

**Exploratory Data Analysis (IPL)**

**Submitted By-**

**Krish Garg**

**(MABSPG24042)**

**IPL Exploratory Data Analysis (EDA)**

# 1. Introduction-

This report presents an Exploratory Data Analysis (EDA) of an Indian Premier League (IPL) dataset. The objective is to uncover patterns, trends, and insights from the data, preparing it for predictive modeling and deeper analysis.

# 2. Objectives-

The primary objective of this EDA is to:  
- Understand the structure and contents of the IPL dataset.  
- Identify key factors influencing match outcomes.  
- Discover patterns and anomalies in the data through statistical measures and visualizations.  
- Clean and prepare the dataset for further modeling.

# 3. Research Methodology-

The research approach includes:  
- Data loading and cleaning.  
- Univariate, Bivariate, and Multivariate analysis.  
- Visualization of distributions and relationships using plots.  
- Summary statistics and correlation studies.

# 4. Data and Variables-

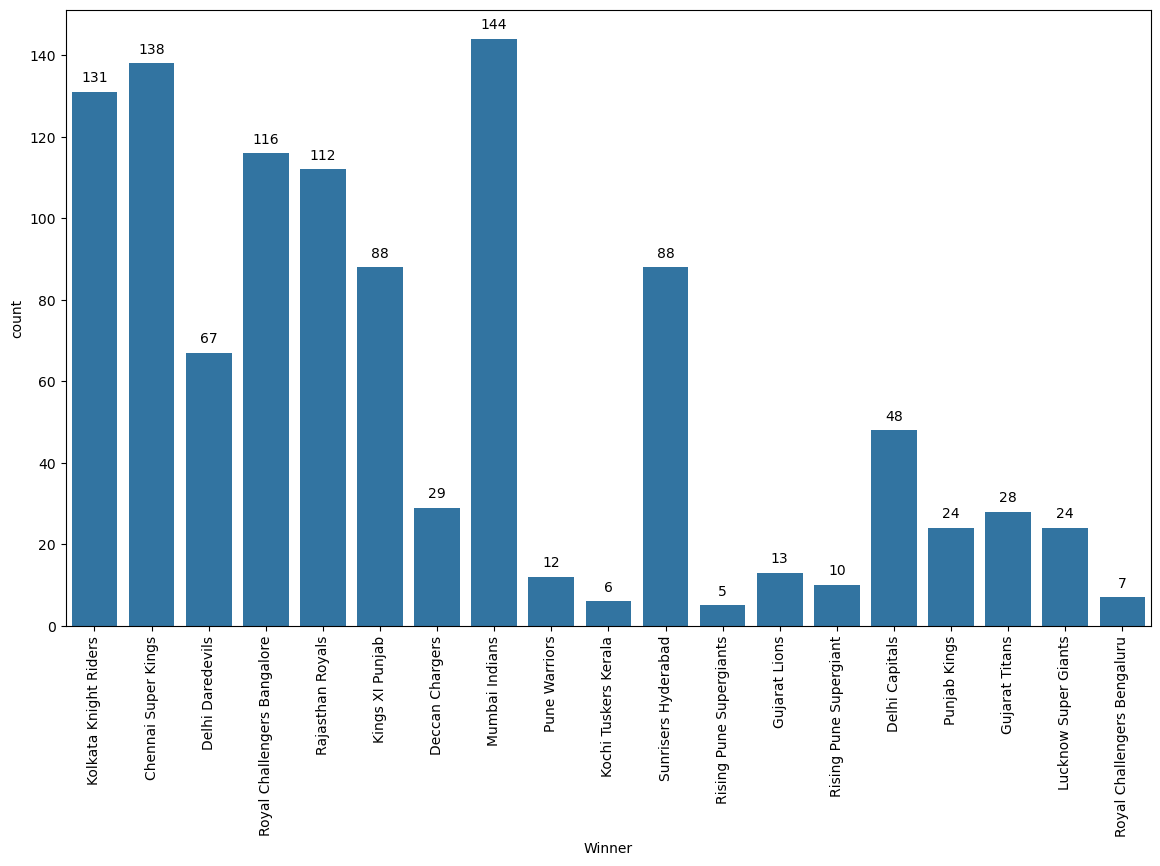
The dataset contains records from multiple IPL seasons with variables such as:  
- Team names  
- Player names  
- Venue details  
- Toss decision and winner  
- Match outcomes  
- Umpires, City, and Season  
- Various performance statistics

