# To what extent is there a correlation between the auction price of a batsman and their strike rate in the Indian Premier League (IPL)?

## Introduction

Mathematics can be found and applied in all aspects of life and plays a significant role in the sports industry, within football, hockey, swimming and many more. In cricket, the foundations of the game are based on mathematics and hence, due to my passion for the sport, I have chosen to focus on the mathematical correlation between the strike rate of batsmen and their auction price in my favourite tournament of the year, the IPL.

## Background Information

The Indian Premier League (IPL) is a domestic professional T20 cricket tournament, established in 2008. It is organised by the BCCI and is based on a round-robin group and knockout format with ten teams representing major cities and states in India.

I have chosen to focus on the IPL since it includes players from all over the world, allowing for a more generalised set of batsmen rather than having them all from the same country. The league is played in the twenty over format of the game, which often means batsmen score runs at a faster rate than in other formats; the average strike rate is therefore higher, which will allow me to notice a correlation more easily. Moreover, the IPL is one of the most recently set up tournaments in the world, so its data represents the current playing style of cricket. After the tournament finishes every year, each team has the opportunity to retain players or release them into the auction for the next year, in which they are bid for and left sold or unsold. This is known as the auction price.

## Hypothesis

In this exploration, it is assumed that if a player gets paid well, they should perform well in the current season to justify their auction price; performing well often means they have a higher strike rate. This means a positive correlation should exist between auction price and strike rate in the same season. I am solely focusing on the correlation for batsmen, not bowlers.

## Strike Rate

The strike rate of a batsman is the average number of runs they would score after facing 100 deliveries in a match. Therefore, the higher their strike rate is, the faster and more aggressively they score runs.

Strike rate is calculated using the following formula:

$$Strike\ Rate\ =\ \frac{Number\ of\ Runs\ Scored}{Number\ of\ Balls\ Faced} \times 100$$

**Pearson Correlation Coefficient**

The Pearson Correlation Coefficient is a measure of the strength of a correlation and the association between a dependent and independent variable. The coefficient is labelled and, in simple terms, indicates how far away the data points on a graph are away from the line of best fit. The value of can be between –1 and 1: a value of 0 or close to zero indicates none or little correlation between the two variables; a value below 0 indicates there is a negative correlation; a value above 0 indicates a positive correlation exists. Hence, the closer the Pearson Correlation Coefficient is to –1 or 1, the stronger the correlation. can be calculated using the following formula:

$$r = \frac{\sum (x - \overline{x})(y - \overline{y})}{\sqrt{\sum (x - \overline{x})^2\ \sum (y - \overline{y})^2}}$$

In this formula, $x$ is the value of the independent variable, the auction price, and $\overline{x}$ is the mean of all the auction prices. $y$ is the value of the dependent variable, the strike rate, and $\overline{y}$ is the mean of all the strike rates.

I am going to be using this method to explore the strength of this correlation and identify if it exists at all between two separate variables.

**Linear Regression**

To find the equation of the trend line, I will use my Graphical Display Calculator to find out the linear regression in the format $y = mx + c$ . This will allow me to assess if the gradient is positive or negative and hence identify any positive or negative correlation between the two variables.

**Data Set**

The table on the next page displays the raw data for the top 35 batters with the highest strike rates in the 2023 IPL season. In the third column, I have displayed the auction prices that they have been bought/retained for in the same season (2023).

