Visualizing the Outcome of Matches of a Tournament in a Single Display

Dibyojyoti Bhattacharjee[a], Hemanta Saikia[b]

a. Department of Statistics, Assam University, Silchar, India
b. College of Sericulture, Assam Agricultural University, Jorhat, India

Abstract: Visualization has by now become an integral part of modern-day research. With the advances in computation technology data display is extensively used in different fields of research, making data more explanatory and insightful. Graphical displays have also become very popular in sports, especially for television audiences. But with advances in technology, various stakeholders in sports have used data to measure player performance, frame strategy, etc. Statistical graphics and data visualization tools have played a significant role in sports analytics. In this paper, the authors aim to develop a semi-graphical display to visualize the tournament schedule and the outcomes of all matches. To explain how the display works, data from the ICC Twenty20 Cricket World Cup of 2022 was used. In the said tournament, 12 teams participated. Though the cricket tournament is used as an example for the display, the presentation can be extended to other sports as well, where more teams participate. Some potential uses for the display are also discussed.
Keywords: Sports Analytics, Data Visualization, Semi-graphic display.

1. Introduction:

Data Visualization is a technique of expressing meaningful numbers through graphs and diagrams. But time has changed and now graphs are called infographics, which are interactive and dynamic as they are no longer restricted to pages of books and journals but are an integral component of websites and social media platforms. People are now working to develop innovative, user-friendly, interactive displays. The domain of data visualization, which was once merely a tool for data display, has now become a subject of research and advanced-level computer programming.

Visual displays of information have become the norm in modern research. High-quality, sophisticated graphs and diagrams are used to communicate results in most research papers in scientific publications. Today, visualization tools play an important role in all aspects of statistical investigation – it begins with exploratory plots, supports various stages of analysis, and helps in the final communication and display of results (Bhattacharjee, 2007). Data, whether generated from an experiment in physical science or from any socio-economic phenomena, needs to be visualized (Bhattacharjee, 2012). In one of the very early works on visualization, Playfair (1801) commented that the purpose of data visualization is to make statistics a little more palatable. But with technological advances and computational power, visualization has become an integral part of data presentation and analysis. In addition to scientific research, visualization tools are used extensively in our classrooms, government reports, newspapers and popular magazines, television news, and sports coverage. Over the last two decades, many websites, books, and other publications have emerged focused on data visualization. One remarkable step in this regard is that of Gap minder - the author and inventor Hans Rosling found out attractive means of communicating several demographic and economic global phenomena to a wide audience using data visualization (Rahlf, 2020).

Data visualization in sports has gained utmost popularity over the years, especially in sports reports appearing in relevant periodicals, on websites, and in television broadcasts. Some recent and excellent examples of data visualization related to cricket and football can be seen in the works of (Perin et al., 2018; Tselova, 2022). provide an extensive review of visualization in the arena of sports. This paper attempts to develop a visualization tool that can be utilized to envisage the outcome of all the matches of a tournament figuring multiple teams, in a single diagram. This shall help the onlooker to understand the run of victories and losses of the different

participating teams when matches in the tournament are played as a league or on a league-cum-knock out basis. For display, real data is taken from the recently concluded ICC Twenty20 Cricket World Cup Tournament of 2022, hosted by Australia.

2. Review of Literature

With access to technological advances such as color, computer graphics, computer animation, dynamic graphs, three-dimensional computer graphics, holography, etc., many statistical problems are solved through the innovative use of both graphical and semi-graphical tools. Most recent versions of statistical software now produce high-resolution, self-explanatory graphics (Bhattacharjee & Das, 2007). Semi-graphical displays are often used in newspapers and reports, and in sports, they are generally used to formulate strategy, mostly in soccer, rugby, basketball, badminton, etc. (Haq et al., 2022; Tani et al., 2015; Zuo, 2022). Visualization has become easier these days, and some websites can generate various types of static and dynamic graphics from user-defined datasets.