# 5. Statistical Tools and Techniques-

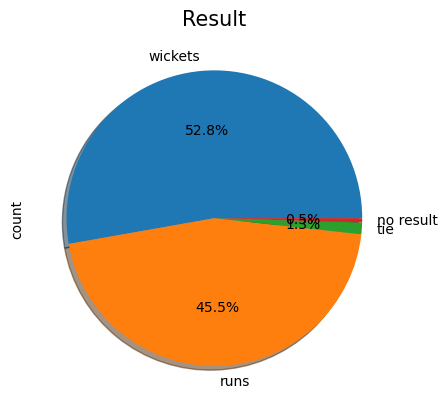
The following tools and techniques are employed:  
- Handling missing values and duplicates.  
- Descriptive statistics: mean, median, mode, standard deviation, and variance.  
- Visualizations: histograms, boxplots, bar plots, pie charts, heatmaps.  
- Correlation analysis.

# 6. Findings and Visualizations-

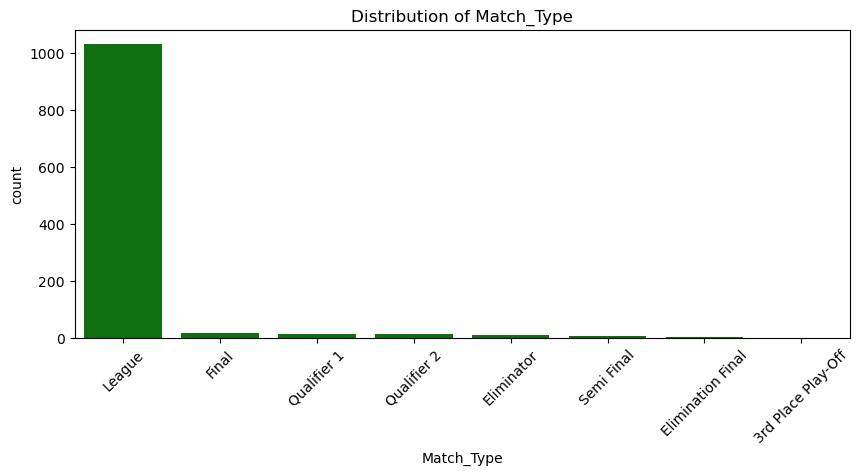
**Q.1) How Many Matches won by the Teams?**

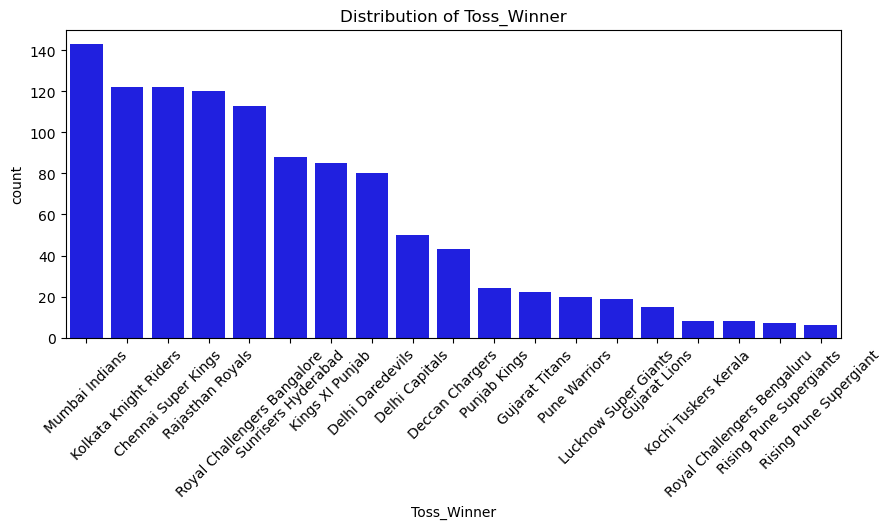


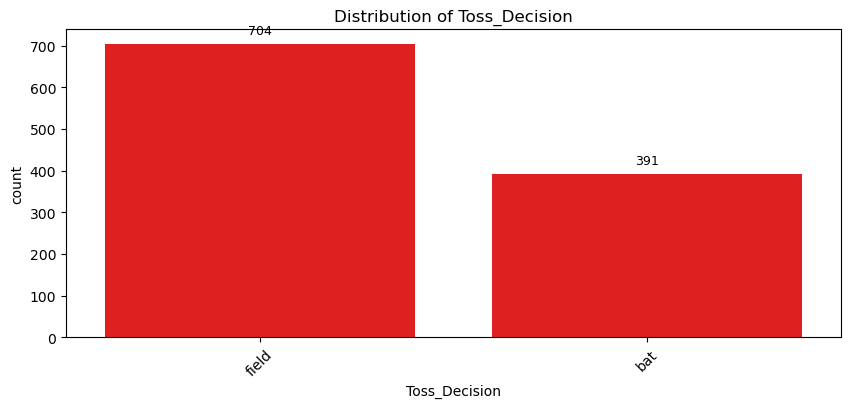
**Q.2) What is the Percentage of the different ways of winning a match?**