| Number | Player | 2023 Strike Rate | 2023 Auction Price (₹ Indian Rupee) |
|---|---|---|---|
| 1 | Glenn Maxwell | 183.48 | 11,00,00,000 |
| 2 | MS Dhoni | 182.45 | 12,00,00,000 |
| 3 | Suryakumar Yadav | 181.13 | 8,00,00,000 |
| 4 | Glenn Phillips | 177.27 | 1,50,00,000 |
| 5 | Heinrich Klaasen | 177.07 | 5,25,00,000 |
| 6 | Nicholas Pooran | 172.94 | 16,00,00,000 |
| 7 | Dhruv Jurel | 172.72 | 20,00,000 |
| 8 | Ajinkya Rahane | 172.48 | 50,00,000 |
| 9 | Shahrukh Khan | 165.95 | 9,00,00,000 |
| 10 | Krishnappa Gowtham | 164.86 | 90,00,000 |
| 11 | Tilak Varma | 164.11 | 1,70,00,000 |
| 12 | Phil Salt | 163.90 | 2,00,00,000 |
| 13 | Yashasvi Jaiswal | 163.61 | 4,00,00,000 |
| 14 | Liam Livingstone | 163.15 | 11,50,00,000 |
| 15 | Shardul Thakur | 161.42 | 10,75,00,000 |
| 16 | Cameron Green | 160.28 | 17,50,00,000 |
| 17 | Vijay Shankar | 160.10 | 1,40,00,000 |
| 18 | Shivam Dube | 158.33 | 4,00,00,000 |
| 19 | Tim David | 158.21 | 8,25,00,000 |
| 20 | Shubman Gill | 157.80 | 8,00,00,000 |
| 21 | Jitesh Sharma | 156.06 | 20,00,000 |
| 22 | Faf Du Plessis | 153.68 | 7,00,00,000 |
| 23 | Sanju Samson | 153.38 | 14,00,00,000 |
| 24 | Rahul Tewatia | 152.63 | 9,00,00,000 |
| 25 | Shimron Hetmyer | 152.28 | 8,50,00,000 |
| 26 | Jason Roy | 151.59 | 2,80,00,000 |
| 27 | Simran Singh | 150.42 | 60,00,000 |
| 28 | Marcus Stoinis | 150.00 | 10,00,00,000 |
| 29 | Rinku Singh | 149.52 | 55,00,000 |
| 30 | Rilee Rossouw | 148.22 | 4,60,00,000 |
| 31 | Ruturaj Gaikwad | 147.50 | 6,00,00,000 |
| 32 | Vivrant Sharma | 146.80 | 2,60,00,000 |
| 33 | Venkatesh Iyer | 145.84 | 8,00,00,000 |
| 34 | Andre Russell | 145.51 | 16,00,00,000 |
| 35 | David Miller | 145.50 | 3,00,00,000 |

**Initial Calculations**

Using this data, I can calculate the mean strike rate and the standard deviation. To calculate the mean, I used the following formula:

$$\overline{x} = \frac{\sum x}{n}$$

$$\overline{x} = \frac{183.48 + 182.45 + 181.13 + \cdots + 145.51 + 145.50}{35}$$

$$\overline{x} = \frac{5611.19}{35}$$

$$\overline{x} = 160.32\ (2\ d.p.)$$

Standard deviation is a measure of how spread-out the values are, using the difference from the mean. There are two formulas which can be used – one for population and the other for a sample. In this case, my values are strike rates and hence, I am going to use the sample standard deviation formula, which is as follows:

$$\sigma = \sqrt{\frac{\sum (x - \overline{x})^2}{n - 1}}$$

$$\sigma = \sqrt{\frac{(183.48 - 160.32)^2 + (182.45 - 160.32)^2 + \cdots + (145.51 - 160.32)^2 + (145.50 - 160.32)^2}{34}}$$

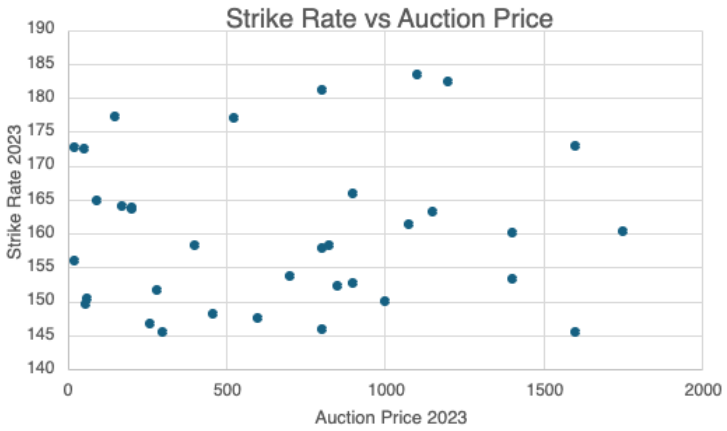
$$\sigma = \sqrt{\frac{4376.0137}{34}}$$

$$\sigma = 11.34\ (2\ d.p.)$$

The mean strike rate of all the batsmen in 2023 is 160.32 and the standard deviation is 11.34. Since the value of the standard deviation is relatively high, there is a wide range of values for the strike rate, suggesting that every player's strike rate varies from the others.

