

Does Hosting the Olympics Provide Home Advantage in Gymnastics?

Project

Han Nguyen

2025-11-12

Contents

1	Introduction	3
1.1	Background	3
1.2	Why Gymnastics?	3
1.3	Research Questions	3
2	Data Overview	4
2.1	Datasets Used	4
2.2	Data Preparation	4
3	Question 1: Does Hosting Provide Home Advantage?	5
3.1	Overview	5
3.2	Sub-question 1: Do Host Nations Win Significantly More Medals?	5
3.3	Sub-question 2: Average Percentage Increase for Host Countries	6
3.4	Question 1 Summary	6
4	Question 2: Gender Differences in Home Advantage	8
4.1	Overview	8
4.2	Sub-question 1: Do Female Gymnasts Benefit More from Hosting?	8
4.3	Sub-question 2: Has the Gender Gap Changed Over Time?	8
4.4	Question 2 Summary	9
5	Question 3: Country Factors	11
5.1	Overview	11
5.2	Sub-question 1: Do Smaller Countries Experience Larger Increases?	11
5.3	Sub-question 2: Do First-Time Hosts Gain a Bigger Boost?	11
5.4	Sub-question 3: Do Host Nations Field More Gymnasts?	12
5.5	Question 3 Summary	12
6	Question 4: Economics & Demographics	13
6.1	Overview	13
6.2	Sub-question 1: Does GDP Per Capita Affect Home Advantage?	13
6.3	Sub-question 2: Do Countries with Larger Working-Age Populations Benefit More?	13
6.4	Question 4 Summary	14
7	Discussion & Conclusions	15
7.1	Summary of Findings	15
7.2	Key Insights	15
7.3	Implications	16
7.4	Limitations	16

7.5	Future Research Directions	16
7.6	Conclusion	17
8	References	18
9	Appendix	19
9.1	R Session Information	19

1 Introduction

1.1 Background

This project investigates whether hosting the Olympics provides a measurable home advantage in **gymnastics**. Unlike track and field events measured by time or distance, gymnastics performances are evaluated by panels of judges, creating opportunities for potential bias that may favor athletes from the host nation.

1.2 Why Gymnastics?

Gymnastics is ideal for studying bias because outcomes depend on **human judgment** rather than purely objective performance metrics. By analyzing Olympic gymnastics data from 1896 to 2020, we can determine whether hosting correlates with statistically significant performance improvements.

1.3 Research Questions

Question 1: Does hosting the Olympics provide a measurable competitive advantage in gymnastics?

- Do host nations win significantly more gymnastics medals during their hosting year compared to historical baselines?
- What is the average percentage increase in total medals and gold medals for host countries?

Question 2: Does the home advantage in gymnastics differ by gender?

- Do female gymnasts benefit more from hosting than male gymnasts?
- Has the gender gap in home advantage changed over time?

Question 3: Does hosting experience and country size influence the magnitude of home advantage?

- Do first-time host nations gain a bigger boost than repeat hosts?
- Do smaller countries (by area or population) experience larger percentage increases in medal counts?
- Does the host nation field more participating gymnasts, potentially influencing outcomes?

Question 4: How do economic and demographic factors affect home advantage?

- Among host countries, do those with higher GDP per capita see different magnitudes of medal increases?
- Do countries with younger median ages or larger working-age populations gain more advantage when hosting?

2 Data Overview

2.1 Datasets Used

1. **Olympics Athletes Dataset (1896-2020)**: 237,674 Olympic performances, filtered to ~28,500 gymnastics records
2. **Olympic Hosts Dataset**: Historical record of host cities and countries
3. **Population Data**: Demographic data from Our World in Data
4. **Country Area Data**: Geographic size information
5. **GDP Data**: Economic data from the World Bank

2.2 Data Preparation

- Standardized country names using NOC codes
- Created medal outcome variables (binary indicators)
- Calculated baseline performance for each country based on non-hosting years
- Merged economic and demographic data matched to Olympic years

Dataset Summary:

- Total gymnastics performances: 28,554
- Number of countries: 102
- Olympic Games analyzed: 30
- Host nation performances: 2,250
- Non-host performances: 26,304
- Time period: 1896-2020

3 Question 1: Does Hosting Provide Home Advantage?

3.1 Overview

We address two key questions:

1. Do host nations win significantly more gymnastics medals during their hosting year compared to historical baselines?
2. What is the average percentage increase in total medals and gold medals for host countries?

3.2 Sub-question 1: Do Host Nations Win Significantly More Medals?

3.2.1 Descriptive Statistics

Medal Winning Rates:

- **Host countries:** 16.04 % (361 medals from 2250 performances)
- **Non-host countries:** 8.03 % (2112 medals from 26,304 performances)
- **Difference:** 8.02 percentage points