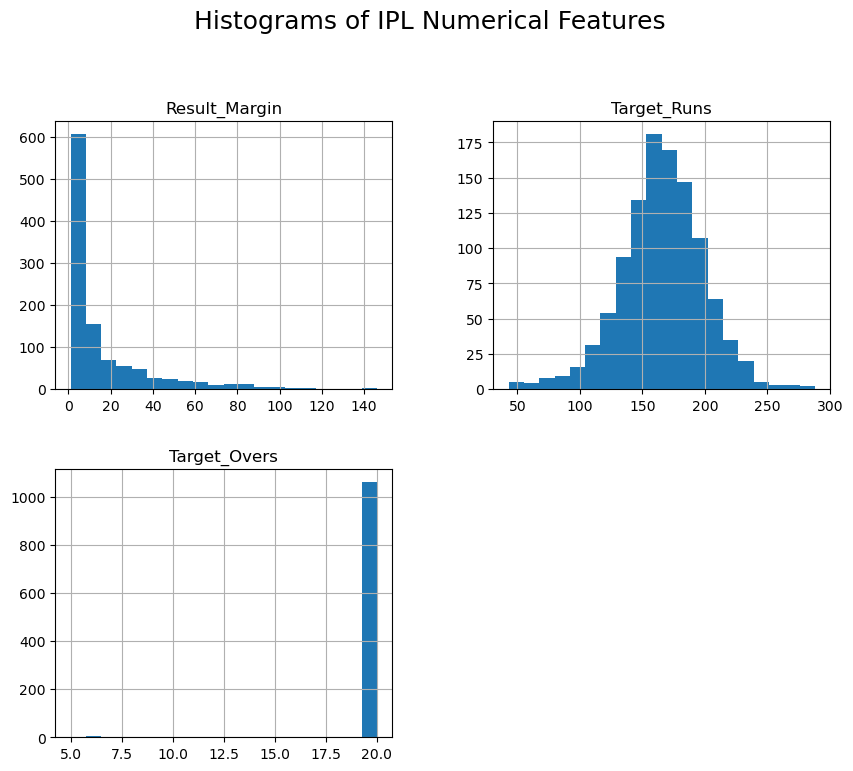
**Q.3) What is the Distribution of Match Type, Toss Winner and Toss Decision?**



