MSc International Business with Data Analytics.

Module Name	Data Analytics for International Business.
Module Code	BMG880 (19628)
Module Tutor	Mr. Mohammad Khurram
Student Name	Shamanth Shripad Hegde.
Student Number	
Student ID	
Word Count	2461. (except cover page, table of contents, and references, figure)

Table of Contents

1. Introduction.

Data science is nothing more than presenting a meaningful story (Srinivasa & Bhatnagar, 2012). The term data analytics refers to the method and practice of analyzing data to gain insights, answer questions, and detect patterns. This is accomplished via the use of a variety of tools, methodologies, and frameworks that vary according to the sort of analysis undertaken (HBR, 2021; Russom, 2011). However, Data preparation, data mining, data management, and data visualization are all aspects of business intelligence (BI) (IBM, 2021). With the use of business intelligence tools and procedures, users can find meaningful information from raw data, enabling businesses of all sizes to make more informed decisions (Graeme & Nargiza, 2012). Business Intelligence is not only used by businesses but also used in marketing, trading, healthcare, and sports.

1.0 Sports Analytics.

In sports, Analytical methods are used by coaches in the selection and formation of sports teams, as well as in the creation of tactics and match-ups. Analytical data in sports refers to a collection of relevant, historical information that may be used to gain an edge over the competition. For teams and players, on-the-field analytics focuses on increasing performance. Game strategy and player conditioning are examined thoroughly by the use of data analytics. Business Analytics in sports is called sports analytics (Morgulev, et.al., 2018).

1.1 Indian Premier League.

An annual cricket tournament in India, the Indian Premier League (IPL), is one of the world's most popular sporting events on the planet (IPL, 2021). It began in 2008. As of 2021, it has a viewership of 380 million people (Economic Times, 2021). This event is put on by The Board of Control for Cricket in India (BCCI).

1.2 Problem Statement.

The Royal Challengers Bangalore (RCB) franchise has participated in every season of the Indian Premier League. There are just a few teams that have participated in every season, and yet they have never won the championship (Hindustan Times, 2021). This research aims to discover why the team failed to win the IPL title with the support of data analysis.

1.3 Opportunity statement.

Every team is required to release players for the auction. Each team can use 90 Crore INR. The Royal Challengers Bangalore team will be able to retain a maximum of 4 players from the current season due to the forthcoming auction in 2022 (IPL, 2021). This assignment's secondary purpose is to assist the Royal challengers Bangalore choose the best squad in the auction in 2022 with the help of Business Intelligence.

2. Entity-Relationship Diagram.

An Entity-Relationship Diagram (ERD), ERDs are data models that describe the relationships between entities, ideas, and objects (Graeme & Nargiza, 2012). An analyst or a business user can utilize ERDs to grasp the business domain, clarify business jargon, and relate business ideas to database structures. It is also known as the Entity-Relationship Model (Sharda, et al., 2014). It can be used for a variety of purposes, including database design, database troubleshooting, and educational purposes.

2.1 Components of ERD.

Entities are usually seen as rectangles, entities might be things, people, ideas, or events that hold data. **Attributes** denoted by a circle or an oval, attributes allude to an entity's features. They are classified as simple, composite, or derived, with an object having one or more properties. **The relationship** demonstrates the interconnectedness of several entities. Entities are linked together by a series of lines with labels (Sharda, et al., 2014; Li & YL., 2009).

2.2 ERD symbol and notation.

An entity's **cardinality** shows the number of occurrences it can have in relation to another entity. Two entities might be One-to-one, Many to One, Many to One, or Many to Many. Cardinality identifies the kind of connection between the two entities (Li & YL., 2009). For instance, a single cricket team has a large number of players. A one-to-many link exists between the entity's Team and Player in an ERD. Chen Notation, Crow's Foot Notation, IDEF1X Style, and Bachman Style are all used to show cardinality between 2 or more entities (Winston, 2010).

2.3 ERD diagram for IPL Players dataset.

Excel uses rows and columns to organize data. Excel is considered a relational database management system (RDBMS). An RDBMS may be observed in Excel when numerous spreadsheets are joined to do a computation (Dasararaju & Taori, 2019). The raw data that was gathered for the assessment was unclean, unprocessed, and messy. It was not suitable for calculations. In order to better comprehend the data and make calculations easier, I've built a small database with the intention of building a player's performance dashboard. Illustration 2.1 shows the database's entity-relationship diagram with crows foot notation.

