Impact of Toss Decision on Winning Games

The dataset I have used is related to my final project, which is the analysis of cricket teams (ODI format) from 1996 to 2005. This dataset contains every ODI (One Day International) cricket game played between 1996 and 2005 with 1237 instances (rows) under the attributes like Team 1, Team 2, winner of toss, toss decision, respective scores, winner, the margin of victory, and location of the cricket ground (city, country).

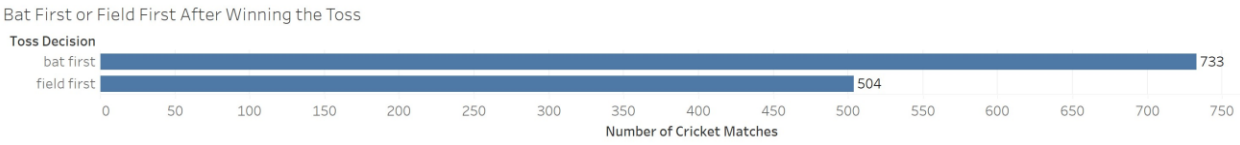
For this Exploratory Visual Data Analysis, I would like to explore external factors (apart from team players' strength) which could have an impact on teams winning more games. In the game of cricket, the cricket pitch (the surface on which bowlers bowl the cricket balls at batters) is a key factor that decides whether it will be easy for batting or more conducive for bowling. Thus, after a team wins the toss, they have the crucial advantage to choose whether to bat or field first depending on the cricket pitch on that particular day. I plan to analyze the relationship between toss decisions (bat or field first) and the number of games won by teams.

As a result, my pitch topic idea is "Impact of Toss Decision on Winning Games" and **my initial question** (not final) **is whether or not winning the toss had an effect on the number of games won by teams**. With respect to data required to visualize this idea, I have "Toss Decision", "Toss" (which team won the toss), "Winner", and "Count of Cricket Matches" attributes. I have created a calculated field "Bat First" which was needed for my final visualization to sort data in descending order. I will explain more on why I created this calculated field during my explanation of the final visualization. This was the step I took as part of my data preparation/cleaning.

Bat First

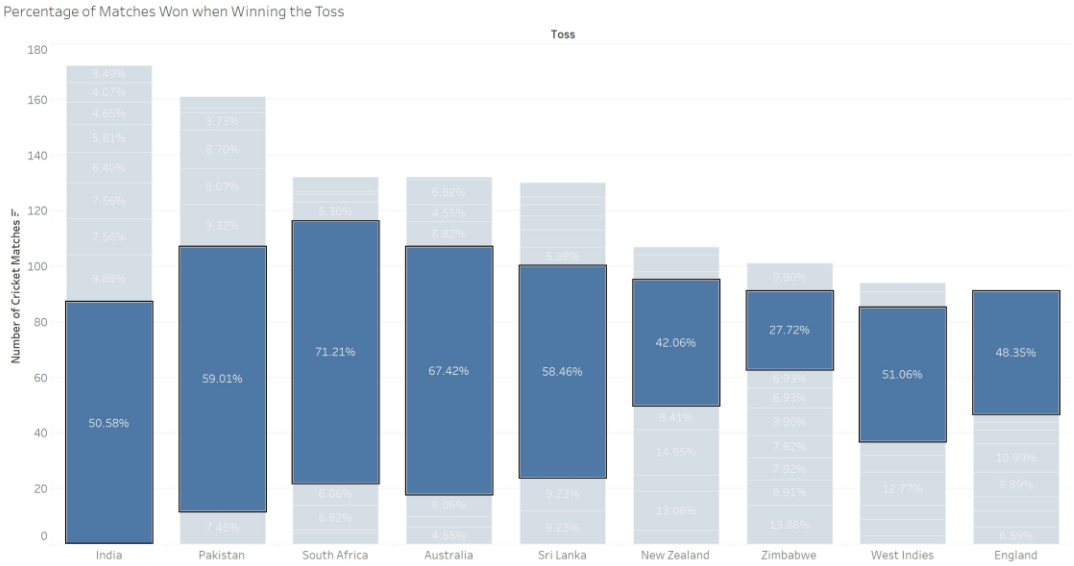
```
IF [Toss Decision]="bat first"
THEN 1
ELSE 0
END
```

To start my exploratory analysis, I created a bar chart with the "Toss Decision" dimension against the "Count of Cricket Matches" measure followed by placing the "Count of Cricket Matches" measure on the "Detail" property of the Marks card. I observed that teams winning the toss have elected to bat first on 733 occasions and field first on 504 occasions. Teams have opted to bat first on 59.26% of the occasions. This data visualization is attached below for reference.



