

# MM217 : Assignment - 4

In this assignment, we use the batting statistics of Sachin Tendulkar and Brian Lara in test matches against five bowlers. The datasets are Our idea is to

	Bowler	Type	Balls	Runs	Wickets
1	Shane Warne	Spinner	279	150	4
2	M Muralitharan	Spinner	366	196	6
3	Brett Lee	Fast	432	242	5
4	Glenn McGrath	Fast	350	208	9
5	Andrew Flintoff	Fast	342	133	2

Figure 1: Sachin's batting stats(Source:espnccricinfo.com)

	Bowler	Type	Balls	Runs	Wickets
1	Shane Warne	Spinner	236	174	2
2	M Muralitharan	Spinner	709	373	3
3	Brett Lee	Fast	243	172	2
4	Glenn McGrath	Fast	233	90	2
5	Andrew Flintoff	Fast	263	109	4

Figure 2: Lara's batting stats (Source:espnccricinfo.com)

make some comparisons about the two batting greats. We have to do some calculations and use bar charts and pie charts for comparing the data. Functions such as `aggregate()`, `mtext()` and `rbind()` will also be handy in places.

1. Plot a bar chart in two rows, showing the number of times each batsmen were out against the five bowlers. The final output of your plot should look like figure 3.
2. Batting average is a stat that is used to compare between batsmen. It is defined as the division of the number of runs scored by the number of times out. Calculate the batting averages for both batsmen against each bowler and plot them in two rows of a bar chart similar to the previous question. The final plot should look like figure 4.

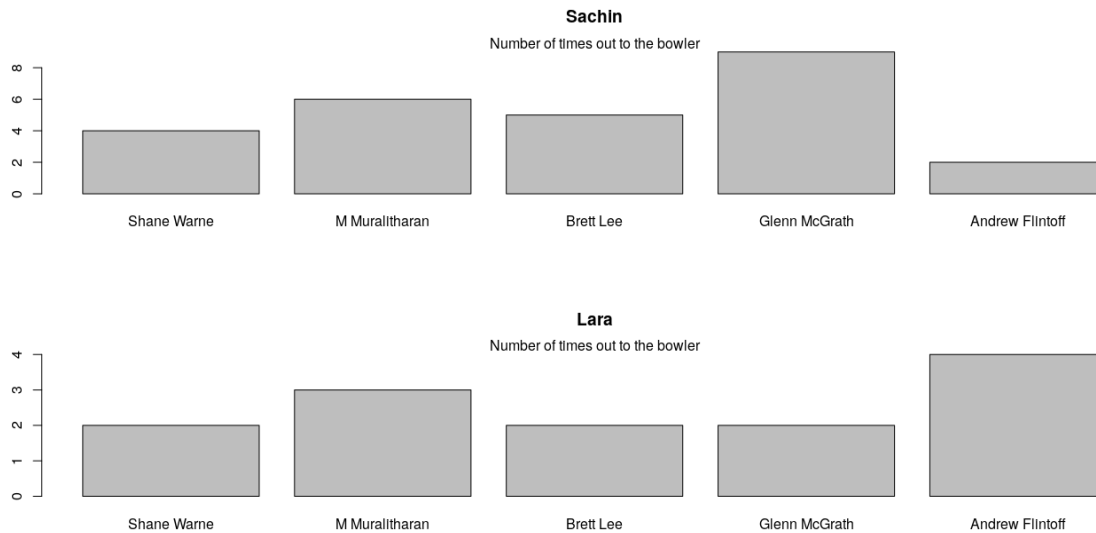


Figure 3: Number of times both batsmen were out against the bowlers



Figure 4: Batting averages against individual bowlers

- It can be seen from Figure 4 that the y-axes are differing for both batsmen and comparison of data is not easy here. Use the function `rbind()` to get the data together and plot the batting average for both batsmen side-by-side in the same plot. The final figure should look like Figure 5.
- From the data, it can be seen that Lara averages better than Sachin against all but one bowlers. This difference is quite evident against the spin bowlers.

