

PGPDSE FT Mini-Project – DBMS

Pie-in-the-Sky

IPL Match Bidding App

Participant List

1. Anurag S. Sawarkar
2. Soham Gotmare
3. Vishwa Sanghavi
4. Dennnis Jacob
5. Shubham Chopde
6. Aman Vishwas

Contents

Industry Review	3
Literature Survey	6
Dataset and Domain	8
Data Dictionary and Dataset Description	8
Business Importance	19
Business Questions Discussion	21
Conclusion	33
References	34
Industry Review	34
Literature Survey	34

Industry Review

According to the data creation statistics for 2022, a mind-boggling 2.5 quintillion bytes of data are generated on a daily basis. The sheer magnitude of this data explosion has granted individuals access to an abundance of information, empowering them to make more informed decisions. Consequently, it comes as no surprise that data has assumed an increasingly influential role across almost every industry. One domain that has particularly flourished in this new era is sports analytics.

For anyone involved in the realm of sports broadcasting, it is evident that data and statistics play an extensive role in the sports industry today. A growing number of sports enthusiasts and professionals are keen on harnessing the power of analytics to gain a statistical advantage, thus giving rise to the thriving sports analytics industry.

Sports Analytics

Sports analytics involves the comprehensive examination of athletic performance and the overall health of a sports organization to enhance its processes and achieve success. This field encompasses two primary components:

- Analysis of on-field data: This aspect revolves around tracking essential on-field data metrics to shape methodologies that can enhance in-game strategies, nutrition plans, and other crucial factors that can ethically elevate athletes' performance levels. It aims to address inquiries related to on-field performance, such as identifying the soccer player with the highest chance creation in Europe or determining the fastest runner over a 20-meter distance.
- Analysis of off-field data: This facet concentrates on the business aspects of sports. It involves monitoring significant off-field data metrics, including ticket sales, merchandise sales, fan engagement, and more. The objective of this data analytics approach is to assist decision-makers within sports teams in making informed choices that drive growth and profitability.

The integration of data analytics has become essential in modern business practices. Numerous organizations across diverse industries have harnessed the power of data analytics to optimize performance. By incorporating data analytics into their business models, these entities now possess valuable insights that enable them to make informed decisions and streamline cost-efficiency.

Data Analytics In IPL

The Indian Premier League (IPL) has redefined cricket in India, attracting top players worldwide and becoming a prominent event in the cricketing sphere. Within this high-stakes environment, the integration of data analytics has emerged as a vital component for teams' preparation and performance analysis. In this article, we will delve into the utilization of data analytics in the Indian Premier League.

Applications in IPL

1. Player Selection and Auction

Prior to the commencement of IPL, teams are tasked with assembling their player squads through auctions. Team owners, coaches, and management heavily rely on data analytics to make informed decisions regarding player selection. They scrutinize player statistics, injury records, and playing styles to curate an optimal combination of players that aligns with their team's gameplay strategy. Furthermore, data analysis aids in evaluating the performance of domestic players in various tournaments, facilitating the identification and recruitment of emerging talent.

2. Performance Analysis

Throughout the IPL, teams amass an extensive volume of data pertaining to player performance. This encompasses ball-by-ball data, player movements, and various other data points. Such data is then utilized to analyze player performance and devise strategies for improvement. Statistical models are employed to identify individual

players' strengths and weaknesses, allowing teams to fine-tune their game plans accordingly. Data analysis further facilitates the identification of patterns in the opposition's tactics, enabling the formulation of counter-strategies. This comprehensive approach not only assists players in enhancing their skills but also elevates the overall competitiveness of the sport.

3. Augmented Coaching and Data-Driven Analysis

Wearable technology has revolutionized the realm of IPL coaching. Teams employ wearables to monitor players' movements and optimize their preparation. For instance, physiotherapists can track vital signs like heart rate and breathing rate during training, pinpointing areas for improvement. Wearables also aid in monitoring players' fitness levels, helping prevent injuries. Augmented reality empowers coaches to provide players with tangible feedback on technique enhancement, facilitating data-driven decision-making during matches.

4. Fan Engagement

Data analytics plays a pivotal role in engaging fans and crafting immersive experiences. IPL teams utilize data analytics to comprehend fan behavior, preferences, and purchasing patterns. This wealth of information is then utilized to create targeted marketing campaigns that enhance the overall fan experience, fostering sustained engagement with the game.

5. Broadcasting

Broadcasters of the IPL harness data analytics to enrich the viewing experience for fans. By leveraging diverse data points such as estimated scores, run rate progression, ball trajectory predictions, and insightful statistics, the televised coverage offers viewers a deeper understanding of the game. These data-driven elements contribute to an enhanced viewing experience for cricket enthusiasts.