**Graphing the data**

I created a scatter graph to display the data visually and identify if there was a correlation between the variables. The graph is as shown below:




Using this graph, I could not identify a noticeable correlation. I had to work out the Pearson Correlation Coefficient for this graph, using the equation. Firstly, I had to find the mean auction price for the players in my list. This equates to ₹ 6,77,00,000.

$$r = \frac{2728}{\sqrt{(4375.985)\,(90{,}21{,}655)}}$$

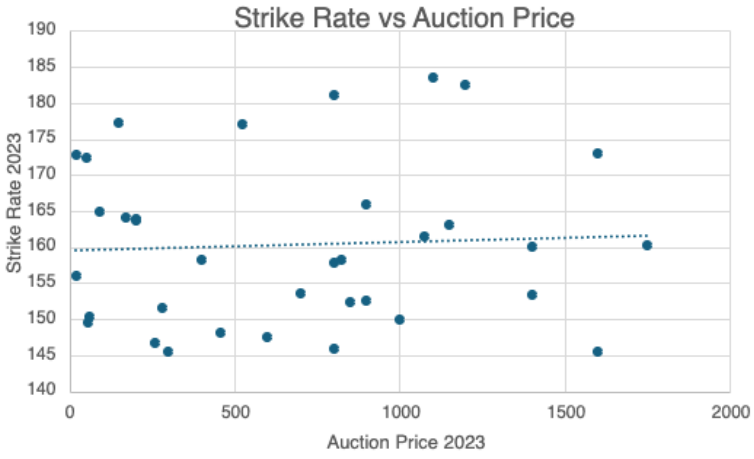
$$r = 0.01373$$

The value of $r$ is close to zero, which nullifies the existence of any linear correlation between the strike rate and auction price.

I used my GDC to work out the linear regression line, which resulted in the equation shown below:

$$y = 0.001119x + 159.53$$

From the graph and this equation, I can see that a slight positive increasing correlation exists, (with a lot of outliers). This can also be seen as the trend line on the graph below:



**New Hypothesis**

Due to the lack of correlation, I decided to explore another hypothesis, using different variables: the strike rate of the same players in the previous 2022 season vs the auction price in the 2023 season. It is assumed that a more successful performance with a higher strike rate in the previous season will lead to a higher auction price in the current season. This would lead to a positive correlation between the auction price and the strike rate.

**Second Data Set**

I had to research the strike rates for the same players in 2022 against their 2023 auction prices. Some of the players listed above were not involved in the 2022 season so I had to remove them from the list, leaving me with the information on 25 batters rather than 35. The following table displays the raw data for 2022 strike rates and 2023 auction prices:

| Number | Player | 2022 Strike Rate | 2023 Auction Price (₹ Indian Rupee) |
|---|---|---|---|
| 1 | Glenn Maxwell | 169.10 | 11,00,00,000 |
| 2 | MS Dhoni | 123.40 | 12,00,00,000 |
| 3 | Suryakumar Yadav | 145.67 | 8,00,00,000 |
| 4 | Nicholas Pooran | 144.33 | 16,00,00,000 |
| 5 | Ajinkya Rahane | 103.90 | 50,00,000 |
| 6 | Shahrukh Khan | 108.33 | 9,00,00,000 |
| 7 | Tilak Varma | 131.02 | 1,70,00,000 |
| 8 | Yashasvi Jaiswal | 132.98 | 4,00,00,000 |
| 9 | Liam Livingstone | 182.08 | 11,50,00,000 |
| 10 | Shardul Thakur | 137.93 | 10,75,00,000 |
| 11 | Shivam Dube | 156.21 | 4,00,00,000 |
| 12 | Tim David | 216.27 | 8,25,00,000 |
| 13 | Shubman Gill | 132.32 | 8,00,00,000 |
| 14 | Jitesh Sharma | 163.63 | 20,00,000 |
| 15 | Faf Du Plessis | 127.52 | 7,00,00,000 |
| 16 | Sanju Samson | 146.79 | 14,00,00,000 |
| 17 | Rahul Tewatia | 147.61 | 9,00,00,000 |
| 18 | Shimron Hetmyer | 153.92 | 8,50,00,000 |
| 19 | Marcus Stoinis | 147.16 | 10,00,00,000 |
| 20 | Rinku Singh | 148.71 | 55,00,000 |
| 21 | Ruturaj Gaikwad | 126.46 | 6,00,00,000 |
| 22 | Venkatesh Iyer | 107.69 | 8,00,00,000 |
| 23 | Andre Russell | 174.47 | 16,00,00,000 |
| 24 | David Miller | 142.72 | 3,00,00,000 |

