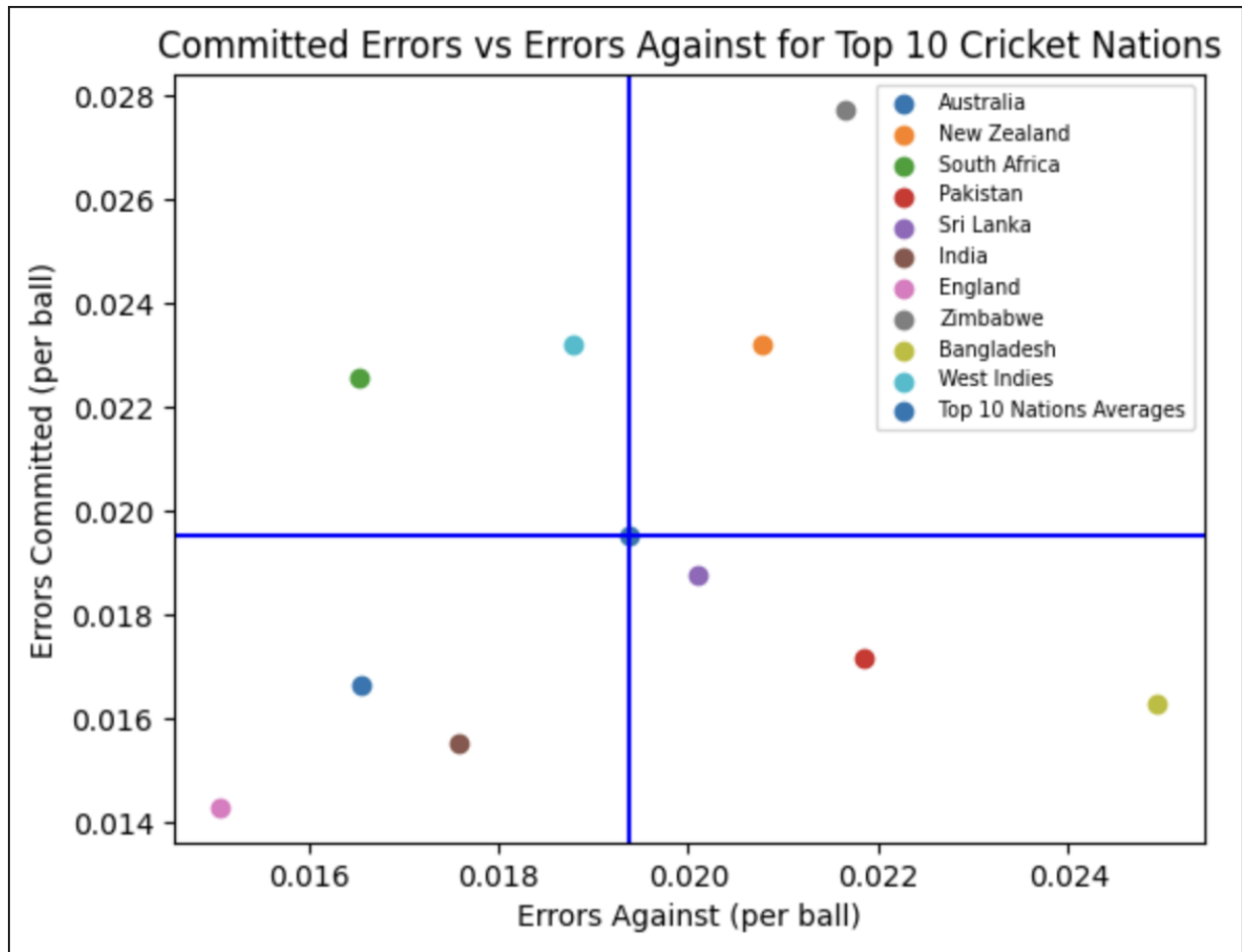To help Coach Ella answer her question about where errors come from in cricket, I looked up the current top 10 ranked nations, and decided to use these as the main focus of my data analysis. These countries are, in no particular order, Australia, New Zealand, South Africa, Pakistan, Sri Lanka, India, England, Zimbabwe, Bangladesh, and the West Indies. I did this to narrow down which teams I would be focusing on, as the data set as a whole is enormous and it would be unreasonable to try to examine every team in it. Additionally, these nations often play against each other, which is ideal because it should theoretically limit the amount of outliers in the data. This is because it is less likely there are extreme amounts of errors in a match as a result of one of the top teams playing against a much lesser opponent.