Literature Survey

Article Name: A team recommendation system and outcome prediction for the game of cricket

Publisher: Journal of Sports Analytics 4 (2018) 263–273 DOI 10.3233/JSA-170196 IOS Press

Author: Sandesh Bananki Jayanth^{a,b}, Akas Anthony^b, Gududuru Abhilasha^b, Noorni Shaik^b and Gowri Srinivasaa,^b

^aDepartment of Computer Science and Engineering, PESIT Bangalore South Campus, Bengaluru, India

^bPES Center for Pattern Recognition, PESIT Bangalore South Campus, Bengaluru, India

Year: 2018

Abstract:

Predicting the outcome of a game using players strength and weakness against the players of the opponent team by considering the statistics of a set of matches played by players helps captain and coaches to select the team and order the players. In this paper, we propose a supervised learning method using SVM model with linear, and nonlinear poly and RBF kernels to predict the outcome of the game against particular side by grouping the players at different levels in the order of play for both the teams. The comparison among different groups of players at same level gives the order of groups which contributes to winning probability. we also propose to develop a system which recommends a player for a specific role in a team by considering the past performances. we achieve this by finding the similar players by clustering all the players using k-means clustering and finding the five nearest players using k nearest neighbor (KNN) classifier. We calculate the ranking index for players using the game and players statistics extracted from a particular tournament. Experimental results demonstrate that, the n-dimensional data considered for modeling is not linearly separable. Hence the nonlinear SVM with

RBF kernel outperforms from the linear and poly kernel. SVM with RBF kernel yields the accuracy of 75, precision of 83.5 and recall rate of 62.5. So we recommend the use of SVM with RBF kernel for game outcome prediction.

Keywords: Cricket, sports analytics, game outcome prediction, recommendation system, performance quantification, Support Vector Machine (SVM)

Conclusion:

In this paper, we have proposed a model for cricket match outcome prediction, team structure analysis and player recommendation system using the statistics of the players extracted from a particular tournament. We have devised a method to find the player ranking system using players statistics in terms of tournament statistics. These ranking of players are used to model the SVM for match outcome prediction and team structure analysis. SVM model is trained using linear, poly and RBF and during the testing phase, it was shown that the accuracy, precision and recall rate for SVM with RBF had outperformed from other two methods. The results indicate that our data set is not linearly separable. So we intend to use the SVM with RBF method for match outcome prediction. In our player recommendation system, we have used player performance measures to cluster all the players using k-means clustering and similar players are found using k-nearest neighbour classifier.

Dataset and Domain

Data Dictionary and Dataset Description

The given dataset has 12 tables. The data dictionary and its description as follows.

Table 1: IPL_User

Data Dictionary:

<i>Variable</i>	<i>Data type</i>	<i>Comments</i>
<u>UserId</u>	VARCHAR	Primary Key - Unique id for the User
Password	VARCHAR	Password to for the User
User_type	VARCHAR	Admin or Bidder
Remarks	VARCHAR	Only null values

Dataset Description:

The IPL_User table contains information about registered users for the Pie-in-the-Sky app. It includes user IDs, passwords, user types, and optional remarks.

Table 2: IPL_Stadium

Data Dictionary:

<i>Variable</i>	<i>Data type</i>	<i>Comments</i>
<u>StadiumId</u>	NUMBER	Primary key - Unique ID's for stadium
Stadium_name	VARCHAR	Unique, Not null - Name of the Stadium
City	VARCHAR	City name of the Stadium present
Capacity	NUMBER	Capacity of the stadium
Address	VARCHAR	Address of the stadium
Contact_no	NUMBER	Contact number of the stadium

Dataset Description:

The IPL_Stadium table stores details about the stadiums where IPL matches take place. It includes information such as stadium IDs, names, cities, capacity, addresses, and contact numbers.

Table 3: IPL_Team

Data Dictionary:

<i>Variable</i>	<i>Data type</i>	<i>Comments</i>
<u>TeamId</u>	NUMBER	Primary key - Unique team id for all the teams
Team_name	VARCHAR	Unique, Not null - Team name for all the teams
Team_city	VARCHAR	Origin city of the team
Remarks	VARCHAR	Short form of the team names

Dataset Description:

The IPL_Team table contains information about the IPL teams participating in the tournament. It includes team IDs, names, origin cities, and optional remarks.

Table 4: IPL_Player

Data Dictionary:

<i>Variable</i>	<i>Data type</i>	<i>Comments</i>
<u>PlayerId</u>	NUMBER	Primary key
Player_name	VARCHAR	Unique, Not null
Performance_details	VARCHAR	Performance details
Remarks	VARCHAR	Remarks like a top performer or second best performer

Dataset Description:

The IPL_Player table stores details about the players participating in the IPL. It includes player IDs, names, performance details, and optional remarks.

Table 5: IPL_Team_players

Data Dictionary:

<i>Variable</i>	<i>Data type</i>	<i>Comments</i>
<u>TeamId</u>	NUMBER	Composite Primary key
<u>PlayerId</u>	NUMBER	Composite Primary key
Player_role	VARCHAR	Captain, Batsman, Bowler, WK, etc.
Remarks	VARCHAR	Team names that he has been played for.

Dataset Description:

The IPL_Team_players table represents the relationship between teams and players, indicating the roles of players within specific teams. It includes team IDs, player IDs, player roles, and team-related remarks.