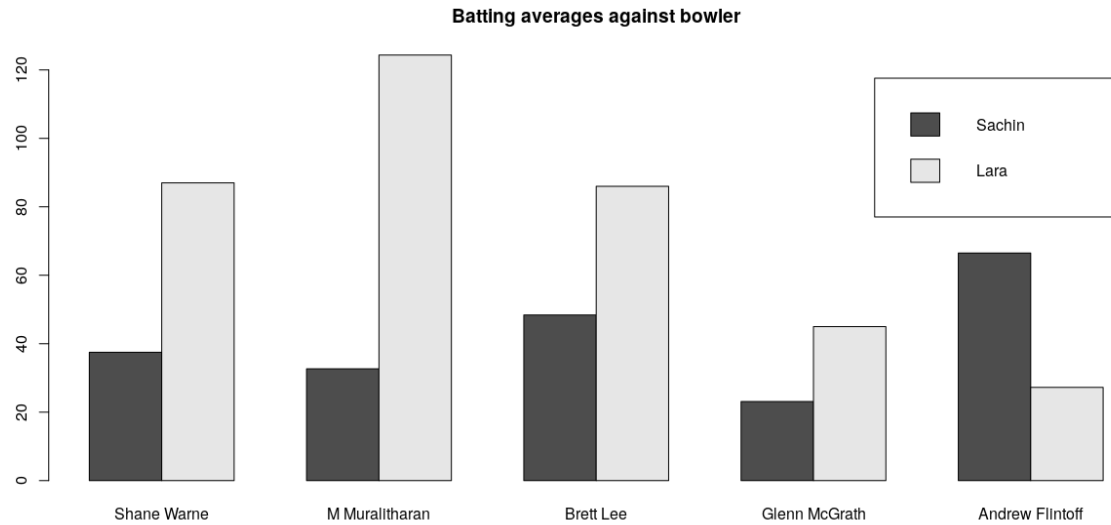


Figure 5: Batting averages by Sachin and Lara against individual bowlers in the same plot

However, we are missing something here : the number of balls faced to score the runs. Use the `aggregate()` function to calculate the number of balls faced against spin and fast bowlers by both batsmen and plot them in two pie charts as shown in Figure 6.

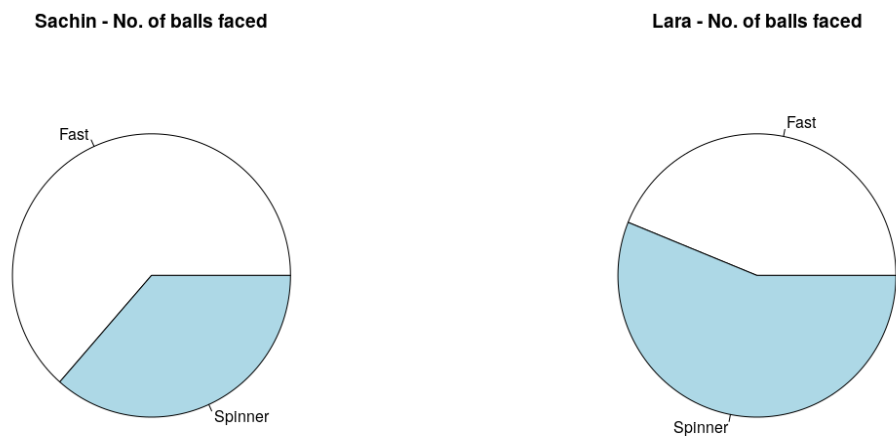


Figure 6: Number of balls faced against spin and fast bowlers

- The statistic used to describe the number of runs per 100 balls is called the strike rate of the batsman. Plot the strike rate of Sachin and Lara against

all five bowlers similar to question 3. The final plot should look like Figure 7.