Some graphs on websites are designed so that a particular section of the graph is highlighted and displays the corresponding figure as tooltip text when the user hovers over that section. One such website is flourish-studio, which can be used to produce various types of sports-related graphical displays. The article by (Tselova, 2022). discusses how to use Flourish Studio to display sports-related data, especially soccer data. In this work, a visualization tool for match schedules and outcomes of the FIFA World Cup 2022 is mentioned. But the author has shown only matches from the round of 16 onward. Displaying the entire tournament schedule and all match outcomes is a daunting task. A proper display of all the matches in a tournament shall show at a glance how each team progressed through the rounds. As the best of our search did not provide us with any such convincing display, it worked as motivation to generate a semi-graphical representation of this type. This is the area where the proposed graph excels over its existing competitors.

3. Objective of the Study

It is evident from the literature review that there is a need for, and an inadequacy of, a visualization tool to display the schedule and outcomes of all the matches in a tournament. This sets the tone for spelling out the study's objective. The paper sought to develop a data visualization tool that displays the outcomes of all matches in a tournament involving multiple teams.

4. The Data Set

The dataset consists of match outcomes from the Twenty20 Cricket World Cup, played in October-November 2022 across different cities in Australia. After the qualifying round, 12 teams were named for the main phase of the tournament. The teams were divided into two groups, viz. Group 1: New Zealand (NZ), Australia (Aus), England (Eng), Sri Lanka (SL), Ireland (Ire), Afghanistan (Afg), and Group 2: India (Ind), Pakistan (Pak), Bangladesh (Ban), Netherlands (Ned), South Africa (SA), Zimbabwe (Zim). Each team in the group plays every other team once, and the top two teams from each group advance to the semi-finals. The teams earn 2 points for a win and 1 point for a tie or a no-result. In case of a loss, no point is awarded. In the first semi-final, the top team from Group 1 plays the second-place team from Group 2. Likewise, in the second semi-final, the top team from Group 2 plays the second-place team from Group 1. The winners of the two semi-finals played each other in the final. The schedule and results of each match are collected from espncricinfo.com.

5. The Plot

The plot, which can be called the Tournament Outcome Display, is semi-graphical and presented as a table, with color used to convey different information. The display can communicate the following information: the schedule of matches, the winners and losers of all matches in the tournament, the teams in the different Groups, the runs of victory of all teams, matches segregated into league matches and knock-out matches. The Tournament Outcome Display for the ICC Twenty20 Cricket World Cup 2022 is shown in Figure 1.

To understand the proposed semi-graphical display, it is essential to have some familiarity with the tournament it visualizes, i.e., the Twenty20 Cricket World Cup of 2022. The tournament was played in Australia from 22nd October to 13th November 2022, with 12 national sides competing in the final round. The teams were divided into two groups (Group 1 and Group 2), each with six teams. Within the groups, round-robin matches were played, and then the knock-out round (semi-finals) began with the two table-toppers from each group. In Group
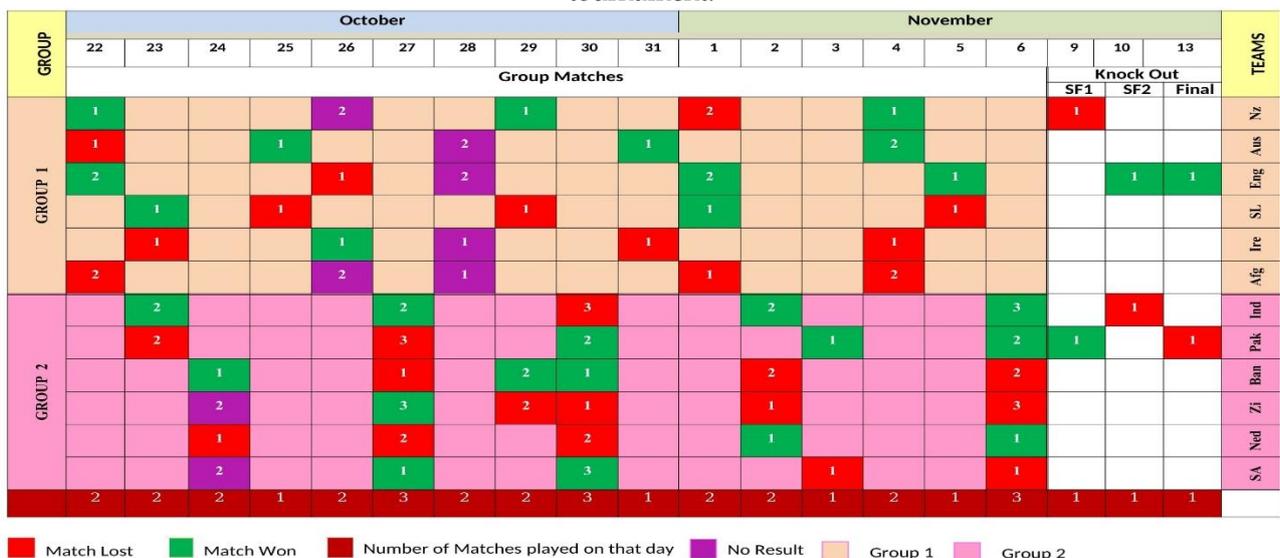
1, New Zealand and England occupied the top positions, and in Group 2, India and Pakistan occupied the top two slots. In the knockout stage, under the crisscross rule, India faced England, and New Zealand faced Pakistan in the semi-finals. Following the semi-victorious teams, Pakistan and England reached the finals of the World Cup, which was played on the 13th of November 2022 at the Melbourne Cricket Ground, Melbourne. Ultimately, England defeated Pakistan in the final by 5 wickets and won the tournament.

