**1. Objective**

To test whether Large Language Models (LLMs) like ChatGPT, Claude, and Gemini show bias when analysing the same IPL player performance data but with different wording or framing of questions.

**2. Dataset**

The dataset contains IPL statistics such as runs, wickets, strike rates, and dismissals of well-known players like Virat Kohli, MS Dhoni, and Shubman Gill.

These stats include performance by year, total runs, dismissals, strike rate distribution, and all-round scores.

**3. Hypotheses**

* Framing Bias:

Asking “Which player underperformed this season?” versus “Which player shows the most growth potential for next season?” will produce different recommendations, even with identical stats.

* Confirmation Bias:

When the prompt suggests “Virat Kohli was the best player every year,” the AI might agree—even if another player had better performance in a given year.

* Selection Bias:

The model may overemphasize batting performance (runs, strike rate) while ignoring bowling or fielding statistics when identifying “top players.”

* Demographic/Status Bias:

Including labels like “senior player” for MS Dhoni or “young player” for Shubman Gill may cause different coaching or improvement suggestions.

**4. Planned Method**

Create paired prompts that differ by only one element (positive/negative phrasing, player role, etc.).

Query ChatGPT, Claude, and Gemini with each prompt version.

Collect 3–5 responses per prompt to control for randomness.

Log all prompts, timestamps, and outputs.

Analyze:

Which players are mentioned most often.

Sentiment/tone using Python (Pandas).

Whether AI recommendations change with prompt framing.