Medal Winning Rates by Hosting Status

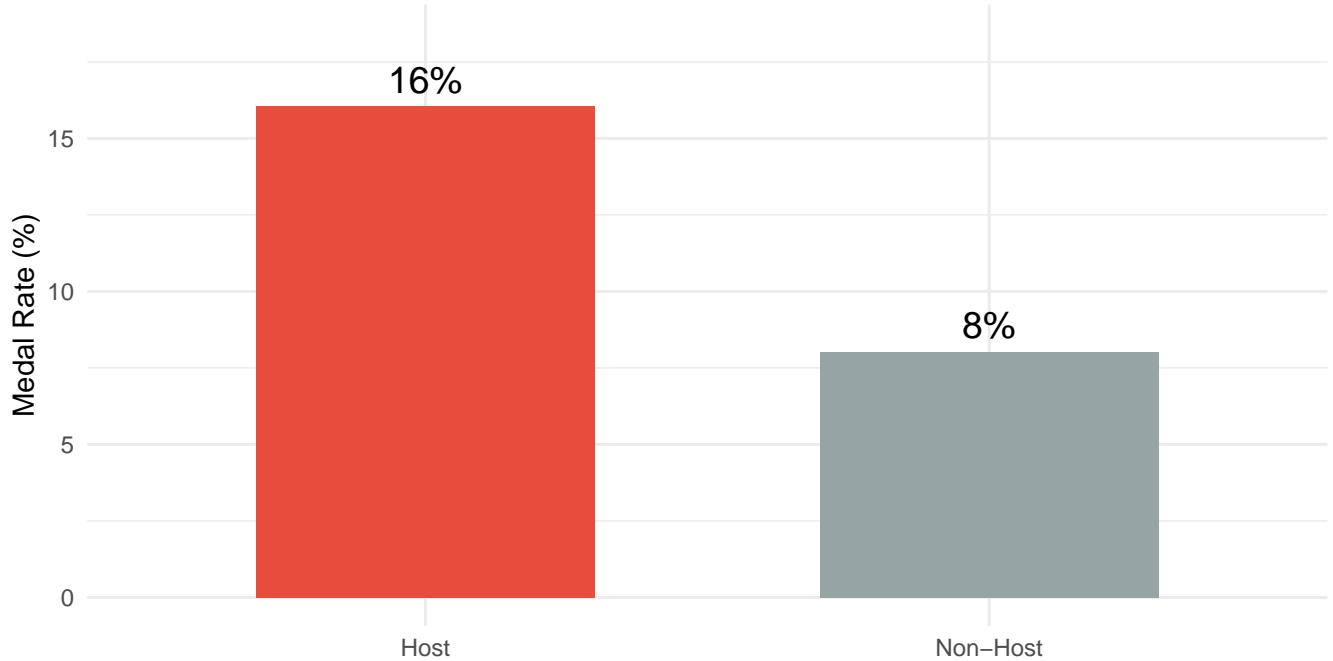


Figure 1: Medal Rates: Host vs Non-Host Countries

3.2.2 Statistical Test: Two-Sample Proportions Test

We test whether the difference in medal rates is statistically significant using a two-proportion z-test.

Hypotheses:

- H₀: Host and non-host countries have the same medal-winning rate
- H_a: Host and non-host countries have different medal-winning rates

Test Results:

- Test statistic (X-squared): 167.316
- P-value: <2e-16
- 95% Confidence Interval for difference: 6.44 % to 9.59 %

Conclusion: The difference is statistically significant ($p < 0.05$). Host nations have a significantly higher medal-winning rate.

3.3 Sub-question 2: Average Percentage Increase for Host Countries

We compare each host country's performance during their hosting year to their historical baseline (average in non-hosting years) to calculate the average percentage increase.

Performance Comparison:

- Average baseline rate (non-hosting years): 7.27 %
- Average hosting year rate: 20.47 %
- Average improvement: 13.2 percentage points

Paired T-Test Results:

- T-statistic: 2.892
- Degrees of freedom: 28
- P-value: 0.00732
- 95% CI for mean difference: 3.85 % to 22.55 %

Conclusion: Host countries perform significantly better during hosting years compared to their own historical baseline ($p < 0.05$).

3.4 Question 1 Summary

Key Findings:

1. Host countries have a medal rate 8.02 percentage points higher than non-host countries
2. Two-proportion test: Significant difference ($p < 0.05$)
3. Paired t-test: Significant improvement from baseline ($p < 0.05$)

Answer: YES, there is strong statistical evidence of home advantage in Olympic gymnastics.