Table 6: IPL_Tournament

Data Dictionary:

<i>Variable</i>	<i>Data type</i>	<i>Comments</i>
<u>TournamentId</u>	NUMBER	Primary key
Tournament_name	VARCHAR	Not null
From_date	DATE	Tournament starting date
To_date	DATE	Tournament ending date
Team_count	NUMBER	Total team count of the seasons
Total_matches	NUMBER	Total matches the team had played for each season's
Remarks	VARCHAR	Champions teams name of the seasons

Dataset Description:

The IPL_Tournament table contains information about the IPL tournaments. It includes tournament IDs, names, start and end dates, team counts, total matches played, and champions' team names for each season.

Table 7: IPL_Match

Data Dictionary:

<i>Variable</i>	<i>Data type</i>	<i>Comments</i>
<u>MatchId</u>	NUMBER	Primary key
TeamId1	NUMBER	FK from Team table. Not null
TeamId2	NUMBER	FK from Team table. Not null
TossWinner	NUMBER	Team no. 1 or 2
MatchWinner	NUMBER	Team no. 1 or 2
WinDetails	VARCHAR	Team 1 or 2 Won by XX runs or X wickets, Match tied.
Remarks	VARCHAR	E.g. Match canceled due to rain.

Dataset Description:

The IPL_Match table stores information about individual matches in the IPL. It includes match IDs, team IDs of the playing teams, toss winner, match winner, match result details, and optional remarks.

Table 8: IPL_Match_Schedule

Data Dictionary:

<i>Variable</i>	<i>Data type</i>	<i>Comments</i>
<u>ScheduleId</u>	NUMBER	Primary key
TournamentId	NUMBER	FK from Tournament table.
MatchId	NUMBER	FK from Match table.
Match_type	VARCHAR	League, Knock out, Final, etc.
Match_date	DATE	This date should be within the from and to dates of the tournament.
Start_time	TIME	
StadiumId	NUMBER	FK from Stadium table
Status	VARCHAR	Scheduled, Completed, Cancelled, etc.
Remarks	VARCHAR	Reasons for the march cancellation.

Dataset Description:

The IPL_Match_Schedule table contains the schedule details of each match in the IPL tournament. It includes schedule IDs, tournament IDs, match IDs, match types, match dates, start times, stadium IDs, match status, and optional remarks.

Table 9: IPL_Bidder_Details

Data Dictionary:

<i>Variable</i>	<i>Data type</i>	<i>Comments</i>
<u>BidderId</u>	NUMBER	Primary key
UserId	NUMBER	FK from User table.
Bidder_name	VARCHAR	Not null
Contact_no	NUMBER	Not null
Emailid	VARCHAR	Email id of the bidders'
Remarks	VARCHAR	Null values

Dataset Description:

The IPL_Bidder_Details table stores information about the bidders in the IPL. It includes bidder IDs, user IDs, bidder names, contact numbers, email IDs, and optional remarks.

Table 10: IPL_Bidding_Details

Data Dictionary:

<i>Variable</i>	<i>Data type</i>	<i>Comments</i>
<u>BidderId</u>	NUMBER	FK from Bidder table. Composite Primary key
<u>ScheduleId*</u>	VARCHAR	FK from Match_Schedule table. Composite Primary key.
BidTeam	NUMBER	One of the team-ids of the match (1 or 2). Composite primary key.
BidDate	DATETIME	Exact date & time of placing the bid. Update this column if a bidder re-bids on the same team for the same match. Composite Primary key.
<u>BidStatus</u>	VARCHAR	Bid, Cancelled, Won, Lost

Dataset Description:

The IPL_Bidding_Details table contains details about the bids placed by the bidders for matches in the IPL. It includes bidder IDs, schedule IDs, team IDs for bidding, bid dates, and bid statuses.

Table 11: IPL_Bidder_Points

Data Dictionary:

<i>Variable</i>	<i>Data type</i>	<i>Comments</i>
<u>BidderId</u>	NUMBER	FK from Bidder table. Primary key
TournamentId	NUMBER	FK from Tournament table.
No_of_bids	NUMBER	Total no. of bids placed by a bidder. Updated after completion of the match on which s/he placed the bid.
No_of_matches	NUMBER	Total no. of matches on which s/he placed the bid. Updated as above.
Total_points	NUMBER	Not null. Default 0

Dataset Description:

The IPL_Bidder_Points table stores information about the points earned by bidders in the IPL. It includes bidder IDs, tournament IDs, the number of bids placed, the number of matches bid upon, and the total points accumulated.

Table 12: IPL_Team_Standings

Data Dictionary:

<i>Variable</i>	<i>Data type</i>	<i>Comments</i>
<u>TeamId</u>	NUMBER	FK from Team table. Primary key
TournamentId	NUMBER	FK from Tournament table.
Matches_played	NUMBER	Not null. Default 0 - Number of matches played
Matches_won	NUMBER	Not null. Default 0 - Number of matches won
Matches_lost	NUMBER	Not null. Default 0 - Number of matches lost
Matches_tied	NUMBER	Default 0 - Number of matches tied
No_result	NUMBER	Default 0 - Number of matches that have no results
Points	NUMBER	Not null. Default 0 - Total points of the team
Remarks	VARCHAR	mentioned champions

Dataset Description:

The IPL_Team_Standings table provides information about the team standings in the IPL. It includes team IDs, tournament IDs, the number of matches played, won, lost, tied, no result matches, total points, and optional remarks.