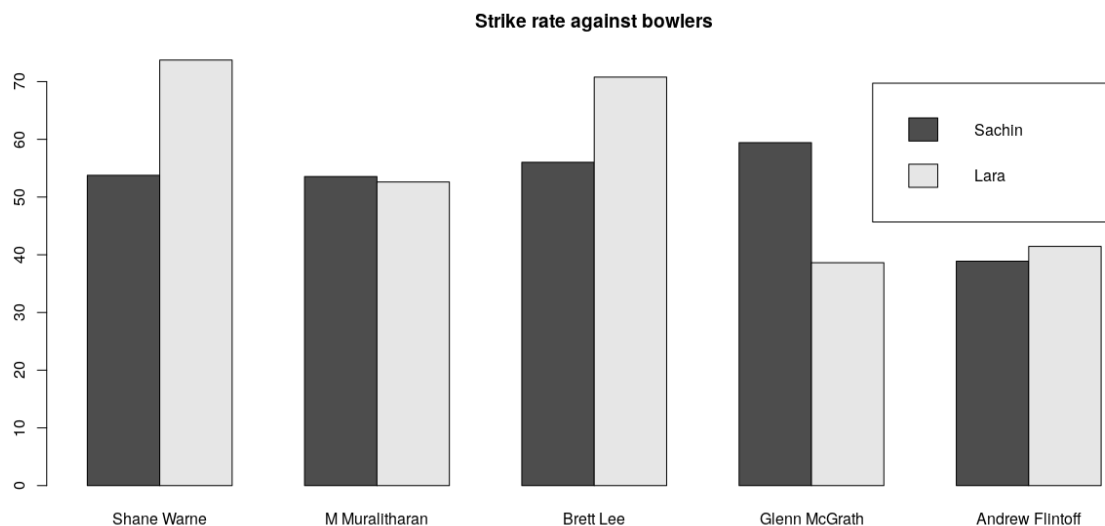


Figure 7: Strike rates of Sachin and Lara against individual bowlers in the same plot

- Use the same `aggregate()` function to plot the batting averages of both batsmen against spin and fast bowlers respectively. The final figure should look like Figure 8.

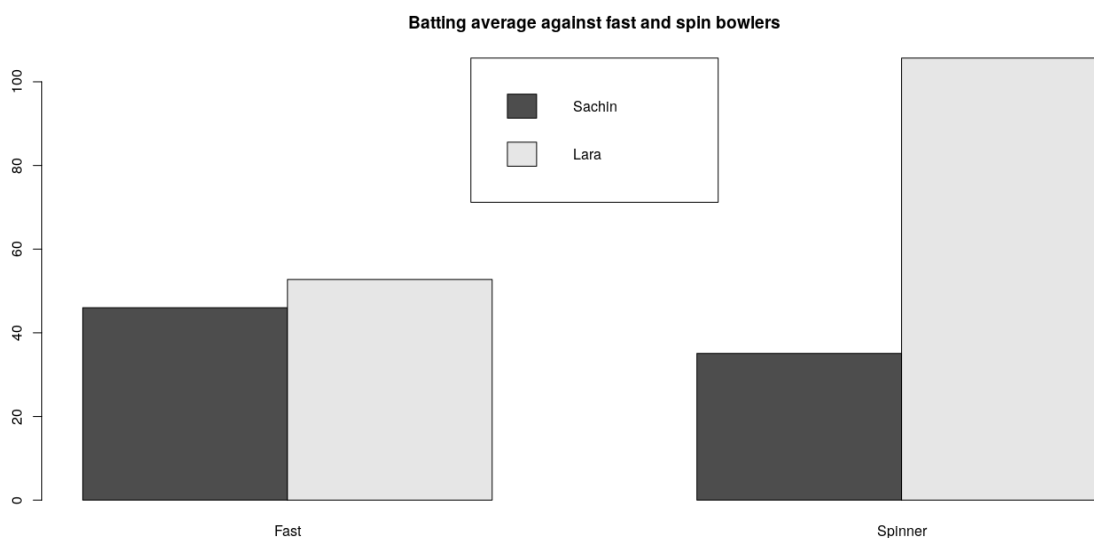


Figure 8: Batting averages against spin and fast bowlers

7. Based on your observations, state whether the following statements made about the data is true or false (or a yes/no):
- a Lara averages higher than Sachin against almost all bowlers because he gets out fewer times than Sachin.
  - b If Sachin and Lara respectively bat for one hour against Glenn McGrath, Sachin is likely to score more runs than Lara.
  - c Based on the strike rate and batting average data, Lara is a better batsman against the spin greats as compared to Sachin.