The first thing I did was find every instance of each of the top 10 countries batting, and then counted the total wides and no-balls and added them. I then divided that error total by the total number of balls that they faced. This gave me the average errors per ball that each nation faced while they were batting. I then repeated this process for all instances in which the top countries were bowling, to find the average errors per ball that they committed. I also found the average of both of these values for the top nations. Finally, I plotted these 11 values in a scatterplot, with errors against on the x-axis and errors committed on the y-axis, as seen in figure 1.1. I also found the average errors committed/faced for all teams in the data set.

Figure 1.1

Committed Errors vs Errors Against for Top 10 Cricket Nations

Legend:
- Australia
- New Zealand
- South Africa
- Pakistan
- Sri Lanka
- India
- England
- Zimbabwe
- Bangladesh
- West Indies
- Top 10 Nations Averages

X-axis: Errors Against (per ball)
Y-axis: Errors Committed (per ball)

* 0.0235 errors per ball for all teams in the data set

As seen from the scatterplot, 4 main categories emerge:

Figure 1.2

Category	Location	Errors Against	Errors Committed
Predictable	Bottom Left	Low	Low
High Margin	Bottom Right	High	Low
Unpredictable	Top Right	High	High
Low Margin	Top Left	Low	High

As expected, the majority of these teams are below the set-wide average for errors committed, and all but one of them faces more errors than average. There is however, a great deal of variation within the top 10 nations, as described in figure 1.2. Figure 1.2 displays the 4 categories which are divided into quadrants by the values of the average errors committed and against for the top 10 nations. Figures 1.1 and 1.2 display interesting data on how many errors teams are facing and committed, but it is difficult to visualize how the values compare to each other for each team. Figure 1.3 shows each nation's error margin, which was obtained by subtracting their errors committed from their errors faced.