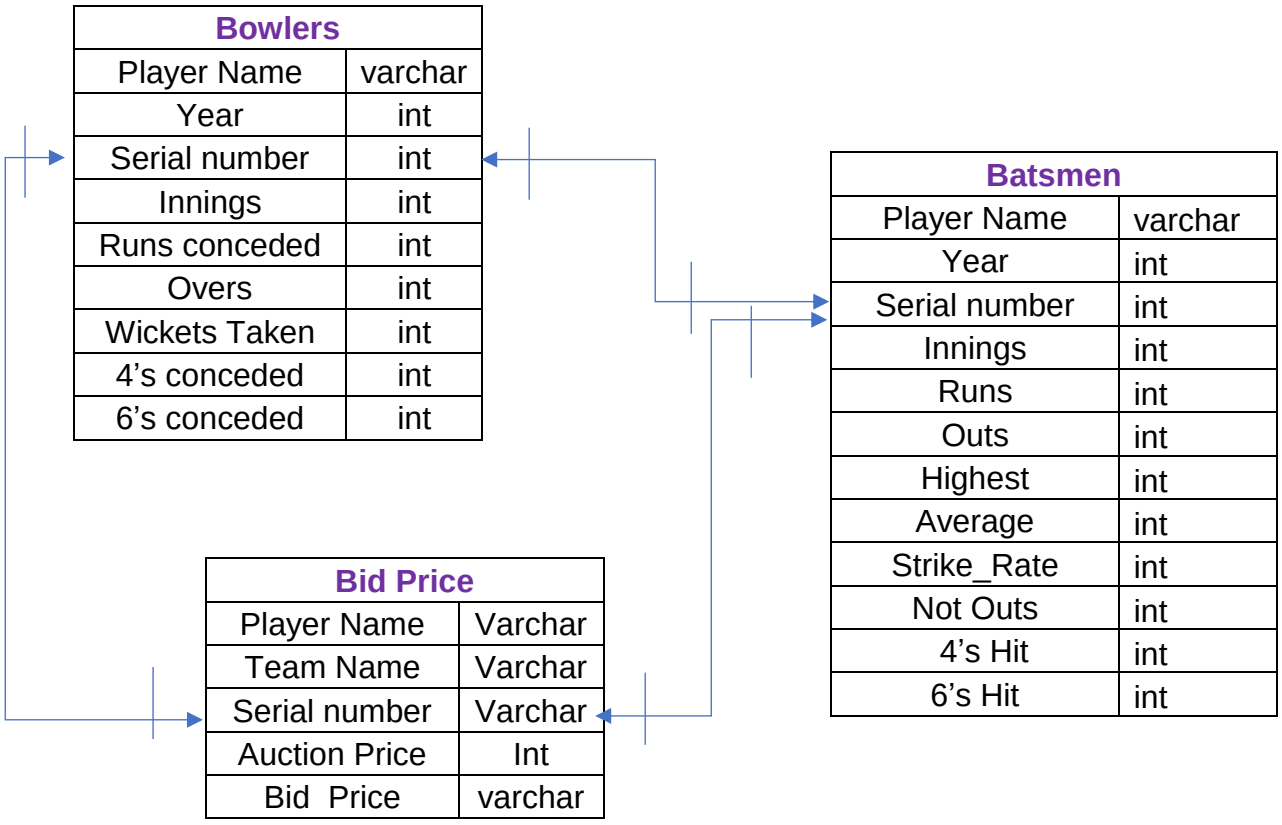

Illustration 2.1 Entity-Relationship Model of Players data in crow's foot notation.

The database's components are outlined below.

Entities – Bid_price, Batsmen, Bowlers.

Attributes -Player's name, Team_name, innings, runs, outs, etc.

Cardinality- one-to-one.

Relationship - It is possible for a single player to play for the same team in different seasons. As a result, a player's name cannot serve as the main identifier. However,

the combined data of year and player's name serve as a primary key. In SQL, merging two columns for the primary key is permitted, however in Excel, the computation is difficult (Winston, 2010; Dasararaju & Taori, 2019). That's why a unique attribute is created called serial_number. the relationship can be defined between the entities with the "serial_number" attribute

Data types – Varchar, Int.

3. Why haven't Royal Challengers Bangalore won a trophy in the Indian Premier League yet?

Surprisingly, Royal Challengers Bangalore haven't won an IPL championship yet. The study attempts to determine why they have not yet won an IPL title through data analysis. The dataset is downloaded from Kaggle (Kaggle, 2021). The data related t all the matches played between 2008 - 2020. This section focuses on IPL matches played between different teams.

3.1 low win percentage.