Medal Rates: Hosting Year vs Historical Baseline

Top 12 Host Countries by Improvement

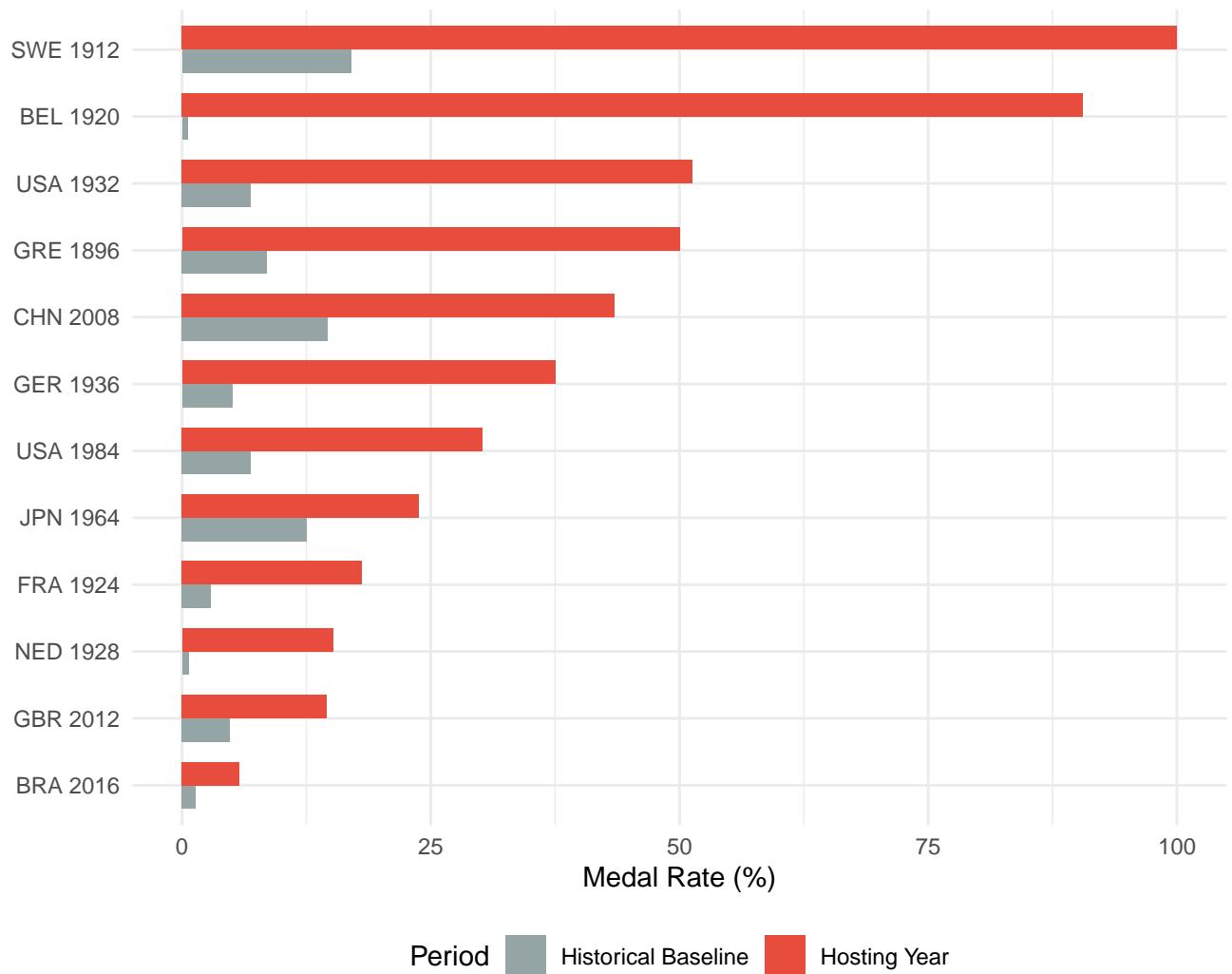


Figure 2: Hosting Year vs Historical Baseline Performance

4 Question 2: Gender Differences in Home Advantage

4.1 Overview

We address two key questions about gender differences:

1. Do female gymnasts benefit more from hosting than male gymnasts?
2. Has the gender gap in home advantage changed over time?

4.2 Sub-question 1: Do Female Gymnasts Benefit More from Hosting?

Medal Rates by Gender:

Male Gymnasts: - Host: 17.17 % - Non-host: 7.95 % - Difference: 9.22 pp

Female Gymnasts: - Host: 13.3 % - Non-host: 8.16 % - Difference: 5.14 pp

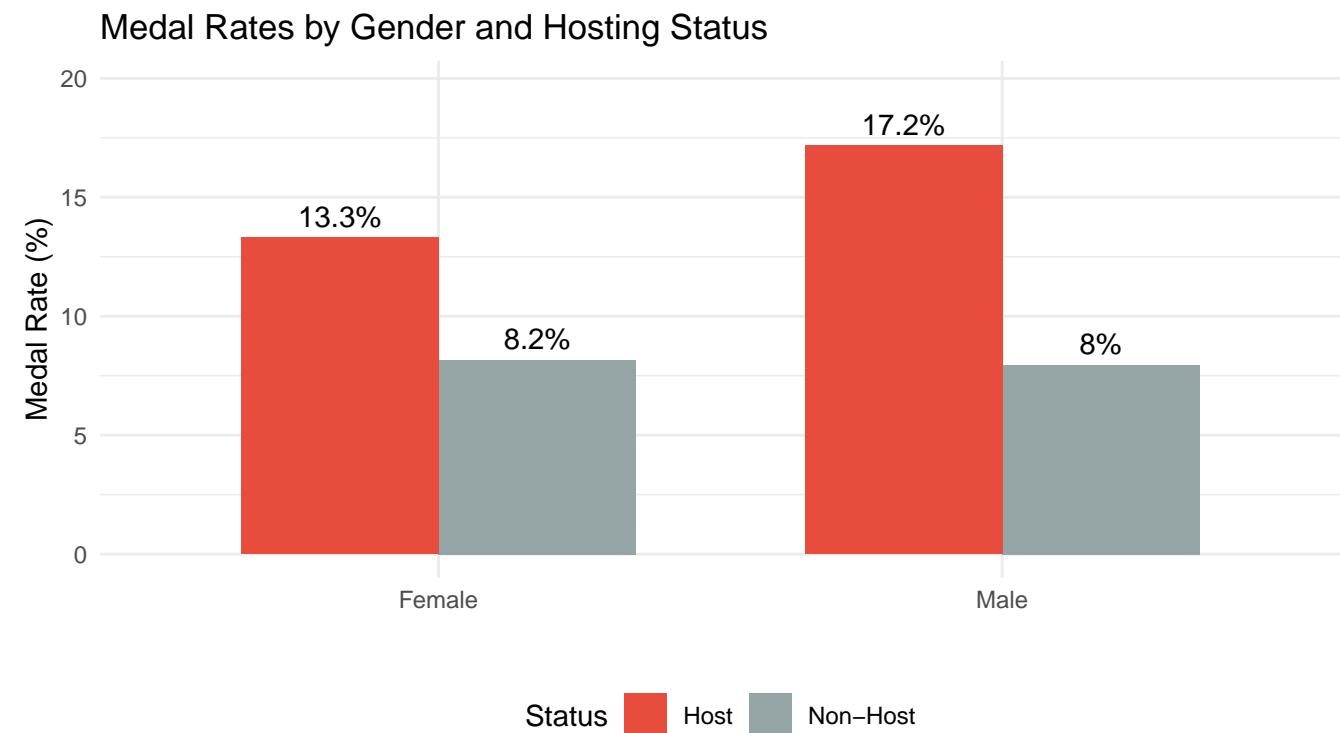


Figure 3: Medal Rates by Gender and Hosting Status

4.2.1 Statistical Tests

T-Test Results:

Males: - T-statistic: -9.529 - P-value: <2e-16 - Result: Significant

Females: - T-statistic: -3.785 - P-value: 0.000167 - Result: Significant

4.3 Sub-question 2: Has the Gender Gap Changed Over Time?

We examine whether the home advantage effect for males vs females has changed across different Olympic eras.