Business Importance

The Pie-in-the-Sky mobile app and its associated dataset play a crucial role in facilitating a fair, engaging, and data-driven bidding experience for IPL enthusiasts. It contributes to user satisfaction, user retention, and revenue opportunities, while also streamlining administrative processes for efficient tournament management.

The Pie-in-the-Sky mobile application holds great significance in the business realm due to its myriad benefits:

1. User Engagement: This app offers a dynamic and interactive platform for IPL enthusiasts to actively partake in matches through bidding and predicting winners. It delivers an immersive experience, captivating users and fostering their deep involvement in the tournament.
2. Fair and Legal Bidding: By providing a compliant platform for bidding on IPL matches, the app ensures transparency and regulation, safeguarding the integrity of the tournament and curbing illicit or unauthorized betting activities.
3. Data-Driven Decision Making: Users gain access to comprehensive match details, team standings, and bidder leaderboards, empowering them to make well-informed decisions when placing bids. Armed with data-driven insights, they can analyze team performances, assess match scenarios, and strategize their bidding tactics.
4. User Retention and Loyalty: The app's personalized features, including individual point tracking, leaderboard rankings, and team standings, fuel healthy competition and encourage users to stay actively engaged throughout the IPL season. This cultivates a loyal user base, bolstering sustained app usage.
5. Revenue Generation Opportunities: The app presents diverse avenues for revenue generation, such as in-app purchases, strategic advertising partnerships, and premium



features that enhance the bidding experience. Leveraging the IPL's immense popularity and the app's active user base, there exists significant potential for revenue growth and business expansion.

6. Streamlined Administration and Efficiency: With its robust administrative functionalities, the app enables seamless management of the tournament, team rosters, match schedules, and result declarations. This streamlines administrative processes, minimizes manual efforts, and ensures smooth operations throughout the IPL season.

7. Real-Time Updates and Communication: Leveraging real-time capabilities, the app promptly delivers updates and notifications to bidders, keeping them informed about match schedule changes, match results, and vital announcements. This seamless communication enhances user experience and facilitates effective engagement between the app and its users.

Overall, the Pie-in-the-Sky mobile application serves as an indispensable tool for driving user engagement, promoting fair bidding practices, facilitating data-driven decision making, fostering user loyalty, and unlocking revenue opportunities. Its efficient administration and seamless communication capabilities further solidify its position as an integral component within the IPL ecosystem.

Business Questions Discussion

Q1. Show the percentage of wins of each bidder in the order of highest to lowest percentage.

In the above question, query calculates the win percentage for each bidder based on the number of 'won' bid statuses and orders the results in descending order to show the bidders with the highest win percentages at the top.

Code Snippet:

```
select ipl_bidding_details.bidder_id, count(ipl_bidding_details.bid_status), no_of_bids,
(count(ipl_bidding_details.bid_status)/no_of_bids)*100 as percentage_win
from ipl_bidding_details
inner join ipl_bidder_points
on ipl_bidding_details.bidder_id= ipl_bidder_points.bidder_id
where ipl_bidding_details.bid_status='won'
group by ipl_bidding_details.bidder_id, no_of_bids
order by percentage_win desc;

with temp( won, BIDDER_ID) as
(select count(NO_OF_BIDS), BIDDER_ID
from ipl_bidder_points left join ipl_bidding_details using(BIDDER_ID)
where BID_STATUS like '%won%'
group by BIDDER_ID, BID_STATUS)
select BIDDER_ID, ((won/NO_OF_BIDS)*100) as percentage_of_won from temp join ipl_bidder_points using(BIDDER_ID)
order by percentage of won desc;
```

Solution:

	BIDDER_ID	percentage_of_won
▶	103	100.0000
	121	90.9091
	118	83.3333
	126	80.0000
	104	71.4286
	122	66.6667

Inference: The query provides valuable information about the performance of bidders in terms of their win percentages. By analyzing the results, we can identify the bidders who have been most successful in winning bids. This information can be used to evaluate bidder performance

Q2. Display the number of matches conducted at each stadium with the stadium name and city.

In the following question, we need to retrieve the count of matches held at each stadium along with their respective stadium names and cities

Code Snippet:

- ```
select count(ims.match_id)as no_of_matches,iss.stadium_name from ipl_stadium iss join ipl_match_schedule ims on
iss.stadium_id =ims.stadium_id
group by iss.stadium_name;
```

Solution:

| Result Grid |               |                                    | Filter Rows: | Export: |
|-------------|---------------|------------------------------------|--------------|---------|
|             | no_of_matches | stadium_name                       |              |         |
| ▶           | 10            | Sawai Mansingh Stadium             |              |         |
|             | 13            | M. Chinnaswamy Stadium             |              |         |
|             | 13            | Eden Gardens                       |              |         |
|             | 18            | Wankhede Stadium                   |              |         |
|             | 16            | Is Bindra Stadium                  |              |         |
|             | 13            | Feroz Shah Kotla                   |              |         |
|             | 13            | Holkar Stadium                     |              |         |
|             | 12            | MS Chidambaram Stadium             |              |         |
|             | 7             | MCA Stadium                        |              |         |
|             | 7             | Rajiv Gandhi International Stadium |              |         |

Inference: It provides valuable information about the distribution of matches across different stadiums, allowing us to analyze the popularity and utilization of each stadium for hosting matches in a particular city.