**Second Round of Calculations**

Using this data, I can calculate the mean strike rate and the standard deviation. The mean is as follows:

$$\overline{x} = \frac{169.10 + 123.40 + 145.67 + \cdots + 174.47 + 142.72}{24}$$

$$\overline{x} = \frac{3470.22}{24}$$

$$\overline{x} = 144.59\ (2\ d.p.)$$

The standard deviation for the new set of values for strike rate is:

$$\sigma = \sqrt{\frac{(169.10 - 138.81)^2 + (123.40 - 138.81)^2 + \cdots + (174.47 - 138.81)^2 + (142.72 - 138.81)^2}{23}}$$

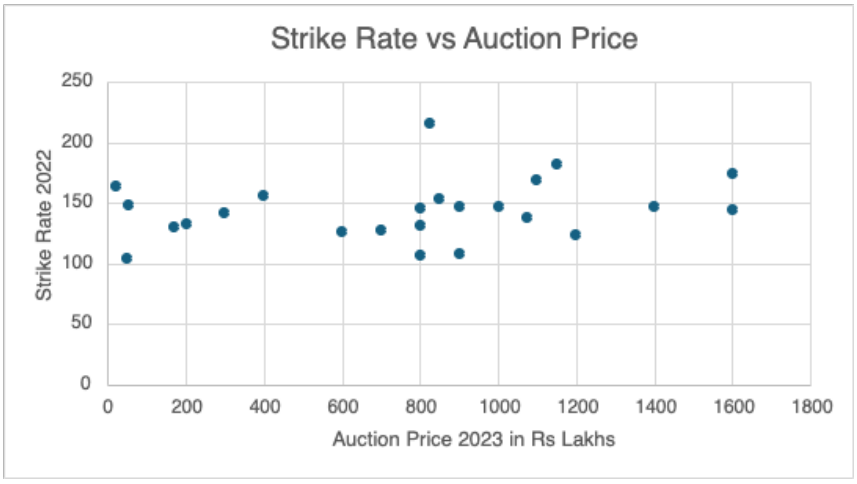
$$\sigma = \sqrt{\frac{15381.4548}{23}}$$

$$\sigma = 25.86\ (2\ d.p.)$$

The mean strike rate of the batsmen in 2022 is 144.59 and the standard deviation is 25.86. Since the value of the standard deviation is high, there is a wide range of values for the strike rate. The standard deviation for this new set of values is higher than before, meaning the values in this set are more varied.

**Graphing Second Set of Data**

I created another scatter graph for the new variables, which resulted in the graph below:



In this graph, there is a more noticeable correlation and there are less outliers than with the correlation between strike rate and auction price in 2023. In order to gain a better understanding of the correlation, I worked out the Pearson Correlation Coefficient for the new set of values:

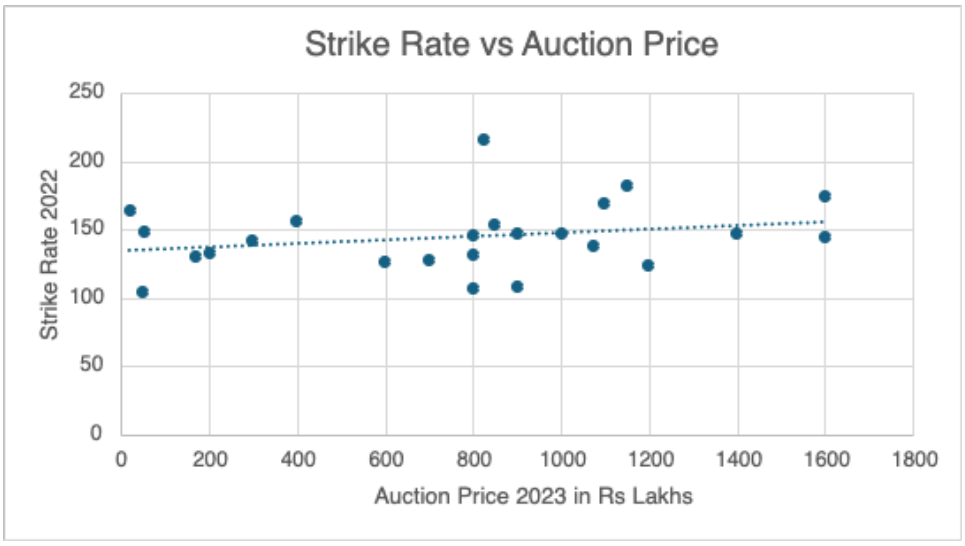
$$r = \frac{64365.26}{\sqrt{(14578.96)\,(50{,}73{,}366)}}$$