Home Advantage by Gender Across Eras:

Early (1896-1948) - Female: 34.07 pp - Male: 1.18 pp

Mid (1952-1988) - Female: 3.86 pp - Male: 11.3 pp

Modern (1992-2020) - Female: 2.76 pp - Male: 8.88 pp

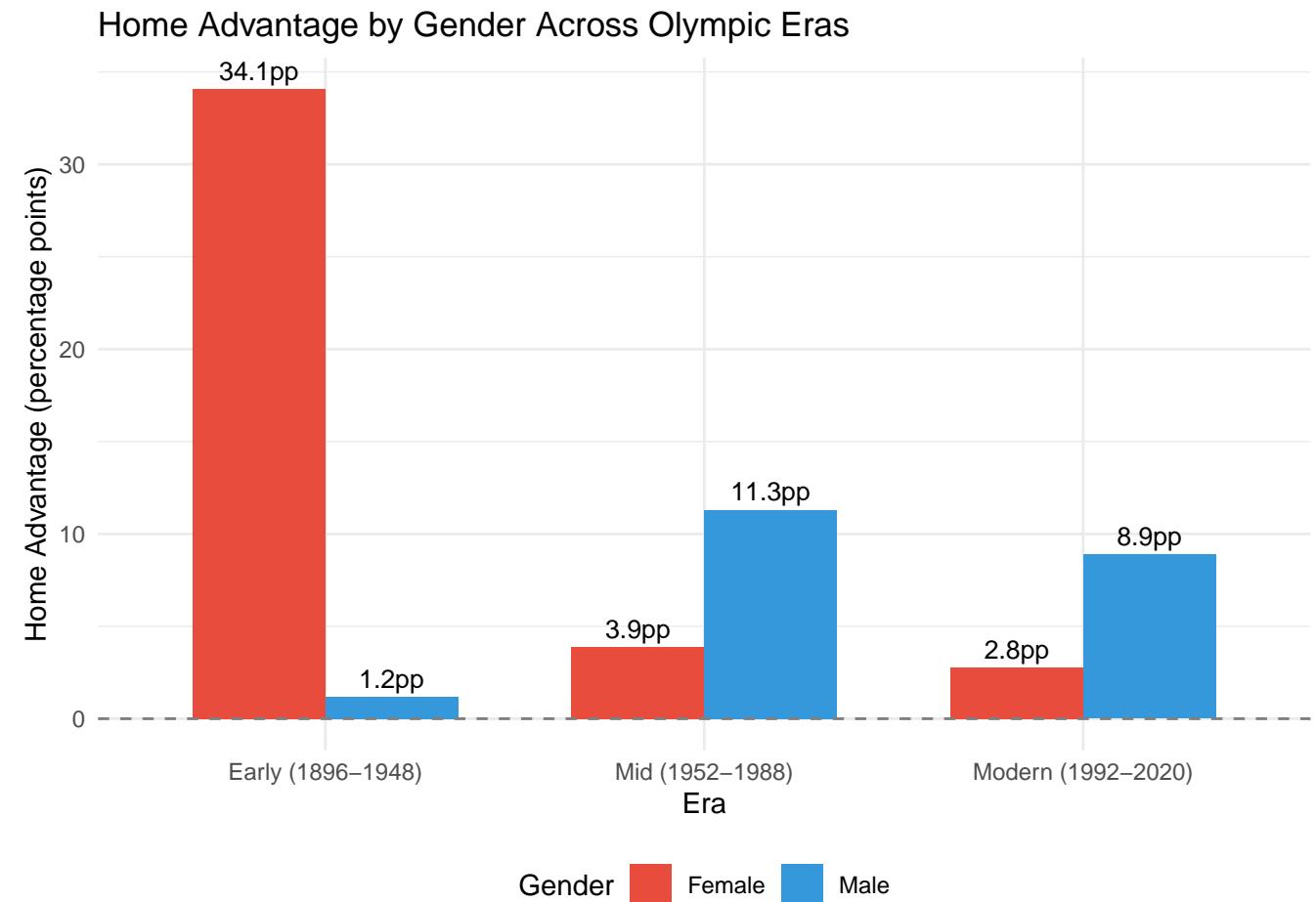


Figure 4: Gender Home Advantage Trends Over Time

4.4 Question 2 Summary

Key Findings:

1. Overall home advantage by gender:
 - Male home advantage: 9.22 percentage points
 - Female home advantage: 5.14 percentage points
 - Males show a larger overall home advantage effect
2. Temporal trends - Era-by-era analysis:

Early (1896-1948):

- Male advantage: 1.18 pp
- Female advantage: 34.07 pp
- Females show 32.89 pp larger advantage in this era

Mid (1952-1988):

- Male advantage: 11.3 pp
- Female advantage: 3.86 pp
- Males show 7.44 pp larger advantage in this era

Modern (1992-2020):

- Male advantage: 8.88 pp
- Female advantage: 2.76 pp
- Males show 6.13 pp larger advantage in this era

3. **Interpretation:** Home advantage effects vary significantly across time periods, with different patterns emerging in early vs modern Olympics.

5 Question 3: Country Factors

5.1 Overview

We address three key questions about country characteristics:

1. Do first-time host nations gain a bigger boost than repeat hosts?
2. Do smaller countries (by area or population) experience larger percentage increases in medal counts?
3. Does the host nation field more participating gymnasts, potentially influencing outcomes?

5.2 Sub-question 1: Do Smaller Countries Experience Larger Increases?

Average Improvement by Country Size:

- Large : 12.06 pp (n = 9)
- Medium : 3.66 pp (n = 4)
- Small : 17.14 pp (n = 15)

