

# Nepal Premier League (NPL) 2024: Comprehensive Data Analysis Report

---



## **Author(s):**

Ajay BK

Prabesh Khatiwada

Ashutosh Dahal

## **Executive Summary**

The inaugural Nepal Premier League (NPL) 2024 marked a significant milestone in Nepali cricket, introducing a franchise-based T20 format with eight regional teams. This report analyzes detailed ball-by-ball, match, and player data from the tournament, providing insights into team performances, player statistics, and key trends. Through rigorous data cleaning, exploratory analysis, and advanced visualizations, we identify factors influencing match outcomes, highlight standout players, and offer actionable recommendations for league organizers and team management. Key findings include the critical impact of toss decisions, the dominance of certain teams and players, and areas for improvement such as reducing extras and optimizing team strategies. Recommendations focus on talent development, tactical planning, and leveraging data-driven insights for future editions.

## **Table of Contents**

1. Executive Summary
2. Introduction
3. Data Description
4. Data Cleaning and Preprocessing
5. Exploratory Data Analysis (EDA)
  - Team Participation and Distribution
  - Points Table and Net Run Rate
  - Toss Analysis
  - Batting and Bowling Trends
  - Player Performance
  - Top 15 Nepali Players Based on Performance
  - Shot Direction and Batting Style
  - Match Progression
6. Results and Evaluation
7. Discussion
8. Conclusion and Recommendations
9. References

## Introduction

### Background and Context

The Nepal Premier League (NPL) 2024 was the first franchise-based Twenty20 cricket league in Nepal, organized by the Cricket Association of Nepal. The tournament aimed to elevate the standard of domestic cricket, attract international talent, and foster regional representation.

### Problem Statement and Objectives

The primary objective of this project is to analyze the NPL 2024 season using comprehensive datasets to uncover performance patterns, evaluate player contributions, and provide strategic insights for stakeholders. Specific goals include:

- Summarizing team and player performances
- Identifying trends and key factors affecting match outcomes
- Recommending actionable strategies for future tournaments

### Scope and Limitations

The analysis covers all 32 matches of the 2024 season, focusing on quantitative data (match, ball-by-ball, and player statistics). Limitations include lack of qualitative factors (e.g., weather, pitch reports), limited access to certain off-field variables, and the inaugural nature of the league, which may affect trend generalizability.

---

## Data Description

### Source(s) of the Data

- Kaggle: Nepal Premier League (NPL) 2024 Cricket Dataset
- Official match records and player lists

### Description of Data Fields and Types

- **matches.csv:** Match-level data (teams, scores, toss, winner, overs, etc.)
- **balls.csv:** Ball-by-ball data (match\_id, inning, batsman, bowler, runs, extras, shot direction, dismissal, etc.)
- **player.csv:** Player information (name, role, team, nationality)

### Initial Data Exploration and Summary Statistics

- 32 matches played, 8 teams participated
- Over 10,000 ball-by-ball records
- Player dataset includes roles (batsman, bowler, all-rounder, wicketkeeper) and nationalities

**Sample Data Fields:**

Field	Type	Description
match_id	Integer	Unique identifier for each match
team_1, team_2	String	Names of competing teams
batsman, bowler	String	Player names
batsman_runs	Integer	Runs scored on a ball
wide_runs, noball_runs	Integer	Extras per ball
player_dismissed	String	Name of dismissed player (if any)
shot_direction	String	Coded field for shot placement

---

**Data Cleaning and Preprocessing**

**Steps Taken to Clean the Data**

- **Null Values:** Removed or imputed missing values in key columns (e.g., runs, wickets, player names).
- **Data Types:** Converted overs from string to float for accurate calculations.
- **Outliers:** Checked for and addressed anomalous entries (e.g., negative runs, duplicate records).
- **Consistency:** Standardized team and player names across datasets.

**Justification for Each Cleaning Step**

- Ensured accurate aggregation and analysis of statistics.
- Prevented skewed results due to data entry errors.
- Enabled seamless merging of datasets for multi-table analysis.

## Data Transformation and Feature Engineering

- Created fields for "batting\_first" and "batting\_second" using toss decisions.
  - Calculated Net Run Rate (NRR) and points for team rankings.
  - Engineered summary statistics (e.g., total runs, wickets per player, extras per bowler).
- 

## Exploratory Data Analysis (EDA)

### Team Participation and Distribution

All 8 teams played an equal number of matches, with balanced representation across Nepal's provinces.

