Six Nations Rugby Prediction

April Ying-Chieh Chiu

Intro

The Six-Nations Rugby Tournament is an annual rugby competition between England, France, Italy, Wales, Ireland and Scotland that runs from February 3 to March 17. In this assignment, we want to **find insights and make predictions about who will win the Six Nations tournament** by diving into the historical data, including the scores of home teams and Away teams as well as the results of each matchup.

In terms of the analysis tool used to approach this project, there're various kinds of ways and tools such as using R or Microsoft Excel to arrange and explore those data. However, I chose **Python** to conduct data exploration and analysis in this project because I believe that it's a precious opportunity to step out of comfort zone, improving my Python skills and learning more about this important analysis tool.

In this project, I plot out the plus and minus points of the home team in each matchup and calculate the overall winning percentage of each team in order to find out which team are more likely to win at home or win away, which can represent the probability of winning the matchup in the future. If the probability is around 50%, representing that the two teams in this matchup are both great, I look up some bets and odds information as well as some rugby fans' posts and discussions for reference. For instance, the round #1 we're predicting focuses on three matches, France v. Wales, Scotland v. Italy and Ireland v. England, so I calculate the winning percentages of these teams in the past five years historical data in order to find out the teams which are most likely to win the matchup.

By analyzing and exploring those historical data, we can make precise predictions as well as draw reliable conclusions.

Analysis

First of all, in order to find out which teams are more likely to win the matchups, I mainly use **HomeTeam, AwayTeam and Score** variables to calculate the winning percentage. After some simple data preparation, we can calculate the winning percentage for each team as they were playing at home or away. My prediction for each matchup and the following actual results are as following. The complete analysis and calculating process is attached in appendix.

Matchup	Prediction	Actual Result
Round #1		
Wales v. France	Wales	Wales
England v. Ireland	Ireland	England
Scotland v. Italy	Scotland	Scotland
Round #2		
England v. France	England	England
Wales v. Italy	Wales	Wales
Ireland v. Scotland	Ireland	Ireland
Round #3		
France v. Scotland	Scotland	France
Italy v. Ireland	Ireland	Ireland
Wales v. England	England	Wales
Round #4		
Wales v. Scotland	Wales	Wales
England v. Italy	England	England
Ireland v. France	Ireland	Ireland
Round #5		
Italy v. France	France	
Wales v. Ireland	Wales	
England v. Scotland	England	

Conclusion

We can always find insights and make predictions according to historical data. However, as regarding the overall winning percentage of each team as the probability of winning the matchup, we need to plot out the data as well as collect some additional information, too.

Although everyone can make predictions, making precise predictions has always been a difficult thing which we try to improve and make every effort to achieve. Thanks for this interesting opportunity, I can improve the technical skills related to Python, learn how to dive deeper into the sports area and conduct data analysis to those data as well as learn how to make more precise predictions!