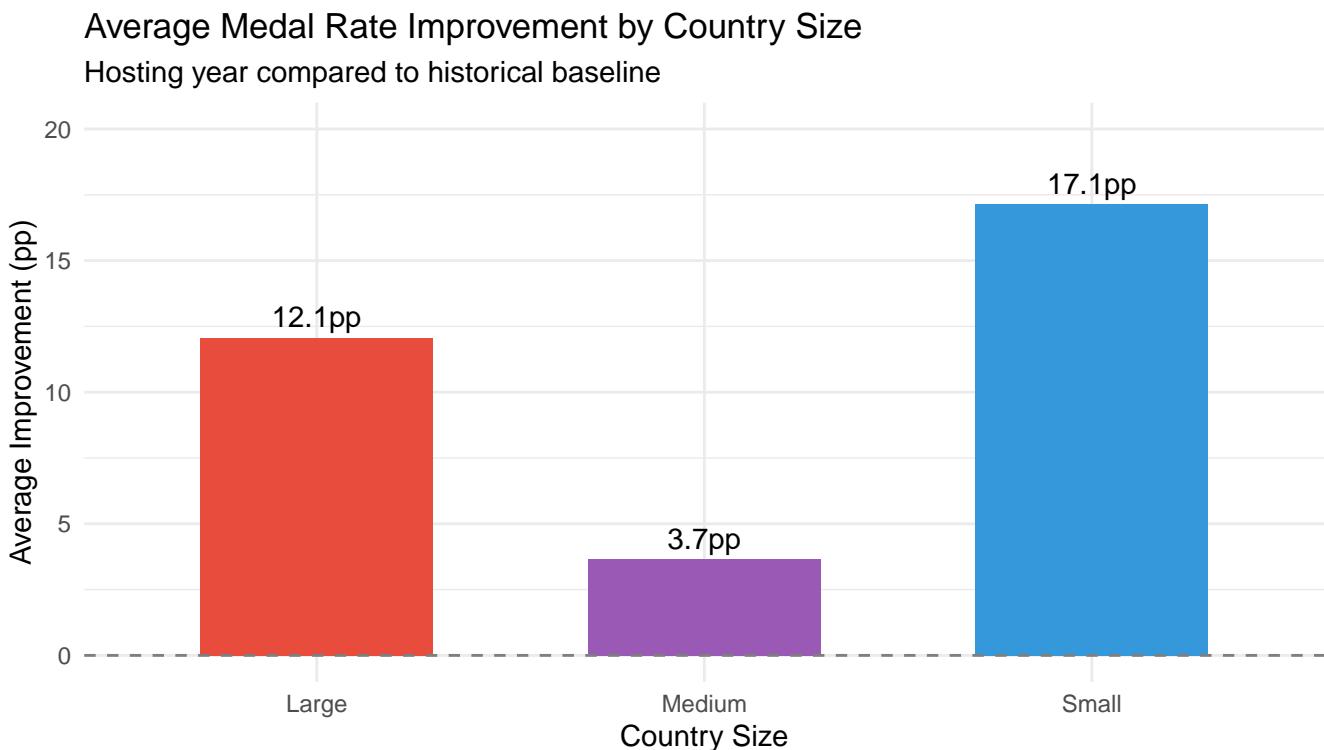


Figure 5: Home Advantage by Country Size

5.3 Sub-question 2: Do First-Time Hosts Gain a Bigger Boost?

Average Improvement by Hosting Experience:

- First-Time hosts: 16.53 pp (n = 18)
- Repeat hosts: 8.27 pp (n = 10)

T-Test: First-Time vs Repeat Hosts

- T-statistic: 0.981
- P-value: 0.336
- Result: No significant difference

5.4 Sub-question 3: Do Host Nations Field More Gymnasts?

We examine whether host nations field larger gymnastics teams, which could potentially influence their medal counts through increased opportunities.

Average Gymnast Count per Country:

- Host countries: 23.5 gymnasts
- Non-host countries: 7.2 gymnasts
- Difference: 16.3 gymnasts

T-Test: Host vs Non-Host Participation

- T-statistic: -3.328
- P-value: 0.00245
- Result: Host countries send SIGNIFICANTLY more gymnasts

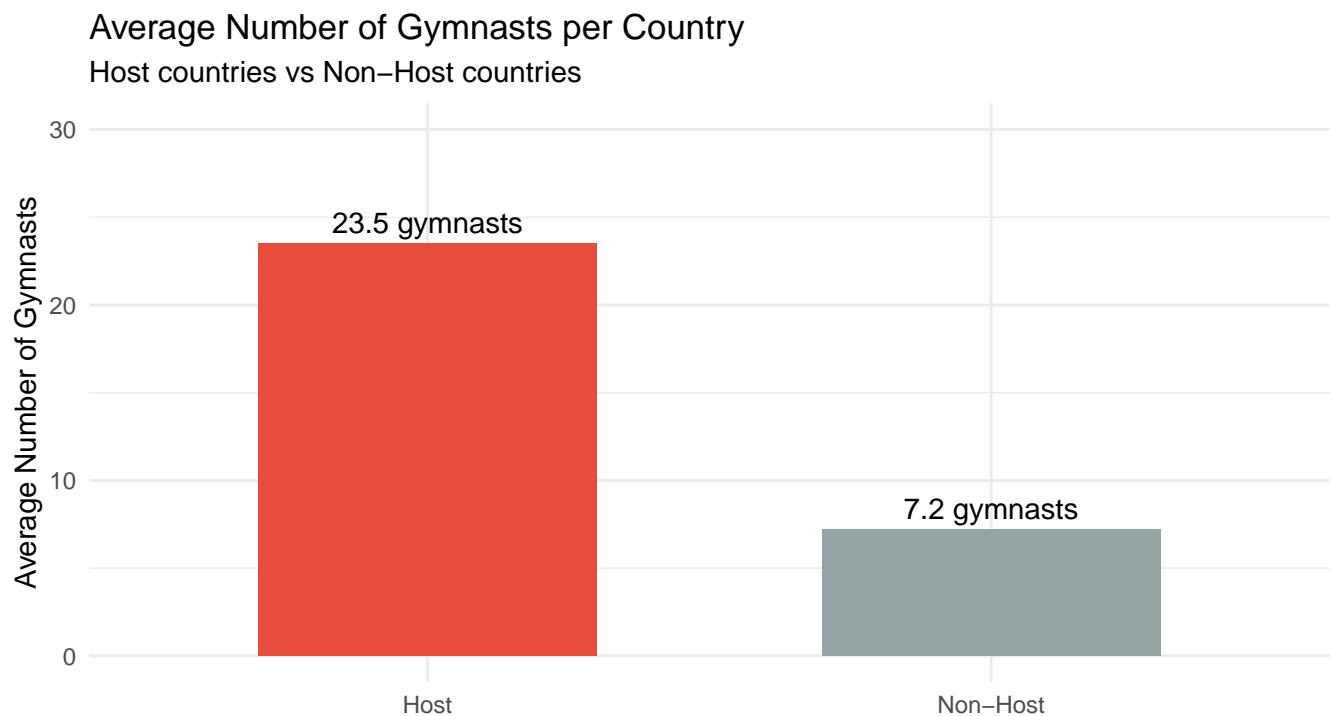


Figure 6: Gymnast Participation: Host vs Non-Host

5.5 Question 3 Summary

Key Findings:

1. **Country size:** Shows varying levels of home advantage across small/medium/large countries
2. **Hosting experience:** No significant difference between first-time and repeat hosts
3. **Gymnast participation:** Host countries send an average of 16.3 more gymnasts - statistically significant

6 Question 4: Economics & Demographics

6.1 Overview

We address two key questions about economic and demographic factors:

1. Among host countries, do those with higher GDP per capita see different magnitudes of medal increases?
2. Do countries with larger working-age populations gain more advantage when hosting?