Q3. In a given stadium, what is the percentage of wins by a team which has won the toss?

Calculates the percentage of wins by a team that has won the toss in a specific stadium.

Code Snippet:

```
select ist.STADIUM_NAME,
(sum((Case when Toss_winner = Match_winner then 1 else 0 end))/count(*))*100
as 'Percent Wins'
from ipl_stadium ist join ipl_match_schedule ms
on ist.stadium_id = ms.STADIUM_ID
join ipl_match im on ms.match_id= im.match_id
group by ist.STADIUM_NAME;
```

Solution:

|   | STADIUM_NAME                       | Percent Wins |
|---|------------------------------------|--------------|
| ► | Sawai Mansingh Stadium             | 70.0000      |
|   | M. Chinnaswamy Stadium             | 38.4615      |
|   | Eden Gardens                       | 38.4615      |
|   | Wankhede Stadium                   | 61.1111      |
|   | Is Bindra Stadium                  | 62.5000      |
|   | Feroz Shah Kotla                   | 53.8462      |
|   | Holkar Stadium                     | 38.4615      |
|   | MS Chidambaram Stadium             | 33.3333      |
|   | MCA Stadium                        | 28.5714      |
|   | Rajiv Gandhi International Stadium | 14.2857      |

Inference: This information provides insights into the correlation between winning the toss and winning the match at a particular stadium

Q4. Show the total bids along with the bid team and team name.

The query aims to display the total number of bids placed along with the corresponding bid team and team name.

#### Code Snippet:

```
#4. Show the total bids on a team along with the bid team and team name.
select count(bidder_id) as total_bid_on_team, bid_team, team_name
from ipl_bidding_details bd inner join ipl_team it
on bid_team=it.team_id
group by bid_team;
```

#### Solution:

| total_bid_on_team | bid_team | team_name                   |
|-------------------|----------|-----------------------------|
| 22                | 1        | Chennai Super Kings         |
| 25                | 7        | Royal Challengers Bangalore |
| 32                | 8        | Sunrisers Hyderabad         |
| 27                | 6        | Rajasthan Royals            |
| 26                | 2        | Delhi Daredevils            |
| 24                | 3        | Kings XI Punjab             |
| 22                | 5        | Mumbai Indians              |
| 22                | 4        | Kolkata Knight Riders       |

Inference: From this we can infer the level of bidding activity for each team and identify which teams have attracted more bids. This information can be used to assess the popularity and demand for different teams among bidders, potentially indicating the teams' fan base or perceived value in the bidding process.

Q5. Show the team id who won the match as per the win details.

Retrieves the team IDs of the teams that have won matches based on the win details.

#### Code Snippet:



```
52 #5. Show the team id who won the match as per the win details.
53 • select* from ipl_match;
54 • select team_id, team_name, team_id1, team_id2, match_winner, ipl_match.win_details
55 from ipl_team
56 inner join ipl_match
57 on substr(ipl_team.remarks,1,3) = substr(ipl_match.win_details,6,3);
```

### Solution:

| team_id | team_name                   | team_id1 | team_id2 | match_winner | win_details              |
|---------|-----------------------------|----------|----------|--------------|--------------------------|
| 1       | Chennai Super Kings         | 1        | 2        | 1            | Team CSK won by 7 Wkts   |
| 1       | Chennai Super Kings         | 1        | 3        | 1            | Team CSK won by 7 Wkts   |
| 4       | Kolkata Knight Riders       | 1        | 4        | 2            | Team KKR won by 35 Ru... |
| 1       | Chennai Super Kings         | 1        | 5        | 1            | Team CSK won by 7 Wkts   |
| 7       | Royal Challengers Bangalore | 1        | 7        | 2            | Team RCB won by 35 Ru... |
| 8       | Sunrisers Hyderabad         | 2        | 8        | 2            | Team SRH won by 7 Wkts   |
| 3       | Kings XI Punjab             | 3        | 2        | 1            | Team KXIP won by 7 Wkts  |
| 1       | Chennai Super Kings         | 3        | 1        | 2            | Team CSK won by 7 Wkts   |

Inference: It helps identify the teams that have achieved success in the matches and can be used to analyze their performance and track their winning streaks throughout the tournament.

Q6. Display total matches played, total matches won and total matches lost by the team along with its team name.

Shows the total number of matches played, matches won, and matches lost for each team, along with their respective team names.