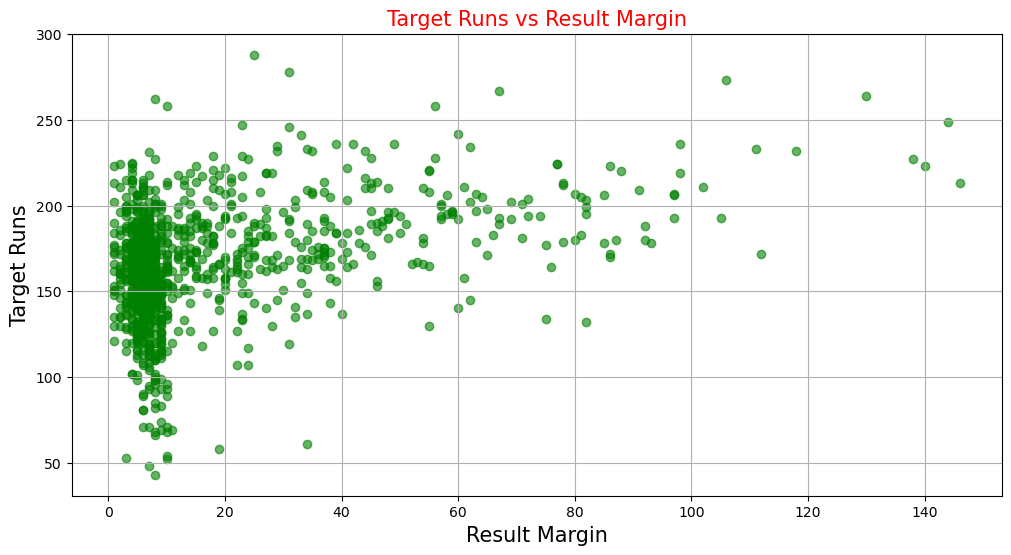




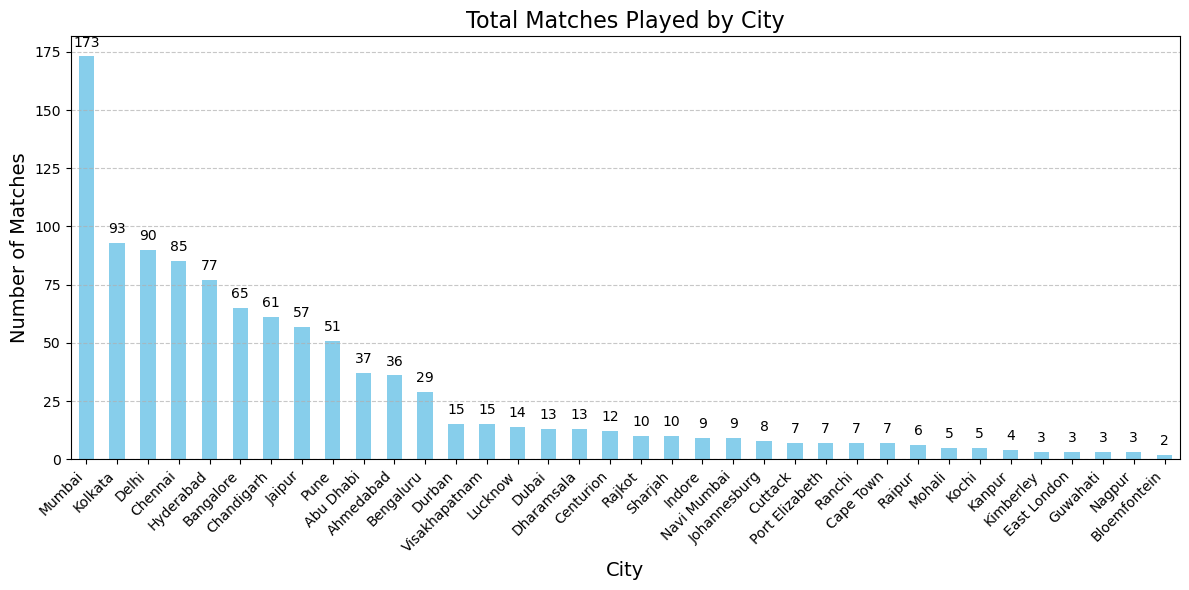
**Q.4) What is the Distribution of Result Margin, Target Runs and Target Overs?**



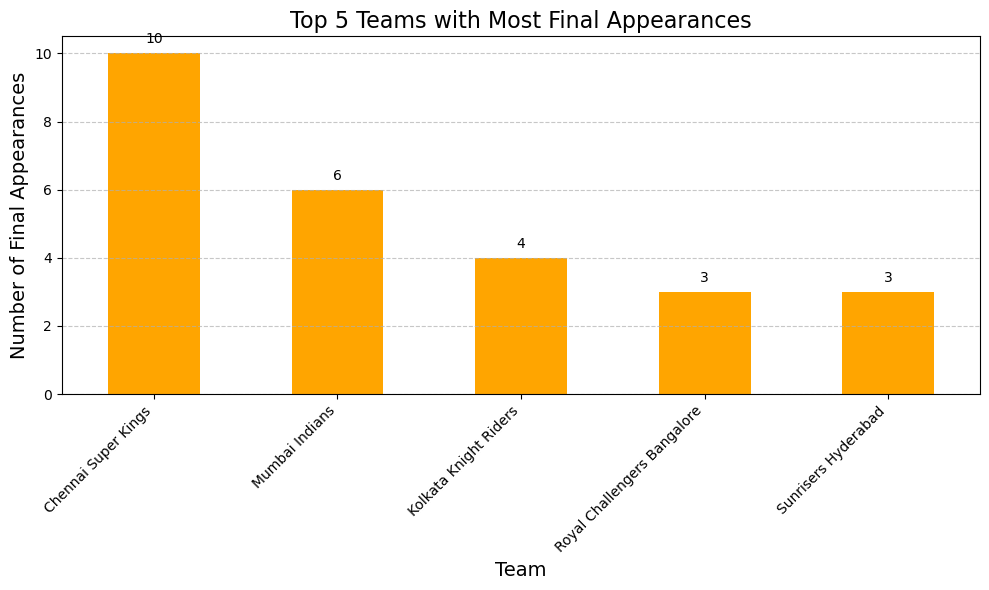
**Q.5) What is the relation between Target Runs vs Result Margin?**