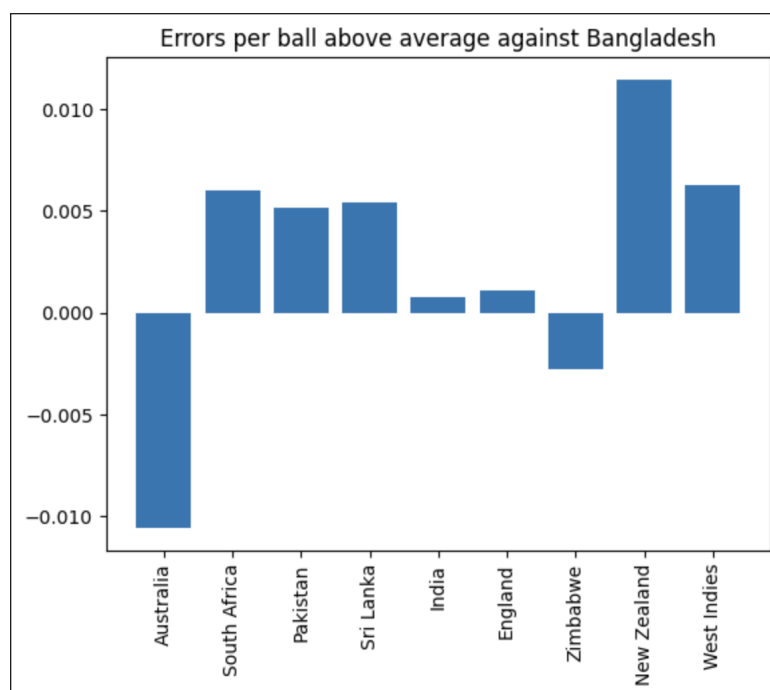
Figure 1.3
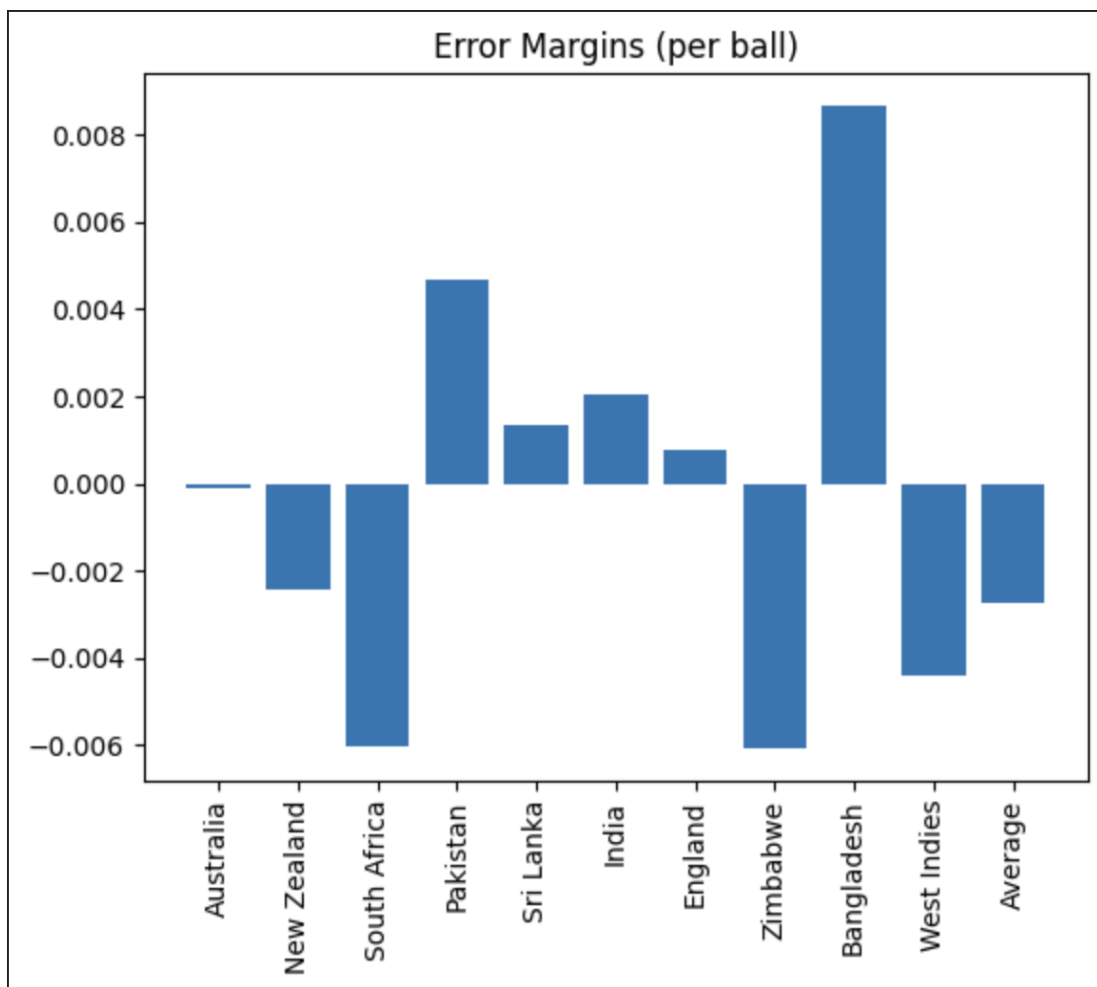
Predictably, the two teams with the highest error margins are both in the "High Margin" quadrant of the scatterplot. What the error margin graph reveals though is that, each of the teams with negative error margins are above average for errors committed. Additionally, both the teams in the "Unpredictable" quadrant have negative error margins. What this suggests is that, despite the fact that teams such as New Zealand and Zimbabwe face a lot of errors, committing errors erases any chance of keeping a positive margin.

Bangladesh is the team with the highest error margin of the top 10 nations, and from the scatterplot, it can be determined that they achieved that margin by facing an enormous amount of errors per ball. To determine if any significant conclusions can be drawn from the high rate of errors that Bangladesh faces, I filtered the data to find every instance of another top 10 nation bowling against Bangladesh. After that, I found the average errors per ball that each team commits against Bangladesh, and then found the difference between that number and the number of errors they committed per ball across the data set. These results are displayed in figure 1.4.

Figure 1.4

From figure 1.4 it is apparent that 7/9 of the other nations committed more errors per ball than they typically do against Bangladesh. Additionally, if it were not for Australia's much lower number of errors against them, Bangladesh's average errors faced would be much higher. Furthermore, despite being 3 of the bottom 5 teams for most errors committed, England, India, and Pakistan all have an increased error rate against Bangladesh. Also, Sri Lanka, the team with the next best error margin after Pakistan and India, has an increased error rate as well. This data all points toward the fact that typically tidy teams that do not commit many errors seem to do so more often against Bangladesh, which would then go on to explain why their error margin might be so high.