### Code Snippet:

```
59 #6. Display total matches played, total matches won and total matches lost by the team
60 #along with its team name.
61 • select distinct(team_id), TEAM_NAME ,
62 sum(MATCHES_PLAYED)over(partition by TEAM_ID) T_MATCHES_PLAYED,
63 sum(MATCHES_WON)over(partition by TEAM_ID) as T_MATCHES_WON,
64 sum(MATCHES_LOST) over(partition by TEAM_ID) as T_MATCHES_LOST
65 from ipl_team_standings join ipl_team using(team_id);
```

### Solution:

| team_id | TEAM_NAME                  | T_MATCHES_PLAYED | T_MATCHES_WON | T_MATCHES_LOST |
|---------|----------------------------|------------------|---------------|----------------|
| 1       | Chennai Super Kings        | 28               | 18            | 10             |
| 2       | Delhi Daredevils           | 28               | 11            | 17             |
| 3       | Kings XI Punjab            | 28               | 13            | 15             |
| 4       | Kolkata Knight Riders      | 28               | 16            | 12             |
| 5       | Mumbai Indians             | 28               | 16            | 12             |
| 6       | Rajasthan Royals           | 28               | 16            | 12             |
| 7       | Royal Challengers Banga... | 28               | 9             | 18             |
| 8       | Sunrisers Hyderabad        | 28               | 17            | 10             |

Inference: We can infer the overall performance of each team based on the total matches played, matches won, and matches lost. It allows us to compare the success rates of different teams and identify teams with a higher win percentage or better overall performance.

Q7. Display the bowlers for the Mumbai Indians team.

Code Snippet:

```
select* from ipl_team it join ipl_team_players itl
on it.team_id =itl.team_id
where team_name ="Mumbai Indians" and player_role="bowler";
```

Solution:

| Result Grid  |                |                     |         |         |                    |             |             |  |
|--------------|----------------|---------------------|---------|---------|--------------------|-------------|-------------|--|
| Filter Rows: |                |                     |         | Export: | Wrap Cell Content: |             |             |  |
| TEAM_ID      | TEAM_NAME      | TEAM_CITY           | REMARKS | TEAM_ID | PLAYER_ID          | PLAYER_ROLE | REMARKS     |  |
| 5            | Mumbai Indians | Mumbai, Maharashtra | MI      | 5       | 8                  | Bowler      | TEAM - SRH  |  |
| 5            | Mumbai Indians | Mumbai, Maharashtra | MI      | 5       | 12                 | Bowler      | TEAM - KXIP |  |
| 5            | Mumbai Indians | Mumbai, Maharashtra | MI      | 5       | 24                 | Bowler      | TEAM - KKR  |  |
| 5            | Mumbai Indians | Mumbai, Maharashtra | MI      | 5       | 33                 | Bowler      | TEAM - KKR  |  |
| 5            | Mumbai Indians | Mumbai, Maharashtra | MI      | 5       | 46                 | Bowler      | TEAM - DD   |  |
| 5            | Mumbai Indians | Mumbai, Maharashtra | MI      | 5       | 56                 | Bowler      | TEAM - RCB  |  |
| 5            | Mumbai Indians | Mumbai, Maharashtra | MI      | 5       | 68                 | Bowler      | TEAM - MI   |  |
| 5            | Mumbai Indians | Mumbai, Maharashtra | MI      | 5       | 98                 | Bowler      | TEAM - CSK  |  |
| 5            | Mumbai Indians | Mumbai, Maharashtra | MI      | 5       | 128                | Bowler      | TEAM - SRH  |  |

Q8. How many all-rounders are there in each team, Display the teams with more than 4 all-rounders in descending order.

### Code Snippet:

```
select it.team_name,count(itl.player_role) as allrounder from ipl_team it join ipl_team_players itl
on it.team_id =itl.team_id
where itl.player_role = "All-Rounder"
group by it.TEAM_NAME
having count(itl.player_role)>4;
```

### Solution:

| Result Grid           |            | Filter Rows: |
|-----------------------|------------|--------------|
| team_name             | allrounder |              |
| Delhi Daredevils      | 7          |              |
| Kings XI Punjab       | 7          |              |
| Kolkata Knight Riders | 5          |              |
| Rajasthan Royals      | 5          |              |
| Sunrisers Hyderabad   | 6          |              |

Q10. Extract the Bowlers and All Rounders those are in the 5 highest number of wickets.

Note:

1. Use the performance\_dtls column from ipl\_player to get the total number of wickets
2. Do not use the limit method because it might not give appropriate results when players have the same number of wickets
3. Do not use joins in any cases.
4. Display the following columns teamn\_name, player\_name, and player\_role.

The question aims to extract the top five bowlers and all-rounders based on the highest number of wickets taken in a given context.

### Code Snippet:

```
3 • SELECT TEAM_NAME, PLAYER_NAME, PLAYER_ROLE
9 FROM ipl_player e,ipl_team_players r,ipl_team t
9 WHERE (e.PLAYER_ID=r.PLAYER_ID and r.TEAM_ID=t.TEAM_ID) and PLAYER_ROLE IN ('Bowler', 'All-Rounder')
1 ORDER BY CAST(SUBSTRING_INDEX(performance_dtls, '/', 1) AS UNSIGNED) desc limit 5;
```

### Solution:

| TEAM_NAME           | PLAYER_NAME    | PLAYER_ROLE |
|---------------------|----------------|-------------|
| Chennai Super Kings | Shane Watson   | Bowler      |
| Chennai Super Kings | Ambati Rayudu  | All-Rounder |
| Chennai Super Kings | MS Dhoni       | Bowler      |
| Chennai Super Kings | Shardul Thakur | All-Rounder |
| Chennai Super Kings | Deepak Chahar  | Bowler      |

Inference: This information is valuable for identifying the most successful players in terms of their bowling performance, indicating their effectiveness and contribution to the team.