Between 2008 and 2020, Royal Challengers Bangalore played 195 matches and won only 91 of them. According to the data analysis, RCB has won only 46% of the matches they have played and which is less than half of the matches they played.

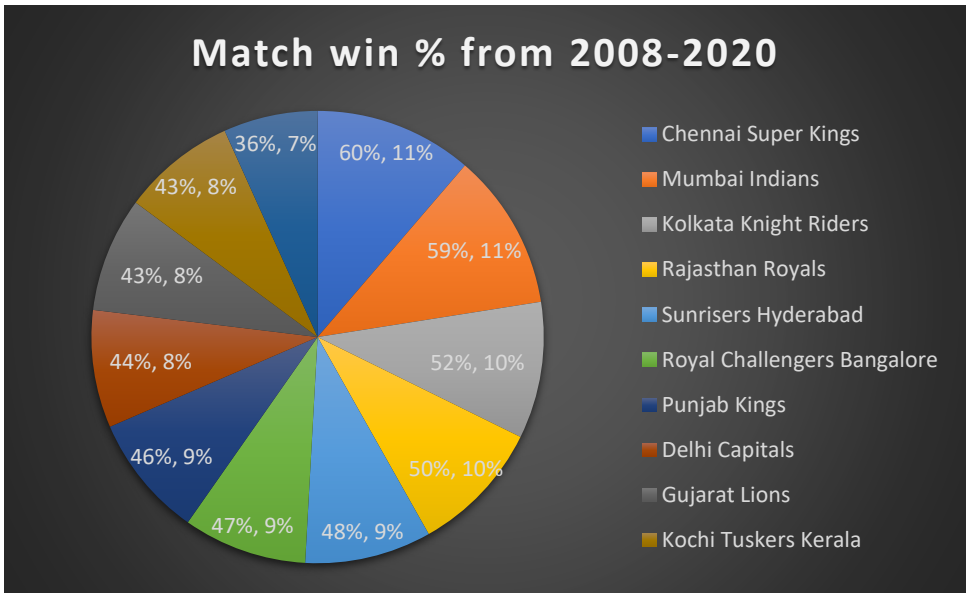

Figure 3.1 match win percentage.

3.2 Game of Luck.

3.2.1 Toss win percentage.

In cricket, "toss win match win", “toss is boss” is a frequent phrase. The coin toss or flip of the coin is a critical aspect of a cricket game. Unlike in other sports, the toss has a significant impact on the eventual outcome of the game. Thus, the captain who wins the toss has the option of batting or bowling, based on the team and pitch conditions (The Economic Time, 2021). RCB has played a total of 195 matches and won the toss only 87 times. Which is approximately 45%. It is the second-lowest when compared to the other teams.

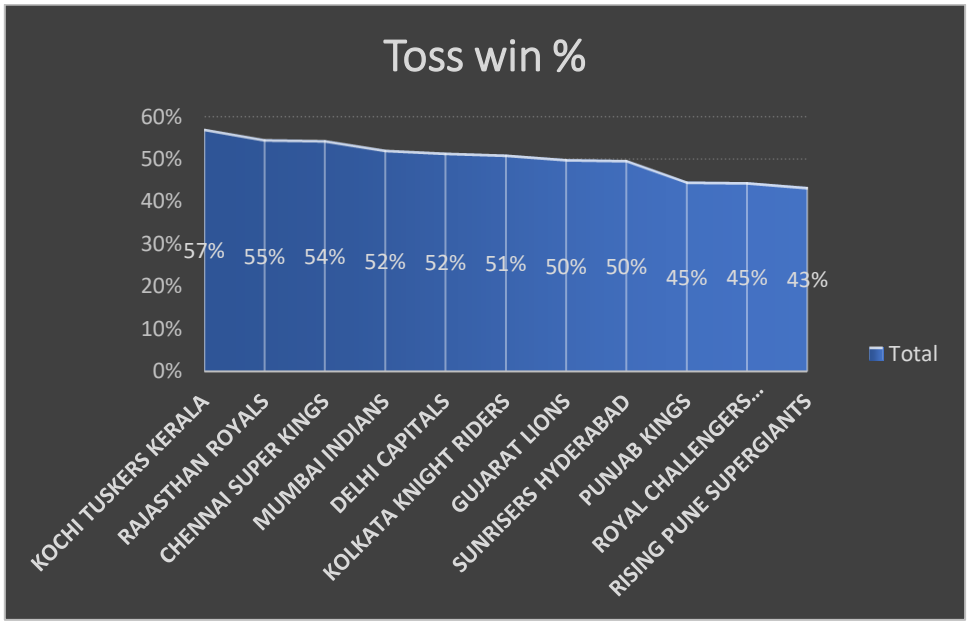

Figure 3.2 Toss win % of all the teams.