As Coach Ella already observed, the teams with the least amount of errors typically perform the best. Additionally, she is probably correct in thinking that aiming to maximize her team's error margin will lead to better results, and the best way to do this is to minimize errors. While it appears that some teams are more prone to making errors than others however, it also appears possible that certain teams influence their opponent in a way that forces them into making errors of their own. Due to both of these facts, I would suggest that Coach Ella employs a dynamic practice schedule that focuses more on one side of the ball depending who her team's upcoming opponent is. Coach Ella should use methods similar to the ones in this analysis to determine which of her opponents have the highest error per ball rate against them. From there she can set a threshold for errors per ball. For all opponents that are above that threshold, she can focus more on bowling in practice to attempt to limit her team's errors against a team that forces a lot of them. For teams that are below the threshold, she should focus more on batting because the other team is less likely to force a great amount of errors and she therefore has a greater

chance of maintaining a positive margin, even without extra practice in reducing errors, as the team is already less likely to make them against these opponents.

Figure 2.1

Errors Per Ball for Each Inning in All Matches
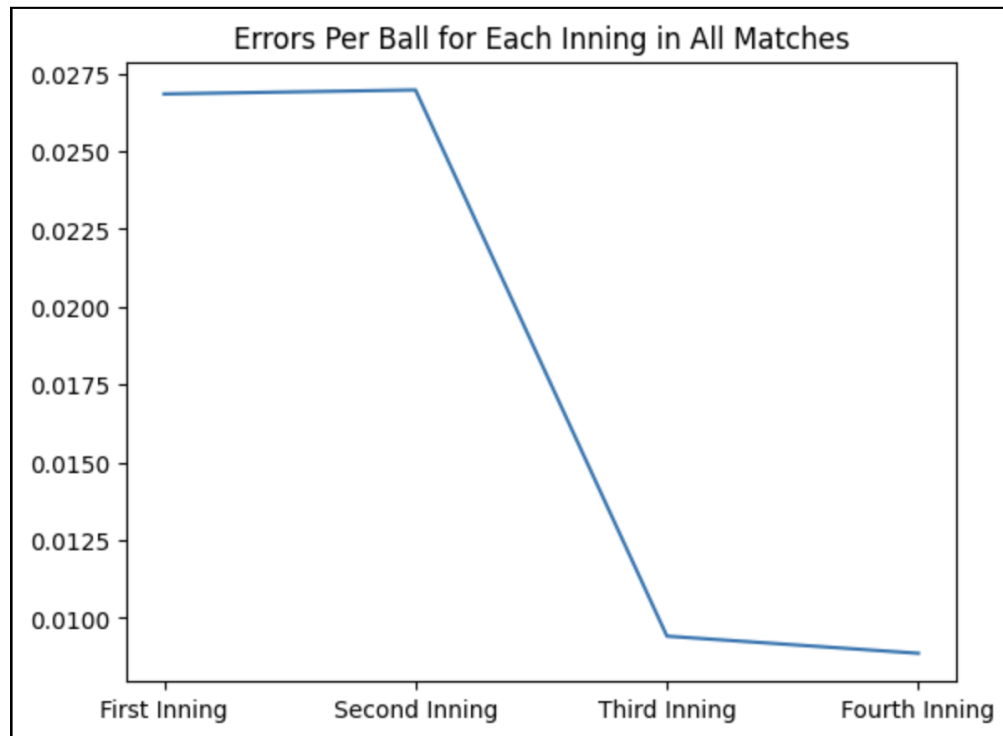
To answer Coach Ella's second question, I split the data set by inning, and then summed the number of wides and no-balls for each inning. I then divided the total number of errors for each inning by the number of balls total from each inning, and got an average error per ball for innings 1-4. The results are displayed in figure 2.1. These data indicate that there is not an increase in errors per ball as the game progresses to the later innings. In fact, the least amount of errors per ball was found in the 4th inning of games. Because of this, I would not suggest that Coach Ella introduces extra cardio training into her practices, and instead focuses on the aforementioned plan.