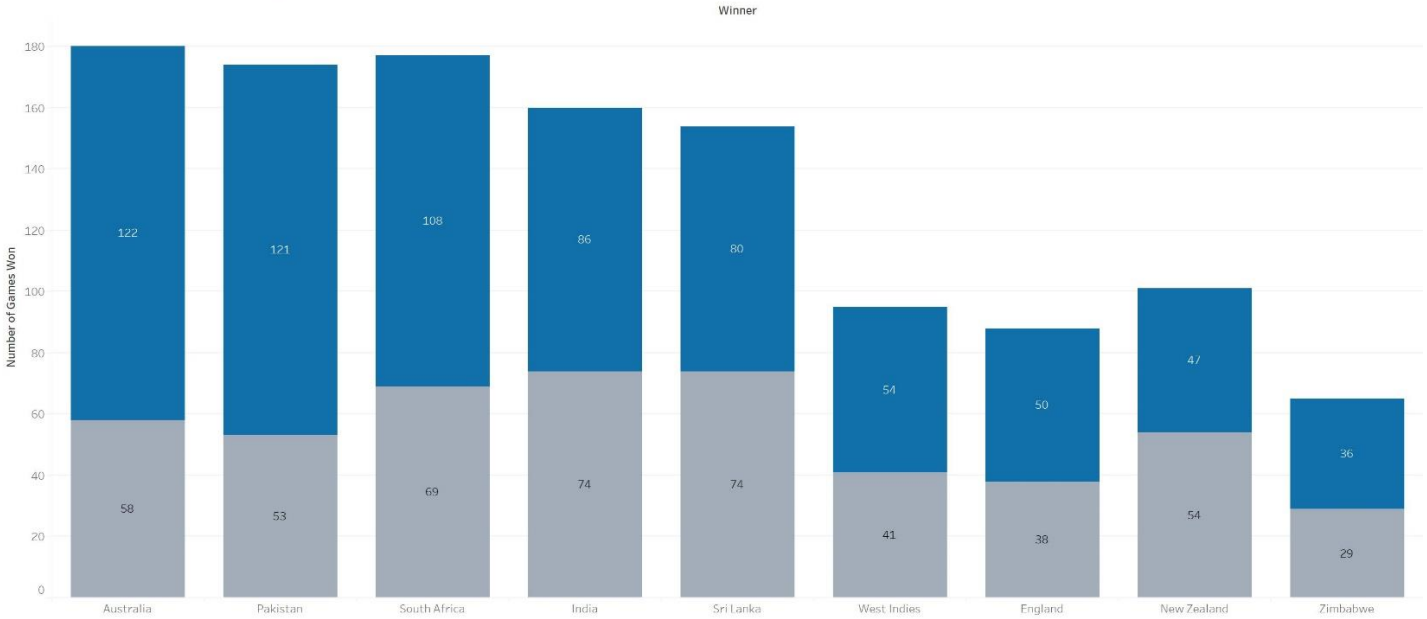
I wanted to gain more information on whether there was a relationship between teams winning the toss and winning the games. To visualize this data, I created a bar chart with the "Toss" dimension against the "Count of Cricket Matches" measure followed by placing the "Count of Cricket Matches" measure on the "Label" property of the Marks card. The "Winner" dimension

was placed on the "Detail" property of the Marks card. 17 cricket teams have played ODI cricket from 1996 to 2005. Out of this, 8 teams are associate nations that do not have any ranking based on their performance and also have played a significantly lesser number of matches compared to the other 9 member nations. So I used a filter to exclude these association nations from this visualization. For this step, I placed "Toss" and "Winner" dimensions under the Filters shelf and excluded the associate nations. I observed that there were nations like South Africa and Australia which won approximately 70% of the matches when they won the toss. Although, there was Zimbabwe which won only 28% of the matches when it won the toss. The nations other than these 3 mentioned before had won between 42% and 59% of the matches when they won the toss. There was no conclusive evidence to say that winning the toss meant a higher chance of winning the game. This data visualization is attached below for reference.



I decided to explore whether batting first or fielding/bowling first after winning the toss had an impact on winning games. As a result, **my final question has evolved to whether or not electing to bat first or field first after winning the toss had an effect on the number of games won by teams**. For this data visualization, I created a bar chart with the "Winner" dimension against the "Count of Cricket Matches" measure. I placed the "Winner" dimension on the Filters shelf in order to remove the 8 associate nations. The "Count of Cricket Matches" measure was placed on the "Label" property and the "Toss Decision" dimension was placed on the "Color" property of the Marks card. The calculated field "Bat First" is used here to sort the number of games won when batting first in descending order. This data visualization led to the observation that winning teams that also won the toss elected to bat first during 60% (704 out of 1194 matches) of the entire matches. All cricket teams except for New Zealand had won more games when opting to bat first than to field first. Winning the toss and electing to bat first led to a significant advantage for cricket teams with a higher probability of winning from 1996 to 2005 and this visualization answers my final question by stating that teams had won more games by winning the toss and opting to bat first. Thus, the toss decision to bat first or field first has had an impact on teams winning more games.

Impact of Toss Decision, 'Batting First' and 'Fielding First', on the Number of Games Won by Teams



This graph represents the number of games won by cricket teams when choosing to bat first or field first after winning the toss.