Q11. Show the percentage of toss wins of each bidder and display the results in descending order based on the percentage

The question aims to calculate the percentage of toss wins for each bidder and sort the results in descending order based on the calculated percentage.

Code Snippet:

```
05 • with temp(won,bidder_id) as
06 (select count(no_of_bids),bidder_id
07 from ipl_bidder_points left join ipl_bidding_details using(bidder_id)
08 where bid_status like '%won%'
09 group by bidder_id,bid_status)
10 select distinct(ipl_bidder_points.bidder_id), ((won/no_of_bids)*100) as percentage_of_toss_won
11 from temp join ipl_bidder_points using(bidder_id)
12 order by percentage_of_toss_won desc;
```

Solution:

| bidder_id | percentage_of_toss_won |
|-----------|------------------------|
| 103       | 100.0000               |
| 121       | 90.9091                |
| 118       | 83.3333                |
| 126       | 80.0000                |
| 104       | 71.4286                |
| 122       | 66.6667                |

Inference: This analysis provides insights into the effectiveness of bidders in winning the toss and their strategic advantage in cricket matches.

Q12. Find the IPL season which has min duration and max duration.

Output columns should be like the below:

Tournament\_ID, Tournament\_name, Duration column, Duration

The query aims to identify the IPL season with the minimum and maximum duration.

#### Code Snippet:

```
8 • select tournament_id, tournament_name, datediff(to_date, from_date) as duration
9 from ipl_tournament
10 where datediff(to_date, from_date) in
11 (select min(datediff(to_date, from_date)) as min_duration from ipl_tournament
12 union
13 (select max(datediff(to_date, from_date)) as max_duration from ipl_tournament));
```

#### Solution:

| tournament_id | tournament_name   | duration |
|---------------|-------------------|----------|
| 2009          | IPL SEASON - 2009 | 36       |
| 2012          | IPL SEASON - 2012 | 53       |
| 2013          | IPL SEASON - 2013 | 53       |

Inference: This analysis provides insights into the variations in the duration of different IPL seasons, allowing for comparisons and understanding of the tournament's timeline over the years.

Q13. Write a query to display to calculate the total points month-wise for the 2017 bid year. Sort the results based on total points in descending order and month-wise in ascending order.

Note: Display the following columns:

1. Bidder ID
2. Bidder Name
3. Bid date as Year
4. Bid date as Month

## 5. Total points

Only use joins for the above query queries.

Calculates the total points earned by bidders in each month of the 2017 bid year. The results are sorted in descending order of total points and ascending order of months

### Code Snippet:

```
with temp(BIDDER_ID, BIDDER_NAME,NO_OF_BIDS, bid_date, t_points) as
(select distinct(BIDDER_ID),BIDDER_NAME, NO_OF_BIDS, BID_DATE, total_points
from ipl_bidder_details join ipl_bidder_points using(BIDDER_ID)
join ipl_bidding_details using(bidder_id)
where BID_DATE like '%2017%')
select BIDDER_ID, BIDDER_NAME,NO_OF_BIDS, bid_date, t_points ,
sum(t_points) over (partition by (BIDDER_ID)) total_points_per_month
from temp ;
```

### Solution:

|   | BIDDER_ID | BIDDER_NAME        | NO_OF_BIDS | bid_date            | t_points | total_points_per_month |
|---|-----------|--------------------|------------|---------------------|----------|------------------------|
| ▶ | 102       | Krishan Valimbe    | 4          | 2017-04-05 00:00:00 | 0        | 0                      |
|   | 102       | Krishan Valimbe    | 4          | 2017-04-17 00:00:00 | 0        | 0                      |
|   | 103       | Megaduta Dheer     | 5          | 2017-04-17 00:00:00 | 19       | 38                     |
|   | 103       | Megaduta Dheer     | 5          | 2017-05-05 00:00:00 | 19       | 38                     |
|   | 104       | Chatur Mahalanabis | 7          | 2017-04-05 00:00:00 | 17       | 85                     |
|   | 104       | Chatur Mahalanabis | 7          | 2017-04-09 00:00:00 | 17       | 85                     |
|   | 104       | Chatur Mahalanabis | 7          | 2017-05-13 00:00:00 | 17       | 85                     |
|   | 104       | Chatur Mahalanabis | 7          | 2017-05-14 00:00:00 | 17       | 85                     |
|   | 104       | Chatur Mahalanabis | 7          | 2017-05-15 00:00:00 | 17       | 85                     |
|   | 105       | Shackcham Bajpeyi  | 9          | 2017-04-30 00:00:00 | 4        | 12                     |
|   | 105       | Shackcham Bajpeyi  | 9          | 2017-05-01 00:00:00 | 4        | 12                     |
|   | 105       | Shackcham Bajpeyi  | 9          | 2017-05-15 00:00:00 | 4        | 12                     |
|   | 106       | Vineet Hegadi      | 10         | 2017-04-05 00:00:00 | 14       | 70                     |

Inference: It allows us to identify the top-performing bidders in each month of the 2017 bid year based on their total points. This information can be useful for evaluating bidder performance, identifying trends, and making strategic decisions in future bidding processes.