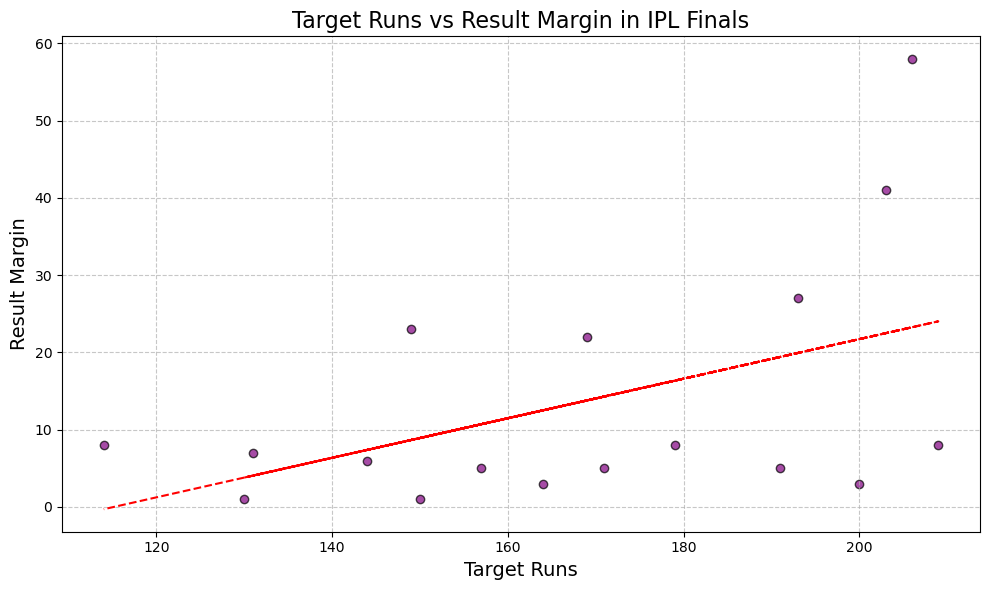
**Q.6) What is the Total Matches Played by City?**



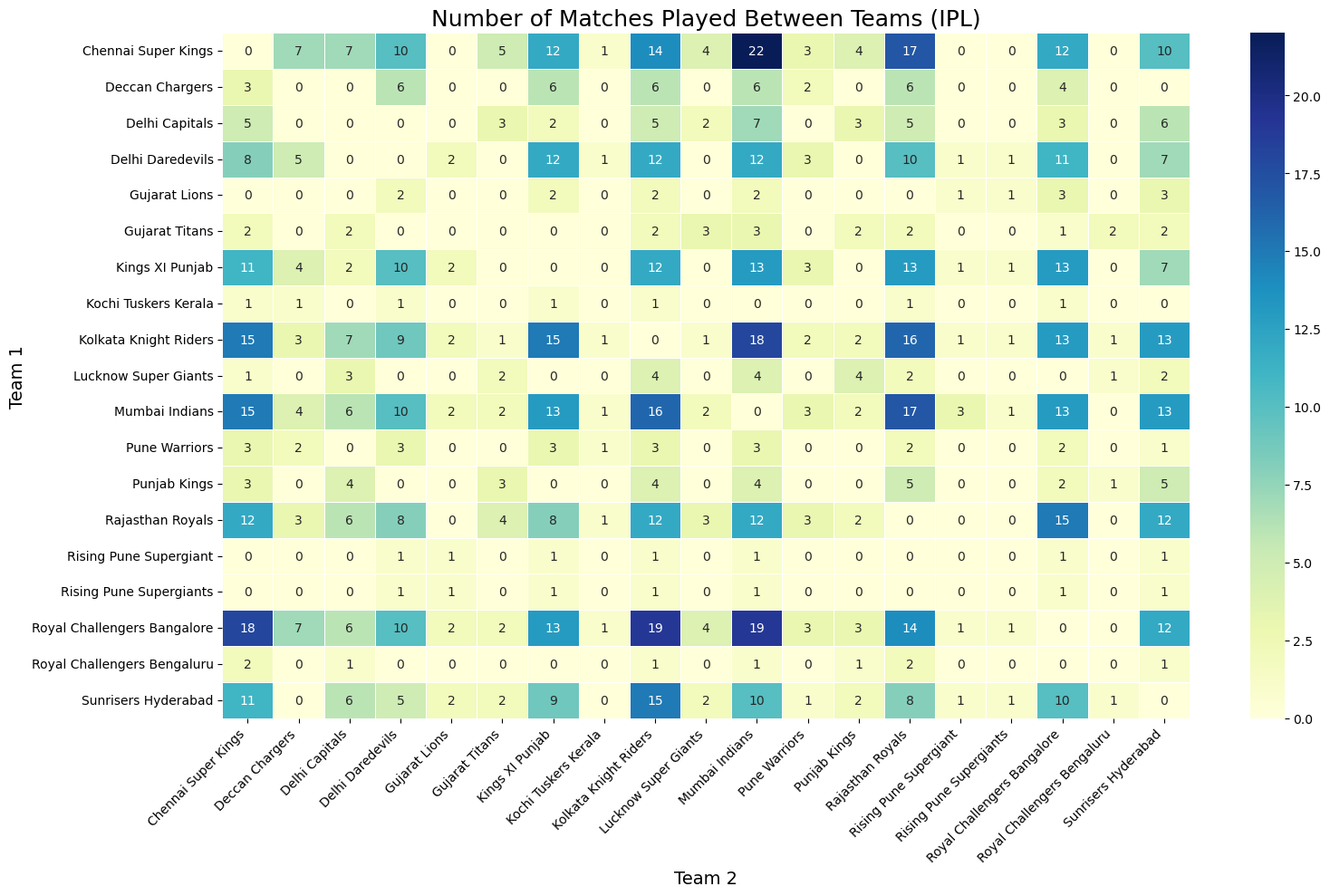
**Q.7) What are the number of top 5 teams with most no. of final appearances?**