In Figure 1, the proposed semi-graphical display is portrayed. It attempts to visualize all the matches of the tournament, viz., the schedule: the participating teams in all matches and their outcomes. The plot is a semi-graphical display in table form, resembling a calendar, showing the winning team in a green cell and the losing team in a red cell. In case there is 'no result' in the match, then the violet color is used in those cells. The first column identifies the tournament's different groups. As there are only two groups (1 and 2), there are only two cells in the first column, each distinguishing the two groups.

The first row specifies the months, and the second row displays the dates. These two rows, when combined, spell out the dates of the different matches in the tournament. The last column of the display is used to name the teams. Additional colors can be used to communicate different stages of the tournament, such as group league or knockout matches. Sometimes, in a single day, multiple matches are played, and they can be differentiated by the integers, viz. -1, 2, etc., used in the boxes to identify the teams that participated in the first match, second match, etc., on that day, respectively. The last row of the display is colored in maroon, and each cell of that row contains an integer, which is the number of matches played on that day. The display may indicate that the second column shows the matches played on 22nd October, with two matches.

Both matches were played by the teams categorized in Group 1. The first match was between New Zealand and Australia, with New Zealand remaining victorious, as indicated by a green cell (in column 2 against New Zealand). Likewise, the second match was played between England and Afghanistan. Marked by cells with integer '2' in the rows dedicated to England and Afghanistan (c.f. last column of the display for country names). In those cells of the second column marked with the integer '2', the cell dedicated to England's row is colored in green, and that of Afghanistan is marked in red, indicating that the latter lost the match to the former. A similar process can be used to interpret the outcome of the other matches of the tournament from the display. Looking at the plot, some interesting observations can be made, such as that all the matches played on 28th October were abandoned. Generally, if more than one match is played on a particular day, then all the participating teams in those matches either belong to Group 1 or Group 2. The exception is on the 23rd and 29th of October. On both these days, two matches were played. The four participating teams in those two matches were from different groups.

Figure 1: A proposed graphical technique for representing the schedule outcome of all the matches of a tournament.

GROUP	22	23	24	25	26	27	28	29	30	31	1	2	3	4	5	6	SF1 (9)	SF2 (10)	Final (13)	TEAMS
GROUP 1	1				2			1			2			1			1			Nz
	1		1				2		1			2								Aus
	2				1		2				2				1			1	1	Eng
		1		1				1			1				1					SL
		1			1		1		1					1						Ire
	2				2		1				1	2								Afg
GROUP 2		2				2			3			2				3		1		Ind
		2				3			2				1			2	1		1	Pak
			1			1		2	1			2				2				Ban
			2			3		2	1			1				3				Zi
			1			2			2			1				1				Ned
			2			1			3				1			1				SA
	2	2	2	1	2	3	2	2	3	1	2	2	1	2	1	3	1	1	1	

Legend: Match Lost / Match Won / Number of Matches played on that day / No Result / Group 1 / Group 2

The figure can also inform the onlooker that the Group matches were played between the 22nd of October and the 6th of November, and the Knock-out phase started on the 9th of November, and the Final was played on

the 13th of November. England won the finals and were named the champions. The team that won the most matches is England (identified by counting the green squares in the row corresponding to their team). The worst performance was Afghanistan's, which ended the tournament without a victory, with no green squares in the row assigned to them. As many as three matches were played on the 27th and 30th of October and on the 6th of November.