6.2 Sub-question 1: Does GDP Per Capita Affect Home Advantage?

Correlation: GDP per Capita vs Home Advantage

- Correlation coefficient: -0.115
- P-value: 0.708
- Result: No significant correlation

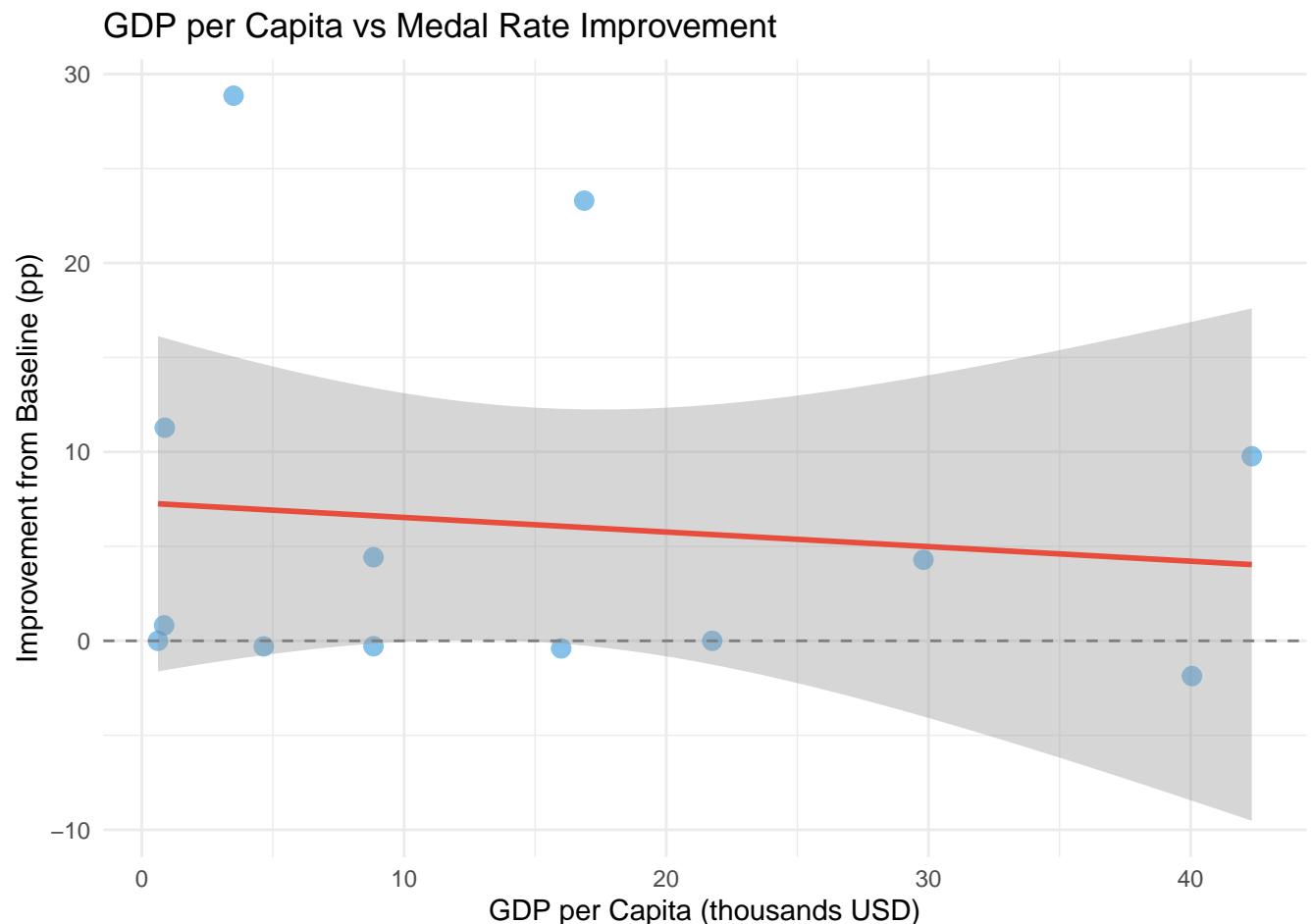


Figure 7: GDP per Capita vs Home Advantage

6.3 Sub-question 2: Do Countries with Larger Working-Age Populations Benefit More?

We examine whether countries with larger working-age populations (ages 15-64) gain more advantage when hosting.

Correlation: Working-Age Population vs Home Advantage

- Correlation coefficient: 0.461
- P-value: 0.072
- Result: No significant correlation