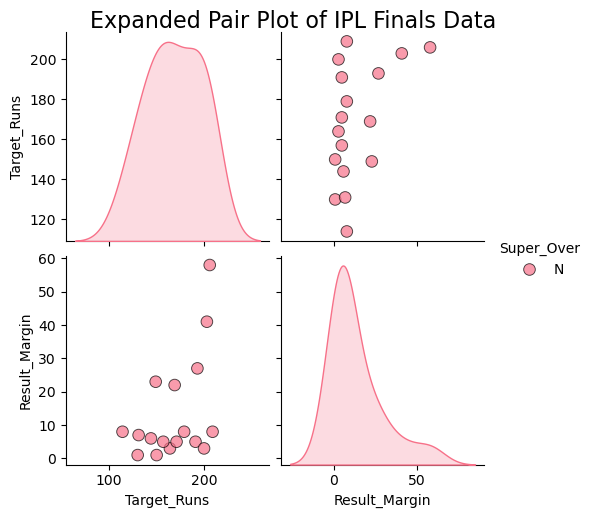
**Q.8) What is the Correlation between Target Runs vs Result Margin in IPL Finals?**



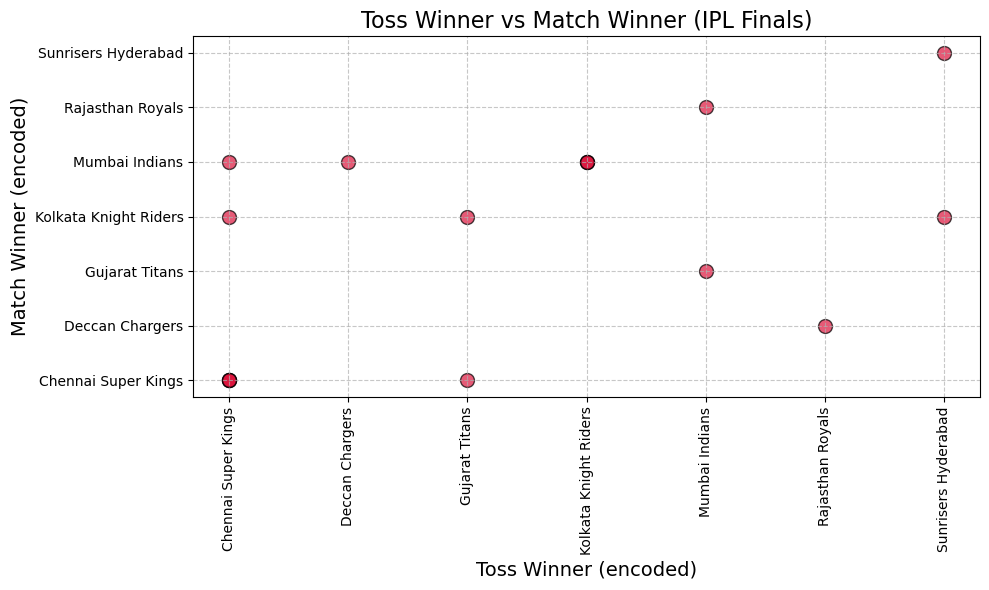
**Q.9) What are the no. of matches played between IPL Teams?**