Q.14 Write a query for the above question using sub queries by having the same constraints as the above question.

Retrieves the total points month-wise for the 2017 bid year and sorts the results in descending order of total points.

Code Snippet:

```
select bd.bidder_id, bd.bidder_name,
year(b.bid_date) as 'year',
month(b.bid_date) as 'month',
(select sum(p.total_points) from ipl_bidder_points p
where p.bidder_id = b.bidder_id and p.tournmt_id = b.schedule_id) as 'total points'
from ipl_bidding_details b join ipl_bidder_details bd on b.bidder_id = bd.bidder_id
where year(b.bid_date) = 2017 order by (select sum(p.total_points) from ipl_bidder_points p
where p.bidder_id = b.bidder_id and p.tournmt_id = b.schedule_id) desc,
year(b.bid_date), month(b.bid_date);
```

Solution:

|   | bidder_id | bidder_name        | year | month | total points |
|---|-----------|--------------------|------|-------|--------------|
| ▶ | 102       | Krishan Valimbe    | 2017 | 4     | NULL         |
|   | 102       | Krishan Valimbe    | 2017 | 4     | NULL         |
|   | 103       | Megaduta Dheer     | 2017 | 4     | NULL         |
|   | 104       | Chatur Mahalanabis | 2017 | 4     | NULL         |
|   | 104       | Chatur Mahalanabis | 2017 | 4     | NULL         |
|   | 105       | Shackcham Bajpeyi  | 2017 | 4     | NULL         |
|   | 106       | Vineet Hegadi      | 2017 | 4     | NULL         |
|   | 106       | Vineet Hegadi      | 2017 | 4     | NULL         |
|   | 107       | Vakpati Munshi     | 2017 | 4     | NULL         |
|   | 107       | Vakpati Munshi     | 2017 | 4     | NULL         |
|   | 108       | Kusika Rajavade    | 2017 | 4     | NULL         |
|   | 109       | Gagan Panda        | 2017 | 4     | NULL         |

Inference: The query allows us to analyze the performance of bidders in the 2017 bid year based on their total points, sorted month-wise. It provides insights into the relative success of bidders during different months, highlighting any patterns or trends in their performance throughout the year.

Q16. Create two tables called Student\_details and Student\_details\_backup.

Code Snippet:

```
184 • CREATE TABLE Student_details (
185 Student_id INT NOT NULL,
186 Student_name VARCHAR(50) NOT NULL,
187 Mail_id VARCHAR(50) NOT NULL,
188 Mobile_no VARCHAR(15) NOT NULL,
189 PRIMARY KEY (Student_id)
190);
191
192 • CREATE TABLE Student_details_backup (
193 Student_id INT NOT NULL,
194 Student_name VARCHAR(50) NOT NULL,
195 Mail_id VARCHAR(50) NOT NULL,
196 Mobile_no VARCHAR(15) NOT NULL,
197 Backup_time TIMESTAMP NOT NULL DEFAULT CURRENT_TIMESTAMP,
198 PRIMARY KEY (Student_id, Backup_time)
199);
```

Using Trigger

```
200 DELIMITER //
201 • CREATE TRIGGER insert_student_details
202 AFTER INSERT ON Student_details
203 FOR EACH ROW
204 BEGIN
205 INSERT INTO Student_details_backup (Student_id, Student_name, Mail_id, Mobile_no)
206 VALUES (NEW.Student_id, NEW.Student_name, NEW.Mail_id, NEW.Mobile_no);
207 END //
208 DELIMITER ;
```



## Conclusion

### 1. Lesson Learned:

The development of the Pie-in-the-Sky mobile app has provided insights into the complexities involved in creating a bidding platform for IPL matches. It highlights the importance of considering dynamic factors such as match schedule changes, match disruptions, and updating match statistics and standings in real-time. Additionally, it emphasizes the need for robust user authentication and data privacy measures to ensure a secure and fair bidding environment.

### 2. Skills Used:

The development of the Pie-in-the-Sky mobile app required a range of skills including database management with MySQL, query optimization, data modeling, and understanding of relational databases.

### 3. Domain Understanding Developed:

Building the Pie-in-the-Sky mobile app provided an opportunity to develop domain understanding in the field of sports bidding and tournament management, specifically focused on the IPL. It involved understanding the rules and dynamics of the IPL, including match scheduling, team standings, and points calculation. Domain understanding also extends to the considerations and challenges faced in implementing a fair and engaging bidding system, ensuring user satisfaction, and maintaining data accuracy and integrity.

To Conclude, the domain understanding developed highlights the vital role of data analytics in the IPL ecosystem. It empowers stakeholders to make informed decisions, optimize player performance, engage fans, and enhance the overall cricketing experience. Leveraging data analytics is essential for driving success and growth in the IPL.

## References

### Industry Review

- Website – <https://www.datacamp.com/blog/sports-analytics-how-different-sports-use-data-analysis>

### Literature Survey

- Google Scholar
- Journal of Sports Analytics 4 (2018) 263–273