



Goal Timing in Men's and Women's World Cups

A Statistical Analysis of Men’s and Women’s FIFA World Cup Tournaments

Authored by join-by-slay-factor
STAT 199 Lab, Duke University

Research Question and Motivation

*"How does **goal timing** differ between **women's and men's** World Cup Tournaments?"*

- FIFA World Cup Women's games get **less attention.** Most of the soccer statistics we came across focused on men's football.
- **Side by side comparison:** The FIFA World Cup lets us look at men's and women's soccer in the same competitive format over a similar timespan.
- Timing differences could **help both games!** If goals tend to come at different points in men's and women's matches, **teams and analysts** could learn from each other's approach to **substitutions and late-game tactics.**

Introducing Our Data - Overview and Cleaning

The Dataset

- Fjelstul World Cup Database - goals.csv
- Authored by **Josh Fjelstul, PhD**
- **3,637** observations (rows) and **27** variables (columns)
- **Pro:** detailed goal information longitudinally
- **Con:** compiled using external sources → may be slightly inaccurate

What Variables Did We Include?

- tournament_name
  - year (derived)
  - gender (derived)
- minute_regulation
- minute_stoppage
- match_period
- own_goal
- penalty

What Changes Did We Make?

- Retained relevant variables
- Tournament_name: separate year and gender variables
- Year → as.numeric
- Penalty and own_goal → as.logical
- Limited observations to 1990-2019
- Ultimately excluded own_goal
- Overall 7 variables (columns) and 2,079 observations (rows)



GitHub

Josh Fjelstul, PhD

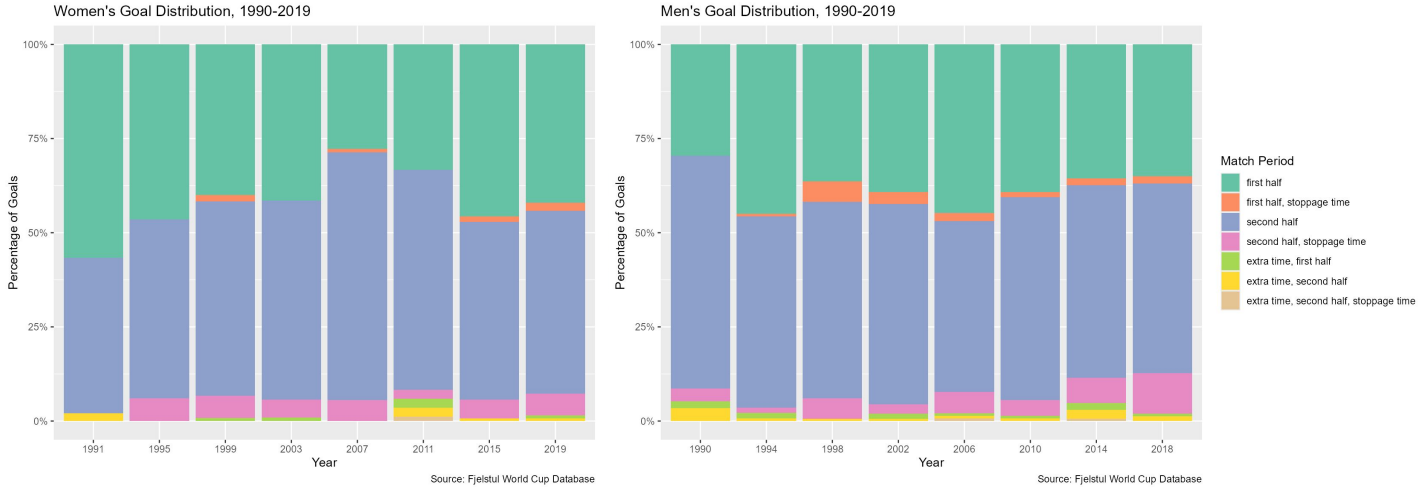
README

The Fjelstul World Cup Database

The Fjelstul World Cup Database is a comprehensive database about the FIFA World Cup created by Joshua C. Fjelstul, Ph.D. that covers all 22 men's tournaments (1930-2022) and all 9 women's tournaments (1991-2019). The database includes 27 datasets (over 1.5 million data points) that cover all aspects of the World Cup.

The database has been featured by The BBC, Sky Sports, The Washington Post, The Times, FiveThirtyEight, The Markup, Data is Plural, Agence France-Presse (AFP), Barron's, Latinometrics, The National, Hindustan Times, DataCamp, and Neo4J.