$$r = 0.23667$$

This $r$ value is still close to zero, so the correlation is still a bit weak but it is greater than the previous value, indicating that more of a correlation exists between the 2022 strike rate and the 2023 auction price. It is a positive value, so the correlation is increasing; the auction price increases as the strike rate from the previous season increases suggesting a player's previous success leads to a higher auction price in the future. I then used my GDC to find the linear regression line. The following equation was displayed:

$$y = 0.012687x + 134.82$$

This line can be seen as a trend line in the following graph:




**Limitations**

During this exploration, there have been a few limitations. Firstly, cricket as a game is affected by different parameters, which have an impact on a batsman's strike rate. These include the weather, pitch quality, the bowler and the game's situation. For this analysis to be perfectly accurate, all these parameters must be controlled, which cannot be guaranteed in my exploration. Different batsmen also have different techniques and game plans which affect their strike rate as well. Hence, there are many other extraneous variables affecting high and low auction prices.

**Conclusion**

Ultimately, I posit that there is not a strong correlation between the auction price of a player and their strike rate in the current or previous season in a cricket game. This therefore proves that the price franchises pay for players in the auction is not solely based on the player's strike rate and performance within the IPL, rather it can be affected by other factors including international experience, games at county-level and other leagues around the world e.g. Ranji Trophy in India, Big Bash League in Australia etc. The high values for standard deviation suggest the strike rate of players vary greatly from one another and from the mean; I only used a small sample of batsmen, however, in the overall game of cricket there are bowlers and all-rounders as well, who often have lower strike rates, increasing the range. I calculated two trendlines for the two separate sets of data for the 2022 and 2023 strike rates. Both lines had small positive gradients, drawing a parallel with the Pearson Correlation Coefficients, which were both relatively close to zero, highlighting the extent of the weak correlation. Overall, this exploration of the relationship between auction price and strike rate proved there is a limited association, suggesting high auction prices are often due to other factors such as experience or stylistic technique instead.

**Bibliography**


"IPL Auction 2023 | IPL Auction Live | IPL Auction Updates." *Cricinfo*, www.espncricinfo.com/auction/ipl-auction-2023.

"IPL Auction 2023: IPL Player Auction, Full Player List - NDTV Sports." *NDTVSports.com*, sports.ndtv.com/ipl-2023/auction. Accessed 28 Mar. 2024.

Laerd Statistics. "Pearson Product-Moment Correlation." *Laerd Statistics*, 2020, statistics.laerd.com/statistical-guides/pearson-correlation-coefficient-statistical-guide.php.

Longmore, Andrew. "Indian Premier League | Indian Cricket League." *Encyclopedia Britannica*, 3 Mar. 2024, www.britannica.com/topic/Indian-Premier-League.

ML/AIIshan | Virginia Tech & IIT Delhi | ✍. "Understanding Pearson Correlation." *Medium*, 3 Oct. 2023, ishanjainoffical.medium.com/understanding-pearson-correlation-bebdc9209885.

Paul, William. "What Is Strike Rate in Cricket: The Term Explained." *SportsAdda*, 9 Apr. 2022, www.sportsadda.com/cricket/features/what-is-strike-rate-in-cricket. Accessed 28 Mar. 2024.

"Players Stats and Records | IPLT20." *Www.iplt20.com*, www.iplt20.com/stats/2023.