### Points Table and Net Run Rate

Team	Points	NRR
Sudurpaschim Royals	12	1.445
Janakpur Bolts	10	0.697
Chitwan Rhinos	8	0.096
Kathmandu Gurkhas	8	-0.107
Karnali Yaks	8	-0.372
Biratnagar Kings	4	-0.445
Pokhara Avengers	4	-0.776
Lumbini Lions	2	-0.502

### Toss Analysis

- Teams winning the toss and choosing to bat or bowl had varying success rates.
- Bar and pie charts illustrate win percentages based on toss decisions and team-wise toss performance.

### Batting and Bowling Trends

- Top 5 team scores in a single innings ranged from 170 to 200+ runs.
- Top batsmen and bowlers identified by total runs and wickets, respectively.
- Extras (wides, no-balls) contributed significantly to total runs in several matches.

## Player Performance

- Standout players: Lahiru Milantha (most runs), Scott Kuggeleijn and Lalit Rajbanshi (most wickets).
  - "Player of the Match" awards concentrated among a few high-performing individuals.
- 

## Top 15 Nepali Players Based on Performance

The NPL 2024 provided a platform for domestic talent to shine alongside international stars. Rankings are based on aggregate performance metrics from the NPL 2024 season, considering both batting and bowling contributions, as well as match impact.

## Shot Direction and Batting Style

- Pie and bar charts reveal preferred shot directions for top batsmen, highlighting tactical tendencies.

## Match Progression

- Interactive line plots show run accumulation and wicket falls over each match, providing insights into momentum shifts.
- 

## Results and Evaluation

### Evaluation Metrics

- Teams and players were evaluated using cricket-specific metrics:
  - Team: Points, NRR, win/loss ratios, toss win percentage.
  - Player: Total runs, wickets, match-winning contributions, extras conceded.

### Comparison Between Teams and Players

- Sudurpaschim Royals and Janakpur Bolts emerged as the top teams, with consistent performances and superior NRR.
- Lahiru Milantha (Janakpur Bolts) led in total runs, while Scott Kuggeleijn (Sudurpaschim Royals) and Lalit Rajbanshi (Janakpur Bolts) topped the wicket charts.

## **Visualizations**

- Bar charts: Team match counts, player runs/wickets, extras.
- Pie charts: Toss decision outcomes, shot direction distributions.
- Line plots: Match run progressions.
- Interactive dashboards: Player shot preferences, match-by-match performance.

## **Interpretation of Results**

- Toss decisions influenced match outcomes but were not deterministic.
  - Teams with balanced contributions from both domestic and overseas players performed better.
  - Extras (wides, no-balls) often shifted momentum, underlining the need for bowling discipline.
- 

## **Discussion**

### **Implications of Findings**

- The NPL's franchise model fostered competitive balance and regional pride.
- Data-driven insights can inform team selection, in-game tactics, and player development programs.
- The presence of international stars raised the league's profile and provided valuable experience for local players.

### **Limitations**

- Lack of contextual variables (pitch, weather, injuries) may obscure certain performance drivers.
- As with the first season, historical comparisons are limited.



- Potential data entry inconsistencies, though minimized through cleaning.

### **Potential Biases and Mitigation**

- Home advantage and scheduling effects not fully accounted for.
  - Player roles and match contexts (e.g., pressure situations) require deeper qualitative analysis.
- 

## **Conclusion and Recommendations**

### **Summary of Key Findings**

- Sudurpaschim Royals and Janakpur Bolts dominated the league, driven by strong batting and bowling units.
- Individual brilliance (e.g., Lahiru Milantha, Scott Kuggeleijn) often determined match outcomes.
- Extras were a recurring issue for several teams, impacting tight contests.
- Nepali players demonstrated significant talent and impact, as highlighted in the Top 15 Nepali Players analysis.

### **Actionable Recommendations**

- **For Teams:**
  - Invest in bowling discipline to reduce extras.
  - Leverage data on shot directions and bowler weaknesses for tactical planning.
  - Foster balanced squads with both experienced and emerging talents.
- **For Organizers:**
  - Continue to attract international players to elevate standards.
  - Enhance data collection (e.g., fielding metrics, player fitness) for deeper analysis.
- **For Future Work:**
  - Incorporate predictive modeling (e.g., win probability, player performance forecasts).
  - Analyze qualitative factors (e.g., leadership, crowd support) for holistic insights.

---

## References

- Kaggle: Nepal Premier League (NPL) 2024 Dataset. <https://www.kaggle.com/datasets>
- ESPN Cricinfo: NPL 2024 Stats and Records. <https://www.espncriinfo.com/>
- Wikipedia: 2024 Nepal Premier League.  
[https://en.wikipedia.org/wiki/2024\\_Nepal\\_Premier\\_League](https://en.wikipedia.org/wiki/2024_Nepal_Premier_League)
- Official NPL Team and Player Announcements