

Olympic Swimming

Background Info

Swimming in the Olympics

- The Olympics is an international multi-sport event that has existed in its modern form since 1896 that was inspired by the ancient Olympic Games in Greece.
- Since 1994, there has been alternating events every two years for winter sports and summer sport.
- Swimming has been an event in the Summer Olympics since the very beginning and has included women since 1912.



Swimming in the Olympics

- In the Olympic, swimming has multiple events.
- The events are split by a few variations:
 - Style: Freestyle, breaststroke, backstroke, butterfly, medley, and relay
 - Distance: 50m, 100m, 200m, 400m, 800m, 1500m, 4x100, and 4x200
 - Gender: Men, women, and mixed
- All together there is currently 35 officially recognized swimming events in the Olympics.
- Since the beginning, there have been 84 different teams who have competed.

Goal of the Study

- Observational study
 - Simply gathering public data from official modern Olympic Games
- Gather information on one of the longest occurring and most popular sports in the Olympic Games.
- Investigate the changes over time in events, athletic ability, and success of specific teams.

The 4 Types of Olympic Swimming Strokes

Backstroke

OR (100m): **51.85 sec**



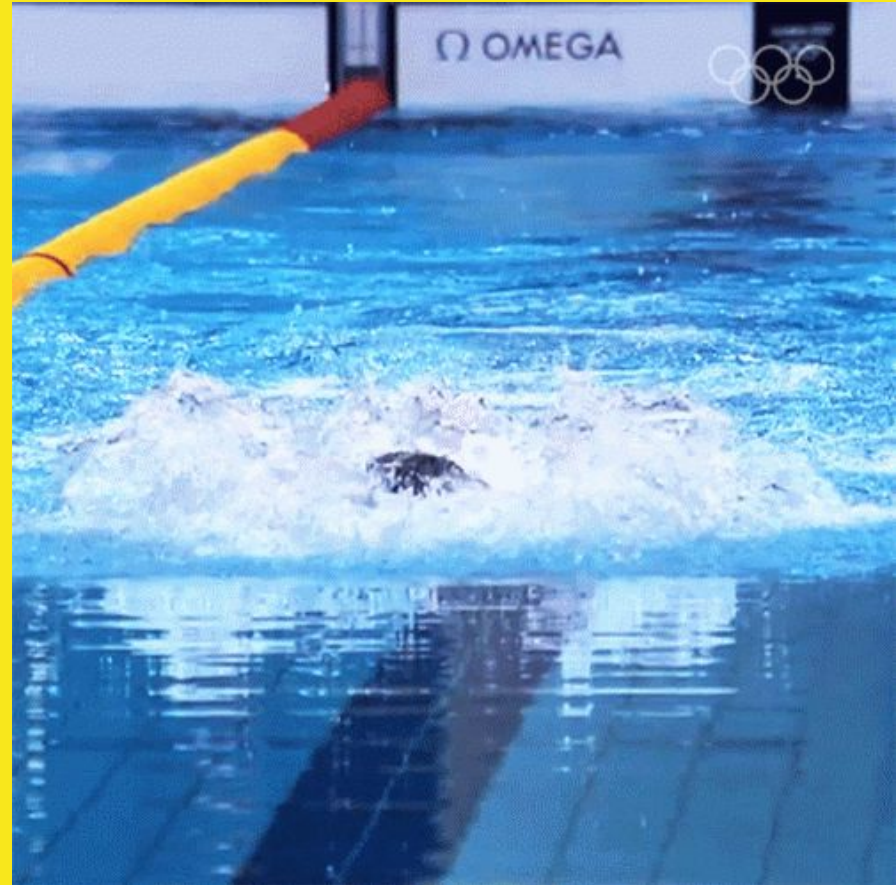
Breaststroke

OR (100m): **57.13 sec**



Butterfly

OR (100m): **49.45 sec**



Freestyle

OR (100m): **47.02 sec**



Data and Methodology

How was the data collected?