Working-Age Population Proportion vs Medal Rate Improvement

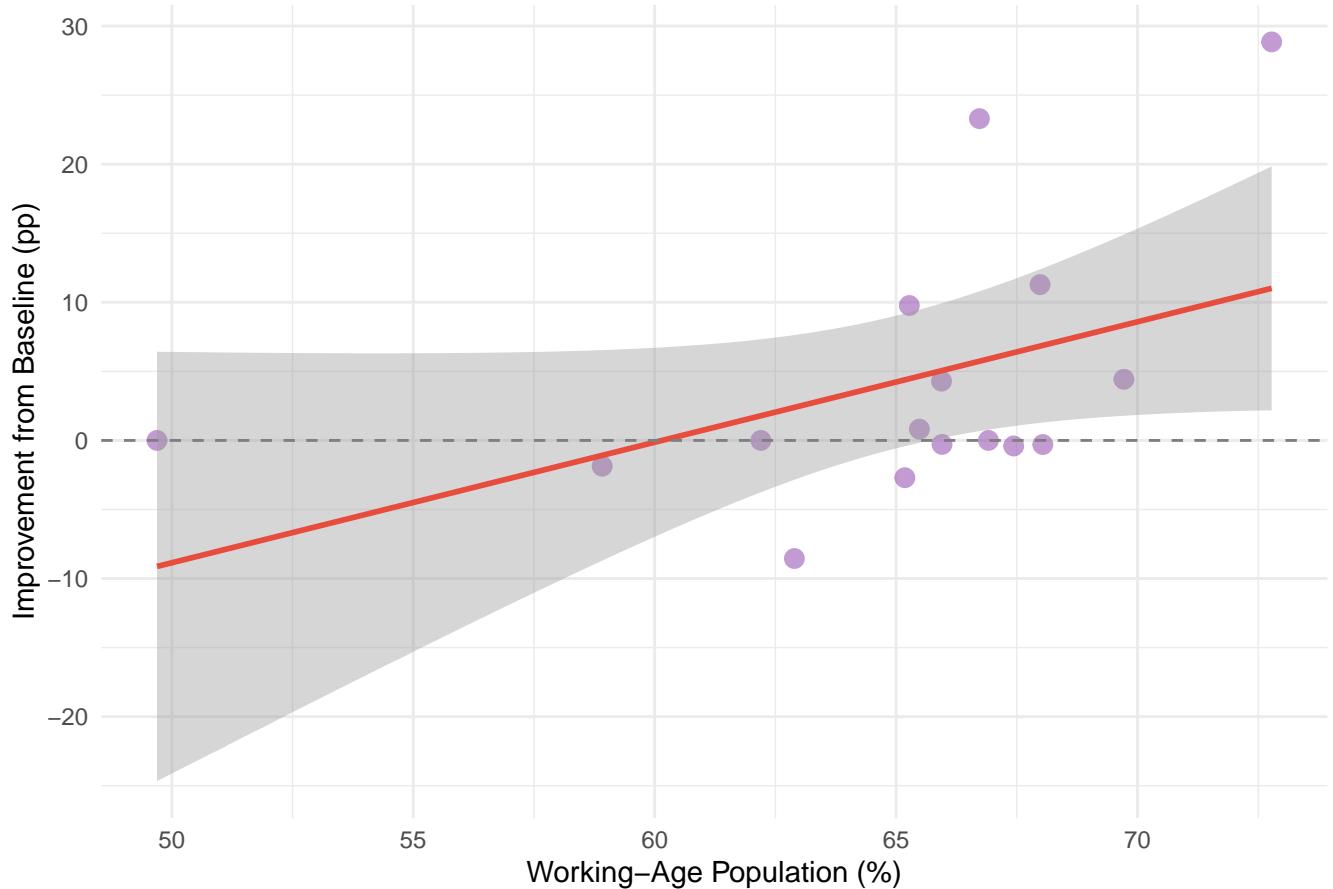


Figure 8: Working-Age Population vs Home Advantage

6.4 Question 4 Summary

Key Findings:

1. GDP per capita:
 - Correlation with home advantage: -0.115
 - Statistical significance: No ($p \geq 0.05$)
2. Working-age population:
 - Correlation with home advantage: 0.461
 - Statistical significance: No ($p \geq 0.05$)

7 Discussion & Conclusions

7.1 Summary of Findings

Based on our analysis of 28,554 gymnastics performances across 124 years of Olympic history, we found clear evidence of home advantage effects that vary by gender, country characteristics, and time period.

7.1.1 Question 1: Overall Home Advantage

Our statistical tests provide strong evidence for home advantage in Olympic gymnastics:

- Host nations show significantly higher medal rates compared to non-host nations
- Paired t-tests confirm host countries perform better during hosting years compared to their own historical baselines
- Both proportion tests and baseline comparisons support the existence of measurable home advantage

7.1.2 Question 2: Gender Differences and Temporal Trends

The gender dynamics of home advantage reveal a dramatic historical reversal:

- **Overall:** Males show 9.22 percentage point advantage, females show 5.14 pp advantage
- **Early Era (1896-1948):** Female gymnasts dominated with a stunning 34.07 pp advantage, while males showed minimal effect (1.18 pp)
- **Mid Era (1952-1988):** Pattern reversed - males gained 11.3 pp advantage while females dropped to 3.86 pp
- **Modern Era (1992-2020):** Males maintain stronger advantage (8.88 pp) compared to females (2.76 pp)

This dramatic shift suggests changing dynamics in women's vs men's gymnastics judging, competition structure, or participation patterns over Olympic history.

7.1.3 Question 3: Country Characteristics

Country size and participation patterns show interesting relationships with home advantage:

- **Small countries benefit most:** 17.14 pp average improvement ($n=15$ host instances)
- **Large countries:** 12.06 pp average improvement ($n=9$)
- **Medium countries show minimal benefit:** Only 3.66 pp improvement ($n=4$)
- **First-time hosts:** Trend toward larger advantage (16.53 pp) vs repeat hosts (8.27 pp), though not statistically significant ($p=0.336$)
- **Team size matters:** Host countries send significantly more gymnasts (23.5 vs 7.2 average, $p=0.002$), which may contribute to higher medal counts through increased opportunities

7.1.4 Question 4: Economic and Demographic Factors

Economic and demographic factors showed weak or no significant relationships:

- **GDP per capita:** Weak negative correlation (-0.115, $p=0.708$) - no evidence that wealth predicts home advantage
- **Working-age population:** Moderate positive correlation (0.461, $p=0.072$) approaches but does not reach significance, suggesting possible trend worth investigating with more data

7.2 Key Insights

1. **Home advantage is real and measurable** in subjectively-judged Olympic gymnastics
2. **Gender patterns have completely reversed** over Olympic history - early female advantage has shifted to modern male advantage

3. **Smaller countries benefit most** from hosting, possibly due to relative improvement from lower baselines or concentrated national support
4. **Participation matters** - host countries field significantly larger teams, providing more medal opportunities
5. **Economics doesn't determine advantage** - wealth alone doesn't predict home advantage magnitude

