Female Rugby Statistics

Joshua C., Jacob D., Helena GB., Eric W., Brandon H.

2022-11-10

Title Page

Female Rugby Statistics

Course 331 01

Group Memebers:

- Joshua

- Jacob

- Helena

- Eric

- Brandon

- The type of analysis we are doing is withing our area of interest.

Introduction

- This set, as its name shows, is based on the fifteens version of rugby, which uses fifteen players on the field at once, with eight substitutes. These rugby games are played by professional rugby players at the world level, with players from different nations. The motivation for the analysis is to determine who the best team in the Female Rugby League is. Our 3 main questions of interest that we are searching for in our data are: 1) Which team has the largest combined margin of victory? (Net) 2) Which venue results in the most points scored? The least? 3) How does being home v.s. away affect win probability?

- The data source is TidyTuesday which can be found at this link: https://github.com/rfordatascience/tidytuesday/tree/master/data/2022/2022-05-24

- The general nature of the data split between strings and integers

- The time period of this data is from June 13th, 1982 to May 6th, 2022

Initial Hypothesis

The variables that we will identify to answer our 3 main questions are:

- Question 1: Which team has the largest combined margin of victory? (Team with the most points scored in a season compared to the team with the least points scored) This will use the data from "score_1" and "score_2" which are related.

- Question 2: Which venue results in the most points scored? The least? This will use the data from "venue", "score_1", and "score_2".

- Question 3: How does being home v.s. away affect win probability? This will use the data from "home_away_win" and "winner".
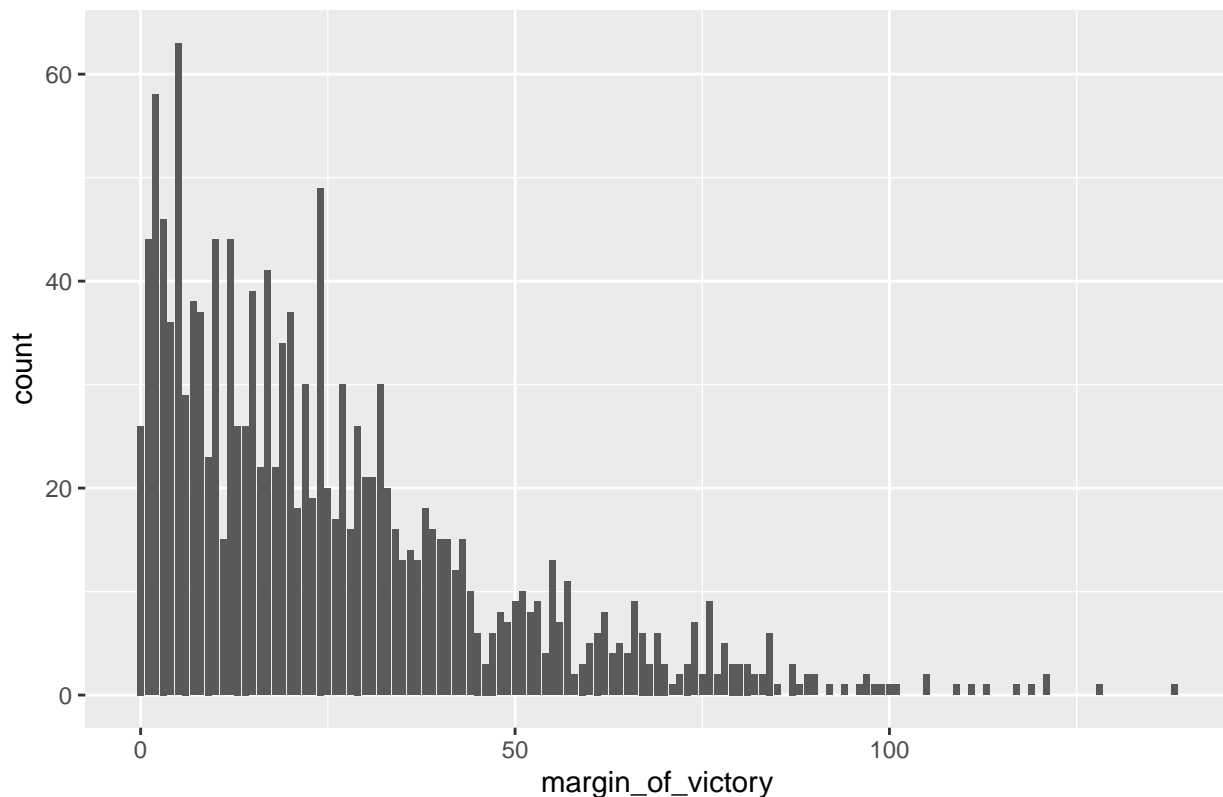
Data Preparation

- Before tidying the data, there are no missing values which makes our job easy. We shouldn't need to overcome any challenge when it comes to tidying the data.

Exploratory Data Analysis (EDA) and Statistical Analysis

- Question 1: Which team has the largest combined margin of victory? (Team with the most points scored in a season compared to the team with the least points scored)

Margin of Victory between 1982–2022 for Womens Rugby

- The team that scored the most points scored 65 and multiple teams scored 0 points during games.