- Data is publicly available Olympic Result data
- Data is from Final Races
 - qualifying round times are not included in the dataset
- Key Variables
 - Country of the Athlete
 - Event Stroke (breaststroke, butterfly, backstroke, etc)
 - additionally, length of event (100m, 200m, 400m, etc)
 - Final Time
- Observational Units are Olympic Swimming Athletes
- 4359 Olympic Swimmers from 1912–2020

Excluded Athletes

- 129 Athletes were excluded from our study
- Some athletes did not have any final time recorded, so were excluded
- Some athletes were disqualified from the race or did not finish the race, and so we excluded these athletes as well
- Total of 4230 Athletes in our dataset

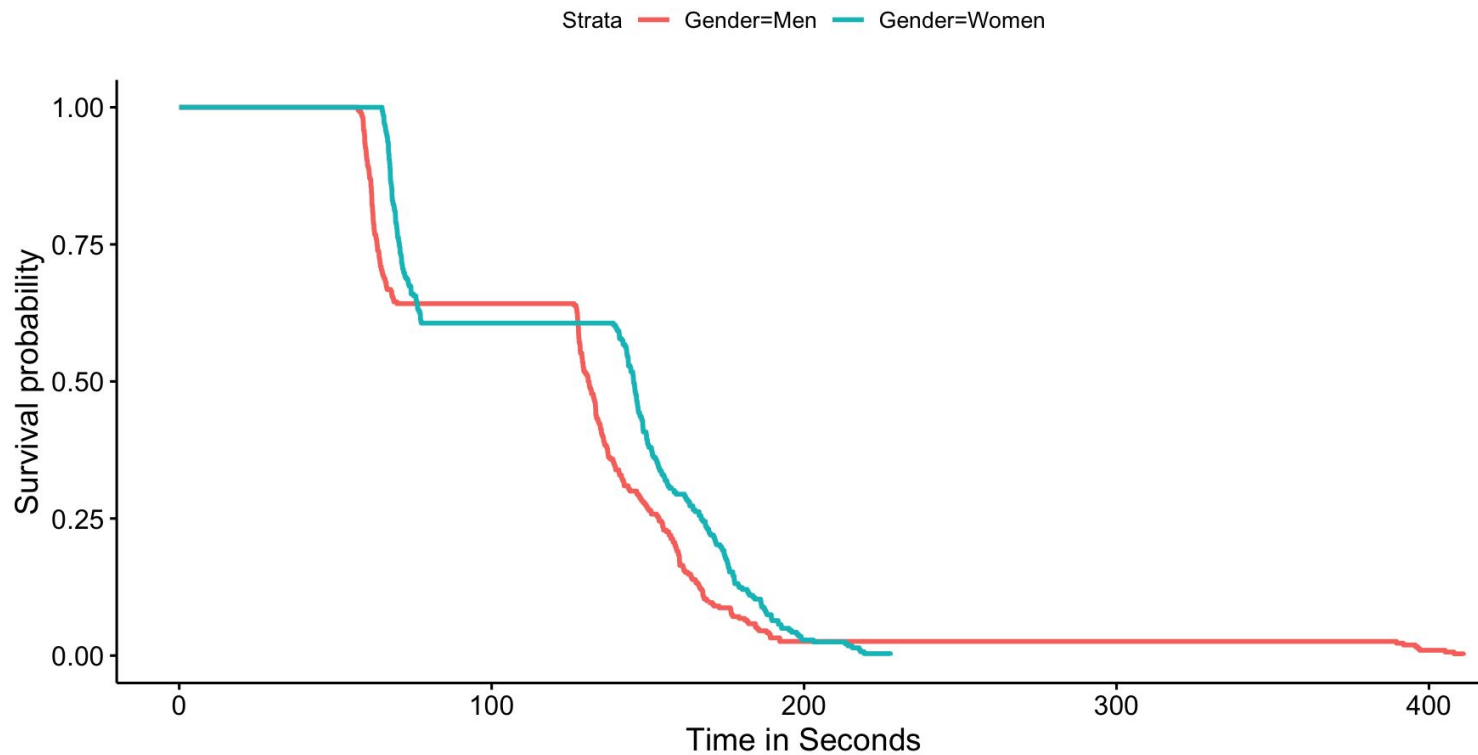


Results/Analysis

**Q₁: How do times differ
between men and women?**

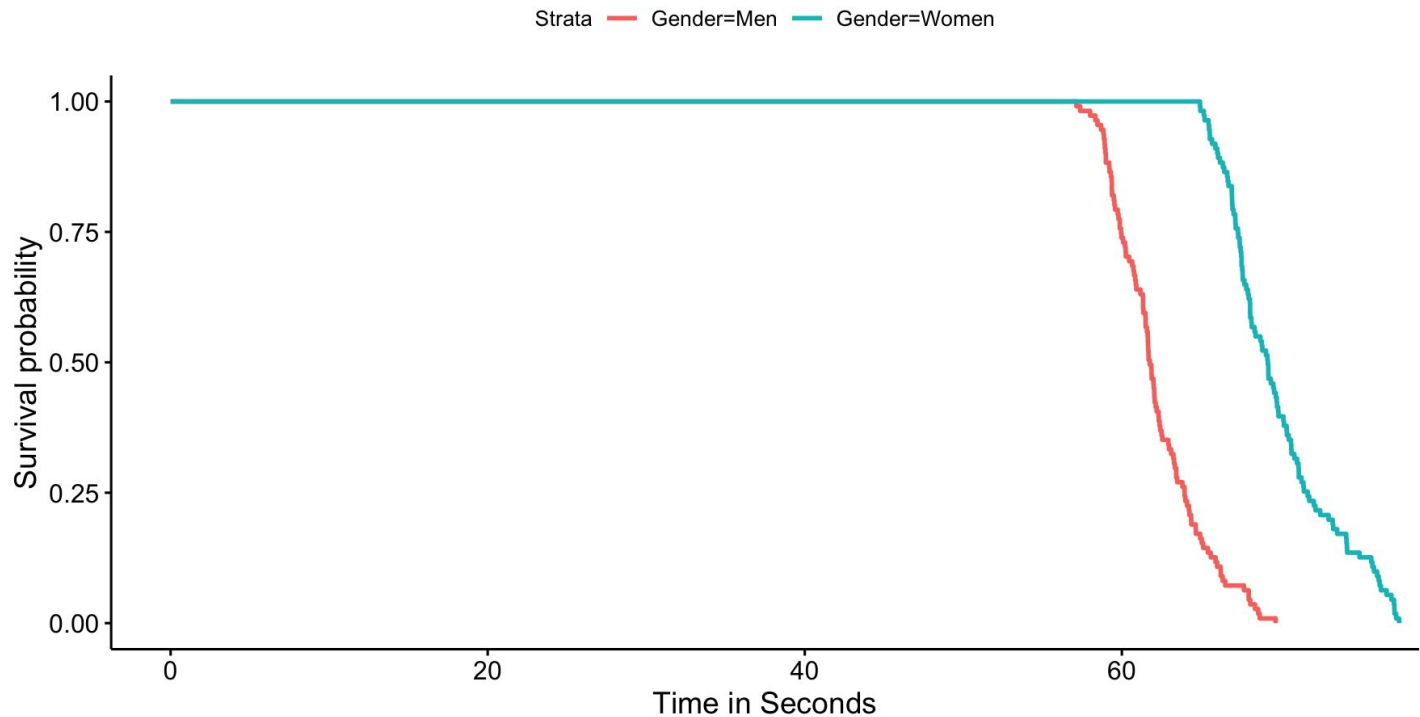
Breaststroke: Men vs Women (KM Curve)