3.2.2 Toss win match win %.

The Royal Challengers Bangalore (RCB) has won the toss 87 times and won 43 games as a result. When compared to the other teams, they fall roughly around the middle of the pack. They are not able to take advantage of a toss like Mumbai Indians and Chennai Super Kings. To conclude, When RCB wins the coin toss, they win 49% of the time. Which is not a gratifying figure to look at. Figure 3.3 depicts A team's percentage of wins when the toss goes their way.

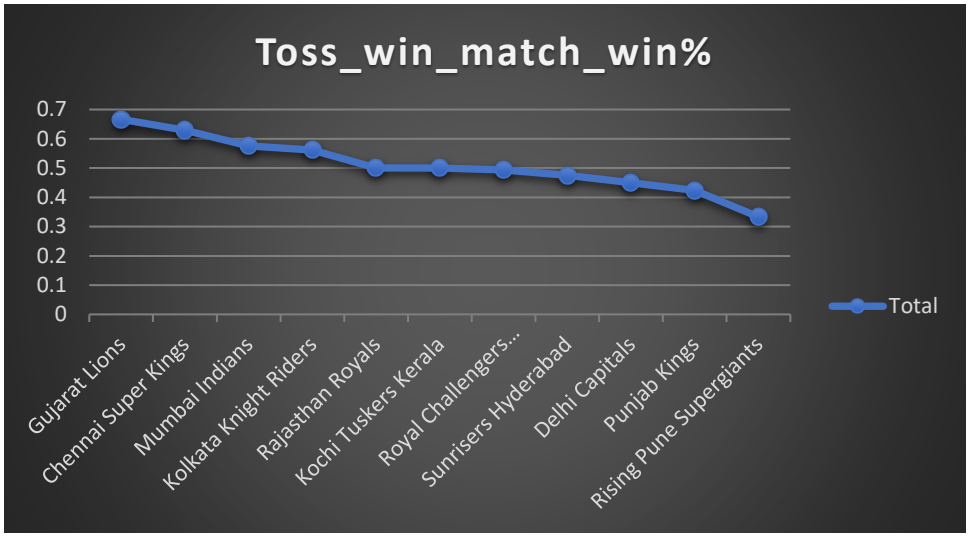

Figure 3.3 The percentage of matches won by teams that won the coin toss.

3.3 Toss decisions.

The choice of whether to bat or field first is vital to the outcome of a cricket match. For example, When day-night or night games are played, the dew or moisture on the field wets the ball as it travels through the ground. Such a ball tends to get slick and harder to grasp with time, and so has a direct effect on the bowler's line and length, placing bowlers at a disadvantage (Dawson, et al., 2009).

3.3.1 Bat first and match win percentage.

When RCB batted first, they won 26% of their matches. Which is very low in comparison to the successful Chennai Super Kings and Mumbai Indians teams. It implied that either batsman are unable of setting a target or bowlers are incapable of defending the total. The percentage of victories when the team bats first are shown in Figure 3.4.

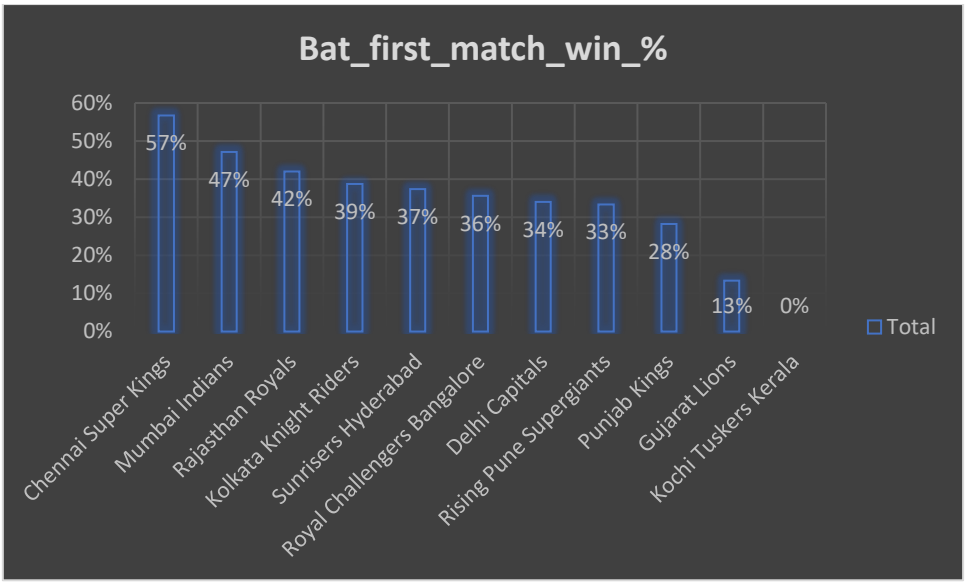

Figure 3.4 Percentage of wins when team bat first.