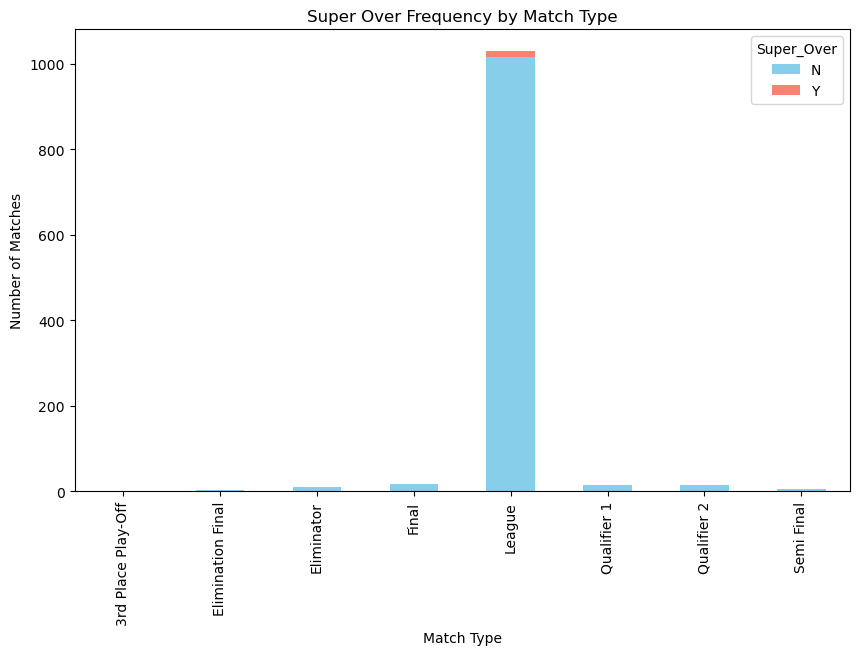
**Q.10) What is the relation between Target Runs, Result Margin, Season, Super Over and Method?**



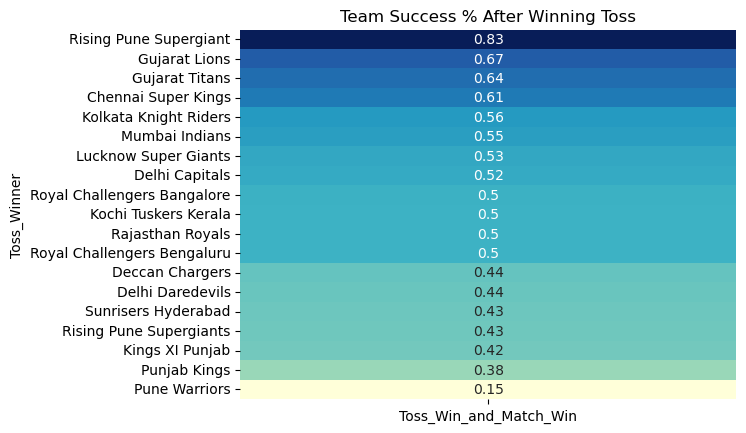
**Q.11) What is the relation between Toss Winner v/s Match Winner (IPl FInal)?**



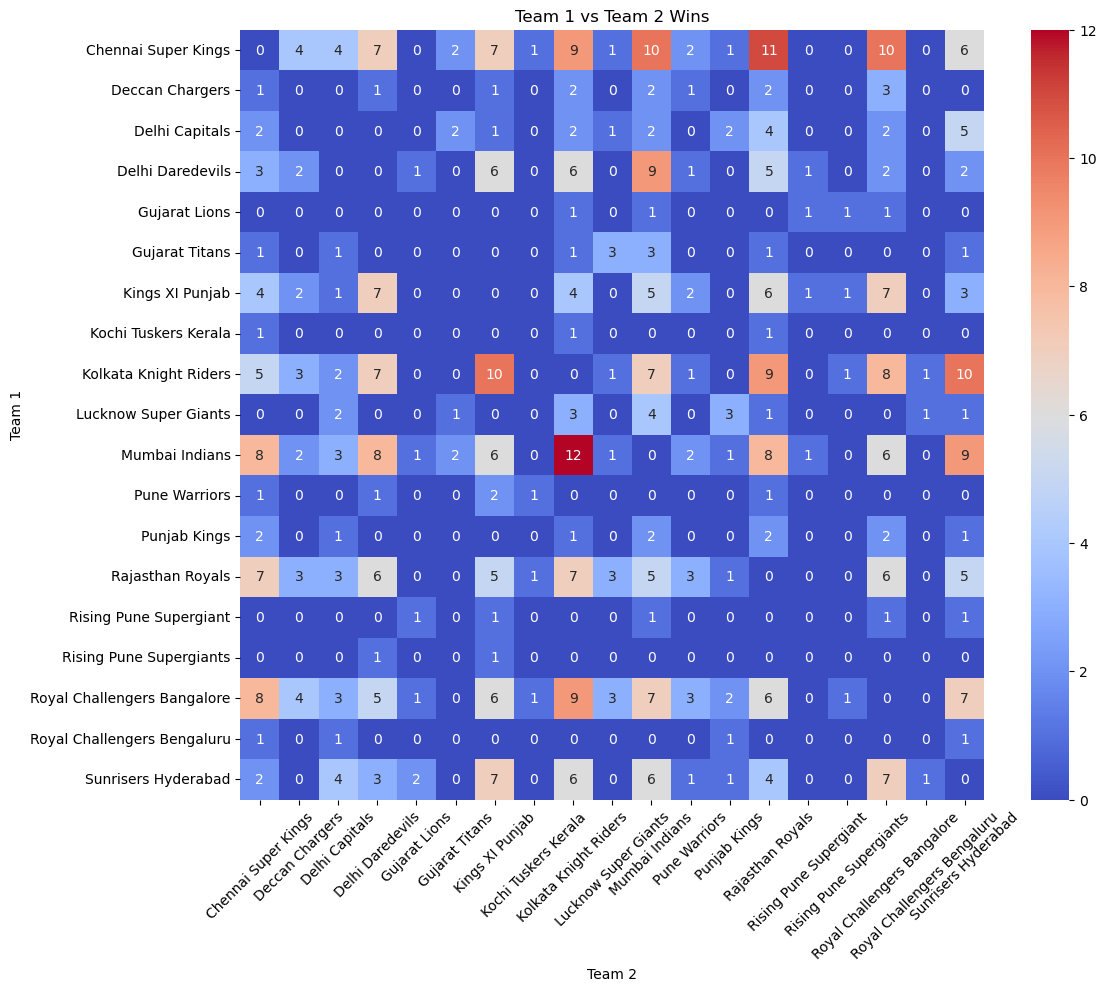
**Q.12) How can you create a stacked bar chart to show the frequency of Super Over occurrences by match type?**