Olympic Breaststroke Survival Probabilities: Men vs Women

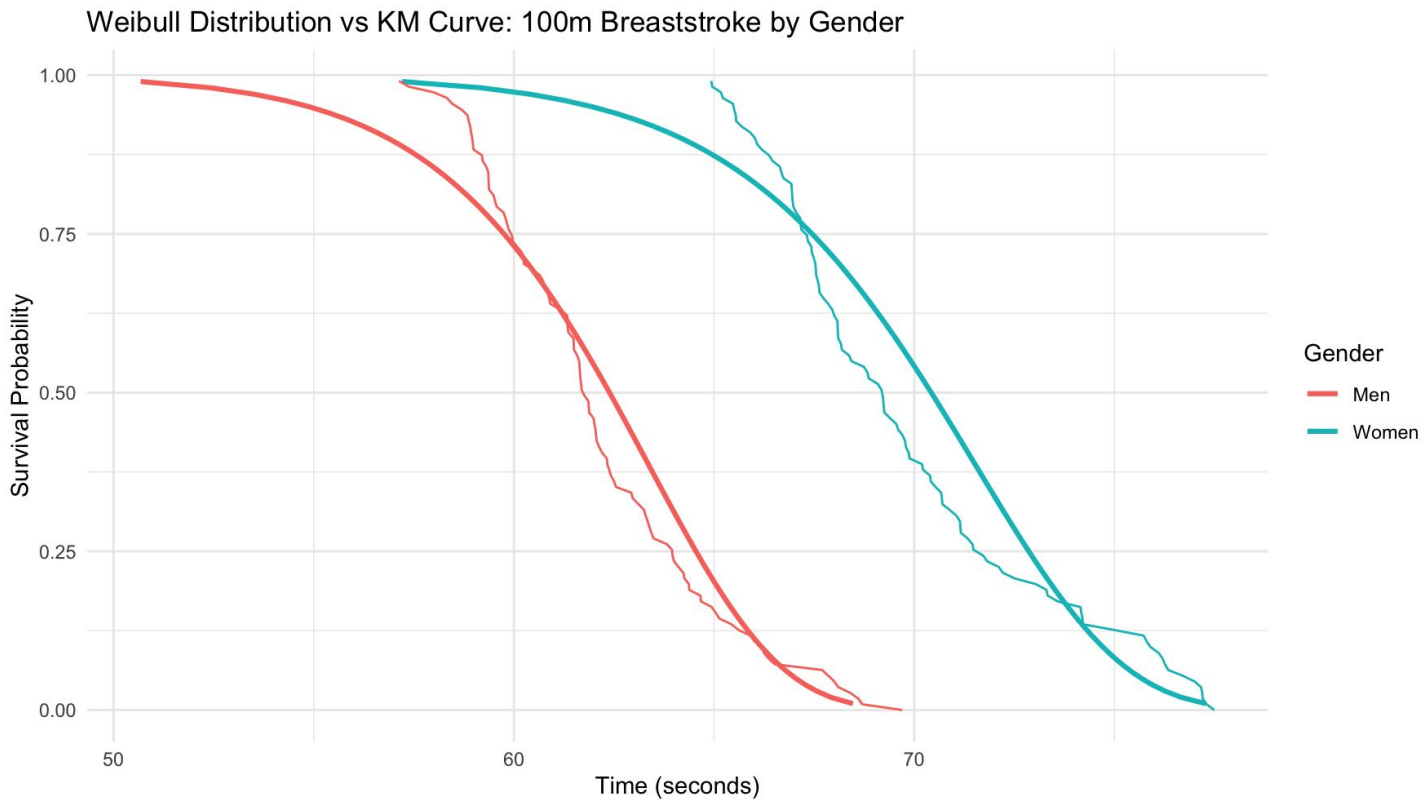


100m Breaststroke: Men vs Women (KM Curve)

Olympic Breaststroke Survival Probabilities: Men vs Women (100m)