3.3.2 Field first and match win %.

63 of the 87 games in which the Royal Challengers Bangalore won the toss, they bowled first. In other words, when the RCB chose to bowl first, they have won 69 percent of their matches. The data suggests that they like to chase the target. Also, the batting strength of the club has been found to be satisfactory. Figure 3.5 shows the percentage of wins when the team batted second.

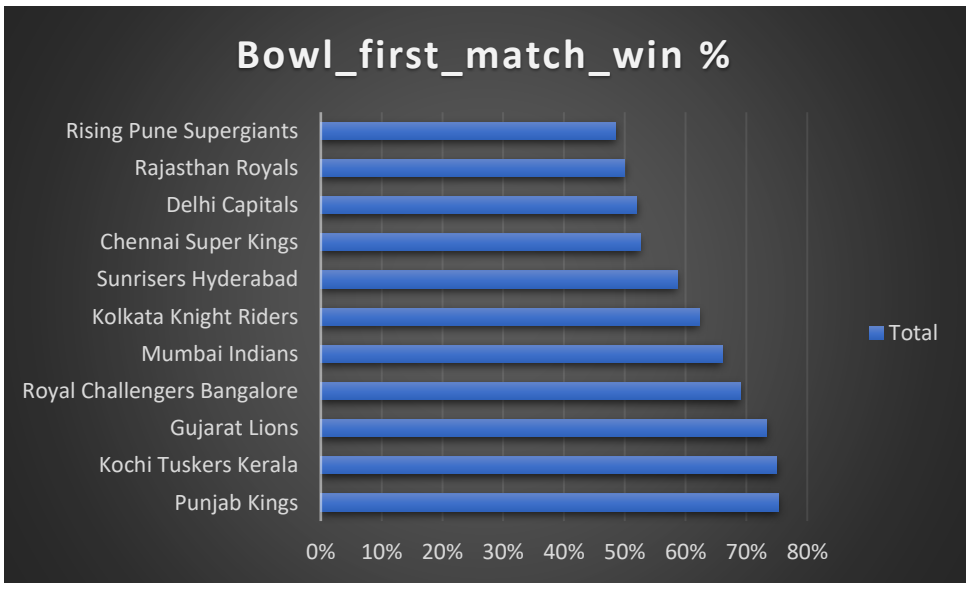

Figure 3.5 percentage of wins when team field first.

3.4 Home ground.

Tens of thousands of spectators are common in cricket-crazed countries like India, and this has a dent in the confidence of the visiting side as the crowd cheers for the home team.

The host team usually has an advantage over the visiting team because of the way the pitches are constructed (Singh, 2014). It is kind of like putting their skills to the test against someone else. For example, In the subcontinent, the pitches will assist spinners. However, in Australia, England pitches will support seamers (fast bowlers). In cricket, the home side will have an edge over the opponent since the opponent will be forced to play to the host team's strengths.

RCB has played 79 matches in Bangalore, winning 36 of them. This demonstrates that the squad is not capitalizing on the home advantage well. The proportion of matches won by RCB on their home ground is shown in Figure 3.6.