6. Limitation of the study, Direction of Future Research, and Conclusion

The paper introduced a novel plot, Tournament Outcome Display, that visually represents the fixtures and outcomes of several matches in a tournament. Although the display in Figure 1 uses a cricket tournament as an example, it can also be applied to tournaments in other team sports. The semi-graphical display is simple to implement. It can be drawn in any word processor that supports table formatting with color. However, for any graphical display, supporting software is helpful. This enables the display to be produced in a much more objective manner.

The current researchers have yet to develop any algorithms or code to produce the graphics discussed. This may be an area of future research concerning this display. Another extension of the display could include the margin of victory for each match. As the margin of victory widens, this can be achieved by intensifying the green hue used in the winner's box and vice versa. Similarly, the intensity of the red hue in the loser's box increases as the margin of defeat widens. Future researchers may also consider creating a web application or software that would enable widespread usage of the display in TV programs and news items presented in magazines and periodicals. Drawing the display without software may prove time-consuming and brainstorming-intensive.

Representing the schedule and outcome of all the matches of the FIFA World Cup using the proposed semi-graphical display can be a big challenge. The last FIFA World Cup, held in 2022 in Qatar, featured 32 national teams. The teams were divided into eight round-robin groups, and the knock-out round began with the two table-toppers from each group. A total of 64 matches were played in the tournament in 29 days (contributors, 2023). It has been learned that the next FIFA World Cup, to be held in 2026, will feature 48 national teams. Though the format of the upcoming FIFA World Cup with 48 nations has yet to be announced, it will be a difficult tournament.

Accordingly, visualizing the tournament schedule and match outcomes using the semi-graphical display discussed here might be challenging, but not impossible. With the idea once in the air, future researchers and data analysts may take the issue forward and apply the display to other tournaments with similar or even more challenging formats, such as the upcoming 2026 FIFA World Cup. With technological advances and as more decision-making studies become data-driven, the need for and skill in data visualization have grown severalfold in the last few decades. This paper is a humble attempt to introduce another application of graphics in the field of sports. The same technique can be carried forward to other domains of knowledge as well. This might be another interesting area for future researchers to explore.

Reference:

Bhattacharjee, D. (2007). *Some Aspects of Statistical Graphics* Gauhati University]. Guwahati, India.

Bhattacharjee, D. (2012). Moving Bubbles: An Example of Dynamic Visualization. In T. K. Dey & A. Nag (Eds.), *Contemporary Trends of Research in Physical Sciences* (pp. 111-116). Department of Physics, Guru Charan College.

Bhattacharjee, D., & Das, K. K. (2007). *Statistical Graphics for Everyone*. South Asian Publishers Pvt. Ltd.

contributors, W. (2023). *2022 FIFA World Cup*. https://en.wikipedia.org/wiki/2022_FIFA_World_Cup

Haq, M. A., Tarashima, S., & Tagawa, N. (2022). *Heatmap visualization and badminton player detection using a convolutional neural network* International Electronics Symposium (IES),

Perin, C., Vuillemot, R., Stolper, C. D., Stasko, J. T., Wood, J., & Carpendale, S. (2018). State of the Art of Sports Data Visualization. *Computer Graphics Forum*, *37*(3). https://doi.org/10.1111/cgf.13447

Rahlf, T. (2020). *Data Visualisation with R: 111 Examples*. Springer.

Tani, T., Huang, H., & Kawagoe, K. (2015). *Sports play visualization system for American football* Proceedings of the International Multiconference of Engineers and Computer Scientists,

Tselova, S. (2022). *Seven creative data visualization ideas for your World Cup coverage.* https://flourish.studio/blog/visualizing-world-cup-2022/

Zuo, X. (2022). Visualization of Football Tactics with Deep Learning Models. *Wireless Communications and Mobile Computing*, 2022.