected question.
Data Visualisation (10%)	Demonstrates excellent use of data visualisation to communicate complex sporting data clearly and effectively. Visualisations are accurate, purposeful, appropriately selected and enhance interpretation of the analysis.	Uses appropriate and generally clear visualisations that support the main findings. Most charts are well constructed and interpretable, although some may lack refinement, contextual benchmarks or analytical depth. Visualisations contribute meaningfully to the overall analysis.	Includes basic visualisations, but their selection, construction or interpretation is inconsistent. Charts may be overly descriptive, cluttered, poorly labelled or insufficiently connected to the analytical argument.	Visualisations are absent, inappropriate, inaccurate or difficult to interpret. Charts add little analytical value or may misrepresent the underlying data.

Judging Criteria (4/4)

Criterion	70-100	50-69	40-49	0-39
Originality and Analytical Approach (10%)	Demonstrates a highly original and purposeful analytical approach. Develops innovative questions, metrics, models or visualisations that reveal insights beyond conventional cricket analysis. Creativity strengthens rather than complicates the submission.	Shows some originality in the choice of metrics, comparisons, methods or visualisations. The approach is relevant and effective, although it may rely partly on established forms of cricket analysis.	Uses mostly conventional metrics and approaches with limited originality. Some attempt at creativity is evident, but it adds little analytical value or is insufficiently developed.	Demonstrates little or no originality. The analysis is largely copied, routine, inappropriate or lacks a coherent analytical approach.
Communication and Presentation (10%)	The presentation is clear, concise, professionally structured and persuasive. The video explanation confidently interprets the evidence, communicates the significance of the findings and remains within the required format and time limit.	The presentation is logically organised and generally easy to follow. The video explanation communicate the principal findings effectively, although some slides may be crowded, unclear or insufficiently integrated into the argument.	The presentation communicates some relevant information but lacks clarity, coherence or professional structure. The video explanation may describe rather than explain the analysis.	The presentation is incomplete, unclear or poorly organised. The video explanation does not communicate the analytical process or principal findings effectively.

Participate in the Data Premier League

INSMA

DATA PREMIER LEAGUE

Contact

+91-8981380461

info@insma.in