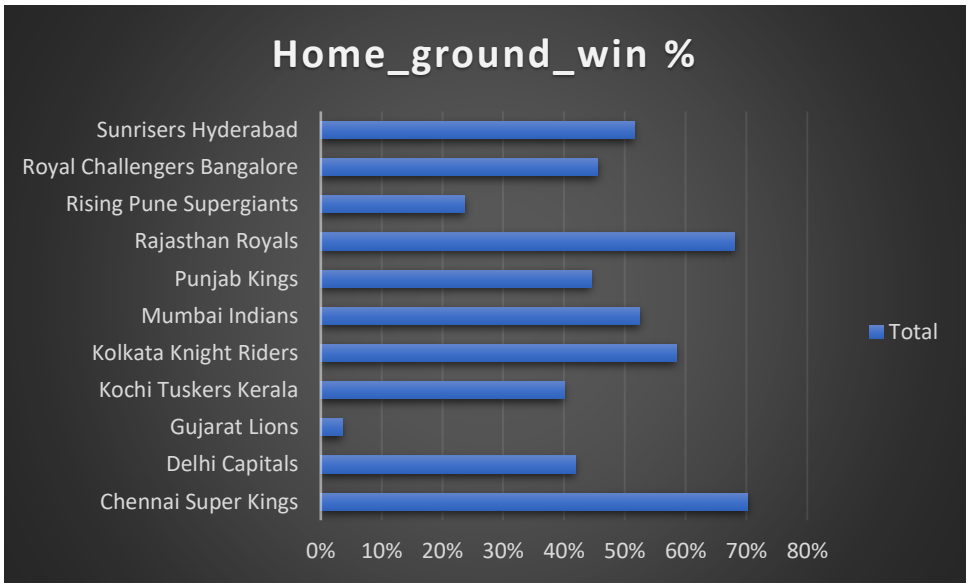

Figure 3.6 Percentage of wins by teams at their home pitch.

4. Opportunity at Upcoming Auction.

The second part of the assignment deals with the different datasets. It is wholly about the players. The dataset is prepared manually from the different datasheets available from the Cricmetric (Cricmetric, 2021). The statistics of players from 2019 to 2021 are investigated. RCB will be able to acquire players at the auction with the assistance of

a complete player analysis, which will also give them an idea of the player's projected price at the auction.

4.1 Rules for the Auction.

- Everyone on the squad must be released and a new team must be assembled from the ground up (IPL, 2021).
- From the 57 Crore INR, Royal Challengers Bangalore must acquire high-quality players (IPL, 2021).
- A maximum of four players can be retained by Royal challengers Banglore (IPL, 2021).

4.1 Data Preparation.

It is defined as the collection, cleansing, and consolidation of data into one file or data table for analytic purposes (Oracle, 2021).

Data cleaned and summarized by some of the excel functions such as summits, averageifs, table, ifna, vlookup, etc.

4.2 Model Building.

It is the visual representations of an information system or its pieces that are used to convey the relationships between data items and structures (Eckerson & Wayne, 2006).

An initial download of the data has no year column. Each year's batting and bowling statistics were stored in a separate file. Despite this, the player's name remained the same in all of the Excel spreadsheets. All of the data is consolidated into a single file and the ERD is created as explained before. The dashboard is created by using tools such as slicers, pivot tables, pivot charts.

4.3 Retention.

RCB can retain a maximum of 3 players from the squad. To retain the players It is important to compare between the players. The key is to look for the recent form of players in IPL.

Runs Above Average (RAA) – RAA quantifies the performance of a batsman/bowler, In terms of how many more runs, a batsman/bowler contributed above an ordinary

player in a particular year/season (Cricmetric, 2021). Table 4.1 shows the top performer of the RCB. The table depicts the top 5 players based on RAA.

Year	Player	Team	RAA
2021	HV Patel	Royal Challengers Bangalore	330
2021	GJ Maxwell	Royal Challengers Bangalore	179
2021	YS Chahal	Royal Challengers Bangalore	95
2021	D Padikkal	Royal Challengers Bangalore	83
2021	V Kohli	Royal Challengers Bangalore	77

Table. 4.1 Top players who are eligible to retain.

On 30th November 2021, The Royal challengers Bangalore has announced that they have retained Virat Kohli, Mohammad Siraj, and Glenn Maxwell for the upcoming IPL 2022 (Royal Challengers Bangalore, 2021). Two of the three players in the above table have been retained, which validates the data analysis above.

4.4 Players' performance.

Two or more players may be compared in a single chart. Which allows for a greater degree of freedom in assessing a player's overall performance. As an example, the number of runs batters score every year. It is shown in Figure 4.2.