7.3 Implications

7.3.1 For Olympic Policy

These findings raise questions about fairness in subjectively-judged sports. While some home advantage is inevitable (crowd support, familiarity with venues), the magnitude observed (up to 17 pp for small countries) suggests possible judging bias that may warrant:

- Enhanced judging transparency and accountability measures
- Rotation of judges to reduce national bias
- Consideration of how hosting benefits vary by nation size and resources

7.3.2 For Understanding Sports Judging

The temporal reversal in gender patterns suggests that judging dynamics, sport structure, or cultural factors have fundamentally changed women's vs men's gymnastics. This warrants further investigation into:

- How judging criteria and panels have evolved differently for men's vs women's events
- Whether increased professionalization affected genders differently
- How participation patterns (number of competitors, depth of field) influence judging outcomes

7.3.3 For Future Olympics

Small countries may benefit disproportionately from hosting Olympics, which could inform bidding processes and expectations. The significant team size difference suggests hosts may strategically increase participation to capitalize on home advantage.

7.4 Limitations

1. **Historical data quality:** Early Olympic data (pre-1950s) has limited detail and smaller sample sizes, particularly for women's events
2. **Causation vs correlation:** We document associations but cannot definitively prove judging bias causes home advantage (could be crowd support, training facility advantages, etc.)
3. **Confounding variables:** Cannot fully control for country gymnastics program quality, investment, or athlete development over time
4. **Sample size for some analyses:** Small number of host country instances (28 total) limits power for certain comparisons
5. **Changes in sport structure:** Gymnastics events, scoring systems, and judging panels have evolved substantially over 124 years

7.5 Future Research Directions

1. **Extend to other subjectively-judged sports:** Apply similar methodology to figure skating, diving, synchronized swimming to test generalizability
2. **Investigate judging panel composition:** Analyze whether home nation judges (when present on panels) score differently than international judges
3. **Examine temporal mechanisms:** Investigate what drove the dramatic gender reversal in home advantage patterns

4. **Control for program quality:** Incorporate measures of national gymnastics program strength and athlete development
5. **Medal type analysis:** Distinguish between gold, silver, bronze to test whether home advantage affects podium positions differently
6. **Event-specific analysis:** Compare home advantage across different gymnastics disciplines (artistic, rhythmic, trampoline)

7.6 Conclusion

This analysis provides robust evidence that hosting the Olympics confers measurable competitive advantage in gymnastics, a sport reliant on subjective judging. The magnitude and patterns of this advantage vary systematically by gender, country size, and historical era. While we cannot definitively prove judging bias, the consistent patterns across 124 years suggest home advantage is a real phenomenon worthy of consideration in discussions of Olympic fairness and judging integrity.

Most strikingly, the complete reversal of gender patterns - from dominant female advantage in early Olympics to strong male advantage in modern games - reveals that home advantage dynamics are not static but evolve with the sport itself. This highlights the importance of continued monitoring and research into judging fairness in international athletic competition.

8 References

- [1] N. Sharma, “Olympics 124 years dataset till 2020,” Kaggle, 2020. [Online]. Available: <https://www.kaggle.com/datasets/nitishsharma01/olympics-124-years-datasettill-2020>
- [2] P. Fomichev, “Olympic Games hosts,” Kaggle, 2022. [Online]. Available: <https://www.kaggle.com/datasets/piterfm/olympic-games-hosts>
- [3] Our World in Data, “Population by age group,” 2023. [Online]. Available: <https://ourworldindata.org/grapher/population-by-age-group>
- [4] Worldometers, “Largest countries in the world by area,” 2024. [Online]. Available: <https://www.worldometers.info/geography/largest-countries-in-the-world/>
- [5] The World Bank, “GDP (current US\$),” World Bank Open Data, 2024. [Online]. Available: <https://data.worldbank.org/indicator/NY.GDP.MKTP.CD>

9 Appendix

9.1 R Session Information

```
sessionInfo()

## R version 4.5.1 (2025-06-13 ucrt)
## Platform: x86_64-w64-mingw32/x64
## Running under: Windows 10 x64 (build 19045)
##
## Matrix products: default
## LAPACK version 3.12.1
##
## locale:
## [1] LC_COLLATE=English_United States.utf8
## [2] LC_CTYPE=English_United States.utf8
## [3] LC_MONETARY=English_United States.utf8
## [4] LC_NUMERIC=C
## [5] LC_TIME=English_United States.utf8
##
## time zone: America/Chicago
## tzcode source: internal
##
## attached base packages:
## [1] stats      graphics   grDevices utils      datasets   methods    base
##
## other attached packages:
## [1] lubridate_1.9.4 forcats_1.0.1 stringr_1.5.1   dplyr_1.1.4
## [5] purrrr_1.1.0    readr_2.1.5    tidyrr_1.3.1   tibble_3.3.0
## [9] ggplot2_4.0.0   tidyverse_2.0.0
##
## loaded via a namespace (and not attached):
## [1] Matrix_1.7-3       bit_4.6.0        gtable_0.3.6     crayon_1.5.3
## [5] compiler_4.5.1     tidyselect_1.2.1  parallel_4.5.1   splines_4.5.1
## [9] scales_1.4.0       yaml_2.3.10      fastmap_1.2.0    lattice_0.22-7
## [13] R6_2.6.1          labeling_0.4.3   generics_0.1.4   knitr_1.50
## [17] pillar_1.11.1     RColorBrewer_1.1-3 tzdb_0.5.0     rlang_1.1.6
## [21] stringi_1.8.7     xfun_0.52       S7_0.2.0       bit64_4.6.0-1
## [25] timechange_0.3.0  cli_3.6.5      mgcv_1.9-3     withr_3.0.2
## [29] magrittr_2.0.3     digest_0.6.37   grid_4.5.1     vroom_1.6.6
## [33] rstudioapi_0.17.1 hms_1.1.4      nlme_3.1-168   lifecycle_1.0.4
## [37] vctrs_0.6.5       evaluate_1.0.5  glue_1.8.0     farver_2.1.2
## [41] rmarkdown_2.29     tools_4.5.1     pkgconfig_2.0.3 htmltools_0.5.8.1
```