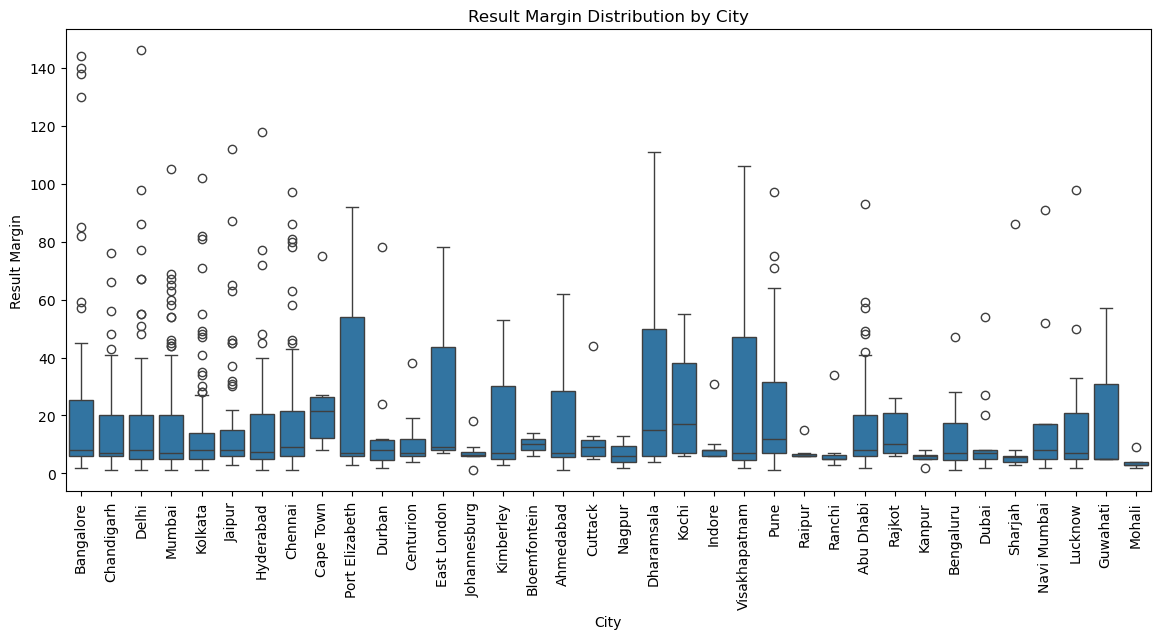
**Q.13) What is the Performance of Teams after Winning Toss?**



**Q.14) What is the relation between Teams Vs Team Win % ?**



**Q.15) What is the relation between City vs Result Margin?**



# 7. Conclusion

The EDA provides a comprehensive overview of the IPL dataset, highlighting key trends such as the importance of toss decisions, the dominance of certain teams, player performance distributions, and venue-specific advantages. These insights are instrumental for building predictive models and making strategic decisions for upcoming matches.