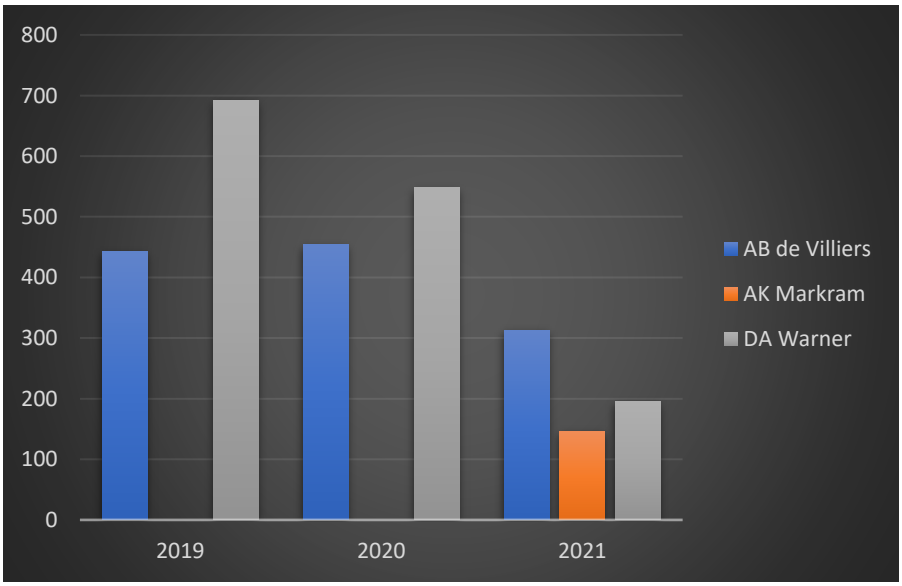

Figure 4.2 shows runs scored by 3 players in each year.

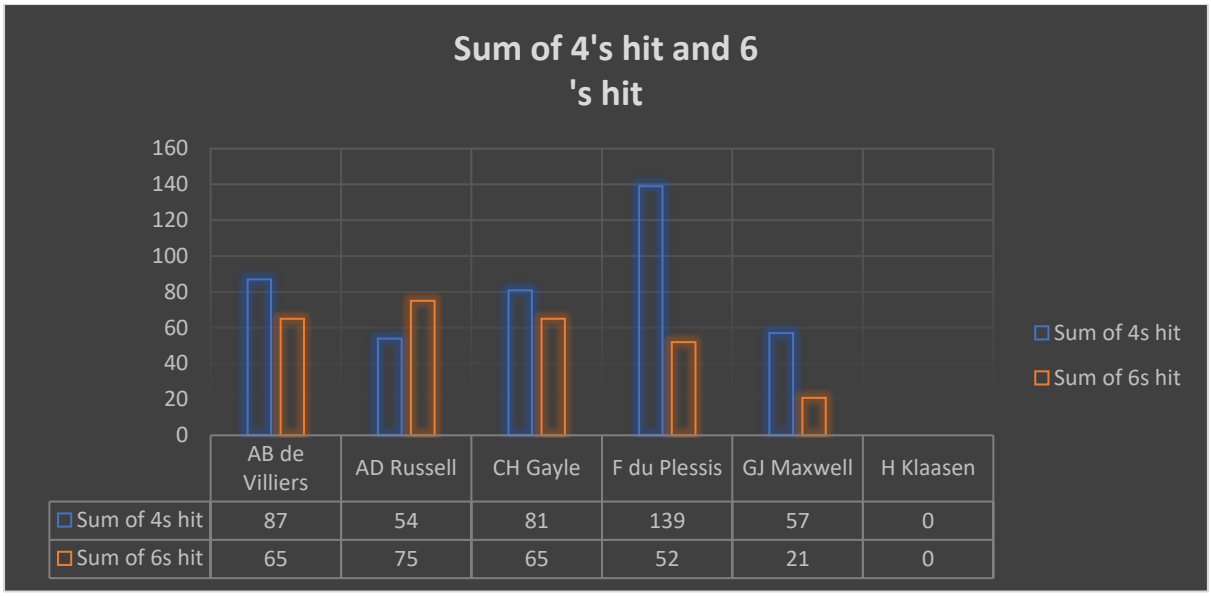

Figure 4.3 shows the number of 4's and 6's hit by the batsmen.

4.5 Expected price for a player in Auction.

This problem can be solved more elegantly by using machine learning algorithms. The task is better accomplished by using predictive analysis. An approximate auction price of players is calculated using the simple Averageifs function in excel. In the upcoming 2022 auction, the players projected bidding prices are shown in Figure 4.4.

Figure 4.4 the players projected bidding prices for 2022.

5. Conclusion.

In today's world, data is the new oil. Data analysis applies to a wide variety of sectors, including healthcare, sports, and others. The first section of the data analysis reveals that Royal challengers Banglore have a poor percentage of home victories. As can be seen, the toss has also played a significant role in winning games. When they bat first, the squad has lost a lot of matches. The squad lost an inordinate number of matches despite winning the toss. However, They have a bright opportunity to pick valuable players from the auction 2022 with 57 crores remaining in their purse. To make an informed decision, a franchisee must first compare players before entering into the auction. The data analytics in retention policy was on point as RCB retained the top 2 players out of 5 players suggested.