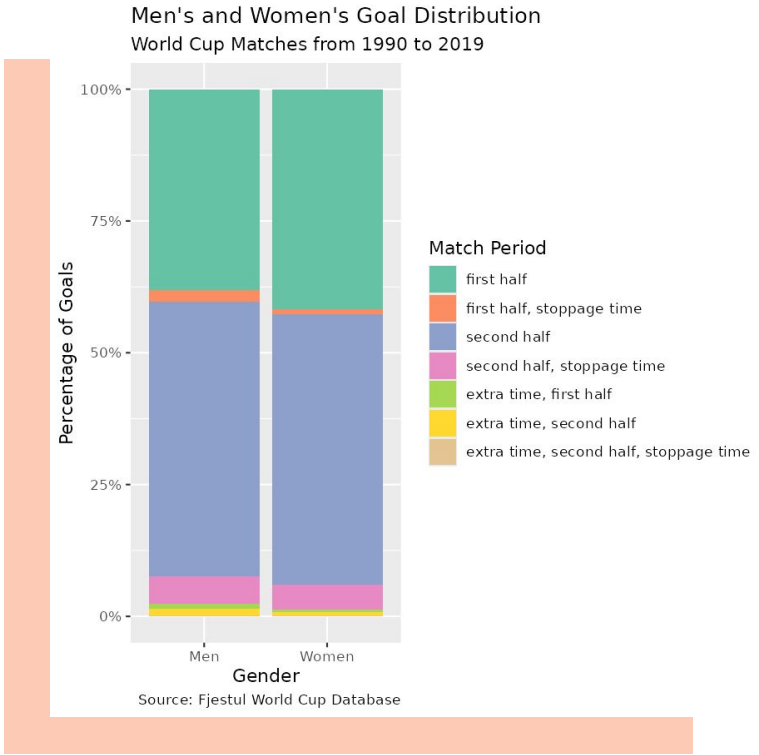
Visualization 1 - Goals Across Years



Key Observations

- **1990–91:** most women's goals came in the first half (57%), while most men's came in the second (65%)
- **2006–07:** most women's goals came in the second half (71%), while men's were split almost evenly (47% first, 51% second)
- **Overall:** most goals came in regular first- or second-half play, not stoppage or extra time (about 90% for men, 93% for women)

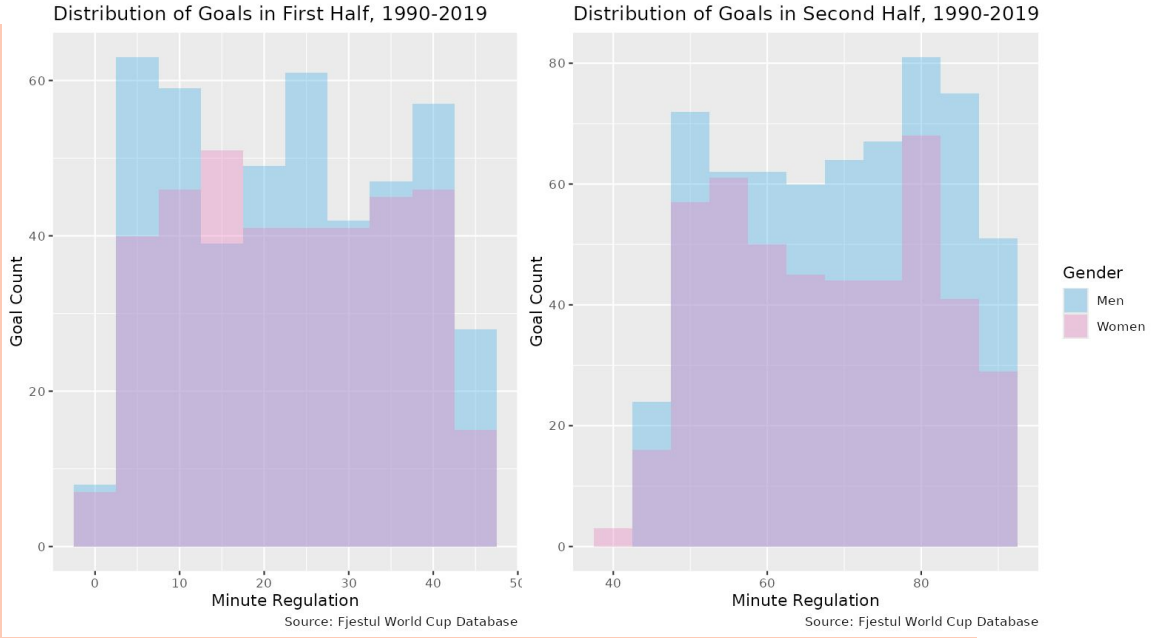
Visualization 2 - Goals Across Gender



Key Observations

- Both women and men score predominantly in the second half.
- Men score more in stoppage and extra time
- A greater proportion of women's goals are scored in the first half compared to men's goals.

Visualization 3 - Goals Across Minutes



Key Observations

- Men have visibly higher counts in several first half bins and in the late second half.
- Goals are less likely to be scored in the first and last 5 minutes of both halves.
- Interestingly, there is a peak in scoring at around minute 80 of regulation time for both women and men's soccer.

Conclusions and Limitations

- Both women and men score predominantly in the second half.
- Goals are less likely to be scored in the first and last 5 minutes of both regulation halves.
- Visually, there seems to be minimal differences in men and women's goal timing in the World Cup – although men's total goals are higher in almost every category.
- **Limitations**
    - Difference in regulation times in 1991 Women's World Cup
    - Women's World Cup has fewer matches and goals





Thank You!

Questions?

Authored by join-by-slay-factor
STAT 199 Lab, Duke University