6. References

Cadle, J., Paul, D. & Turner, P., 2014. *Business analysis techniques: 99 essential tools for success British Computer Society (BCS.* London: s.n.

Cricmetric, 2021. *Cricmetric.* [Online]
Available at: http://www.cricmetric.com/blog/glossary/
[Accessed 13 December 2021].

Cricmetric, 2021. *IPL.* [Online]
Available at: http://www.cricmetric.com/ipl/
[Accessed 10 December 2021].

Dasararaju, H. & Taori, P., 2019. Data Management—Relational Database Systems (RDBMS). In: Pochiraju B., Seshadri S. (eds) Essentials of Business Analytics. International Series in Operations Research & Management Science,. *International Series in Operations Research & Management Science,* Volume 264, pp. 41-69.

Dawson, P., Morley, B., Paton, D. & Thomas, D., 2009. To bat or not to bat: An examination of match outcomes in day-night limited overs cricket,. *Journal of the Operational Research Society,* 60(12), pp. 1786-1793.

Eckerson & Wayne, W., 2006. *Performance dashboards: measuring, monitoring, and managing your business.* New Jersey: John Wiley & Sons.

Economic Times, 2021. *The biggest T20 league in the world.* [Online]
Available at: https://economictimes.indiatimes.com/news/sports/big-growth-ipl-2021-viewership-count-hits-380-million-mark/the-biggest-t20-league-in-the-world/slideshow/86674520.cms
[Accessed 4 December 202].

Graeme, S. & Nargiza, B., 2012. Achieving benefits with business analytics systems: an evolutionary process perspective,. *Journal of Decision Systems,* 21(3), pp. 231-244.

HBR, 2021. *WHAT'S THE DIFFERENCE BETWEEN DATA ANALYTICS & DATA SCIENCE?.* [Online]
Available at: https://online.hbs.edu/blog/post/data-analytics-vs-data-science
[Accessed 02 December 2021].

Hindustan Times, 2021. *IPL winners.* [Online]
Available at: https://www.hindustantimes.com/cricket/ipl/indian-premier-league-history-winners-list
[Accessed 1 December 2021].

IBM, 2021. *What is Business Intelligence..* [Online]
Available at: https://www.ibm.com/uk-en/analytics/business-intelligence
[Accessed 02 December 2021].

IPL, 2021. *2021.* [Online]
Available at: https://www.iplt20.com
[Accessed 4 December 2021].

IPL, 2021. *Player Retension.* [Online]
Available at: https://www.iplt20.com/news/3751/vivo-ipl-2022-player-retention
[Accessed 5 December 2021].

Kaggle, 2021. *IPL complete dataset (2008 - 2020).* [Online]
Available at: https://www.kaggle.com/patrickb1912/ipl-complete-dataset-20082020
[Accessed 1 December 2021].

Li, Q. & YL., C., 2009. Entity-Relationship Diagram. In: Modeling and Analysis of Enterprise and Information Systems. *Springer.*

Morgulev, E., Azar, O.H. & Lidor, R. Sports analytics and the big-data era. Int J Data Sci Anal 5, 213–222 (2018).

Oracle, 2021. *Analytics.* [Online]
Available at: https://blogs.oracle.com/analytics/post/what-is-data-preparation-and-why-is-it-important
[Accessed 20 December 2021].

Royal Challengers Bangalore, 2021. *Squad.* [Online]
Available at: https://www.royalchallengers.com/rcb-squad
[Accessed 10 December 2021].

Russom, P., 2011. Big data analytics.. *TDWI best practices report, fourth quarter,* 19(4), pp. 1-34.

Sharda, R., Turban, E. & Delan, D., 2014. *Business intelligence and analytics; systems for decision Support.* Boston: Pearson..

Singh, S. B., 2014. Cricket pitches—Science behind the art of pitch making. *International Journal of Science and Research,* 3(7).

Srinivasa, S. & Bhatnagar, V., 2012. Big data analytics In Proceedings of the first International Conference on Big Data Analytics BDA .. pp. 24-26.

The Economic Time, 2021. *Toss is boss.* [Online]
Available at: https://economictimes.indiatimes.com/news/sports/when-toss-is-boss-cricket-loses/articleshow/87724430.cms?from=mdr
[Accessed 5 December 2021].

Winston, W. L., 2010. *Excel : Data analysis and business modelling, Microsoft Press,.* Washington, USA.: s.n.

Wisniewski, M., 2010. *Quantitative methods for decision makers with math.* s.l.:Pearson Education..