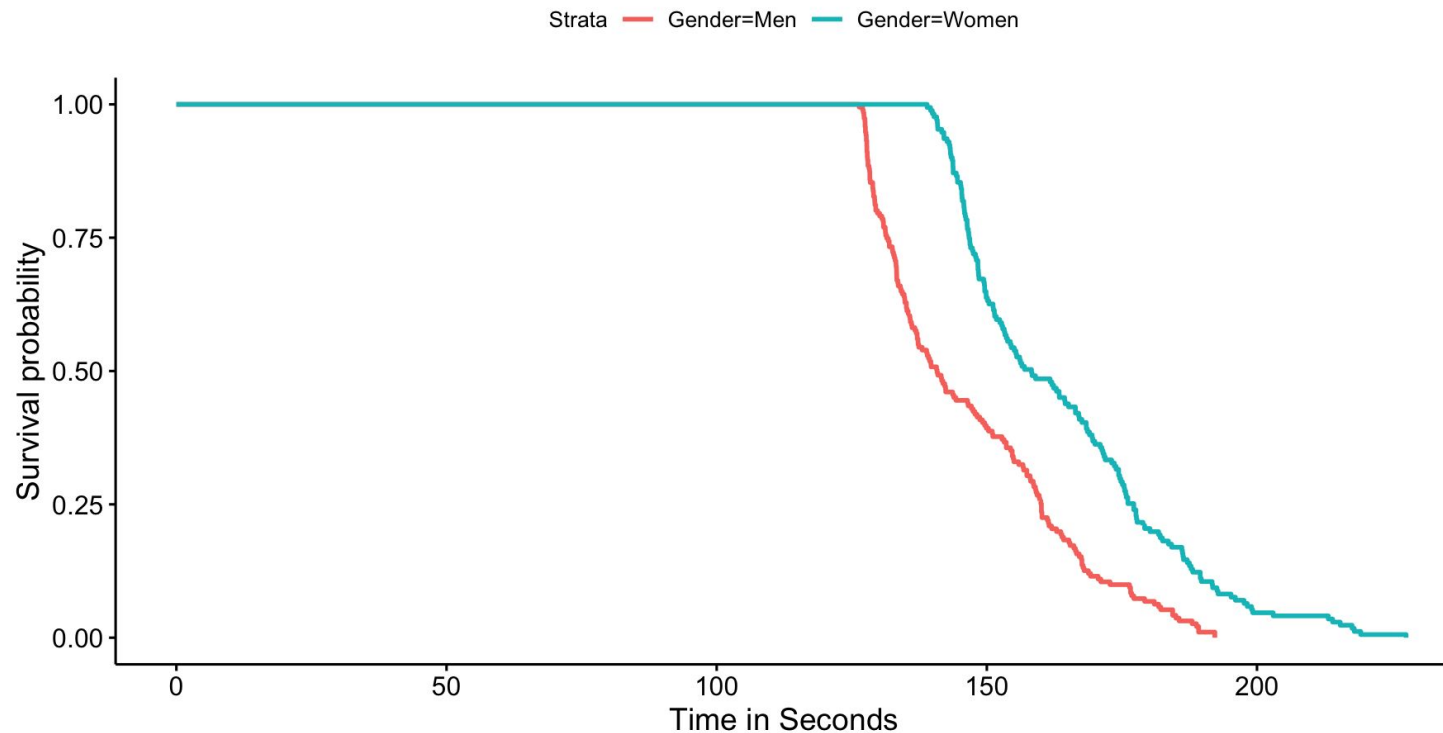


Weibull Model vs KM Curve: 100m Breaststroke

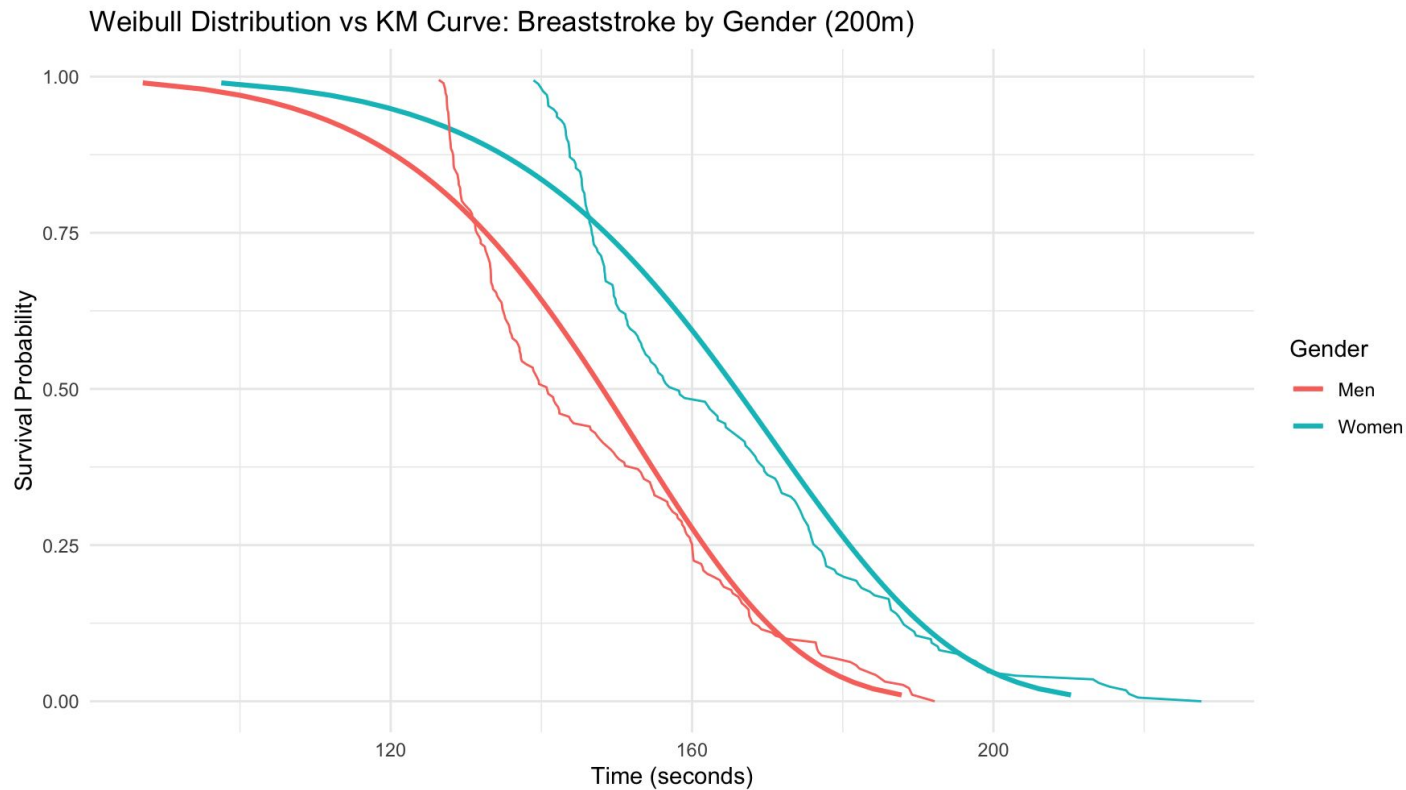


200m Breaststroke: Men vs Women (KM Curve)

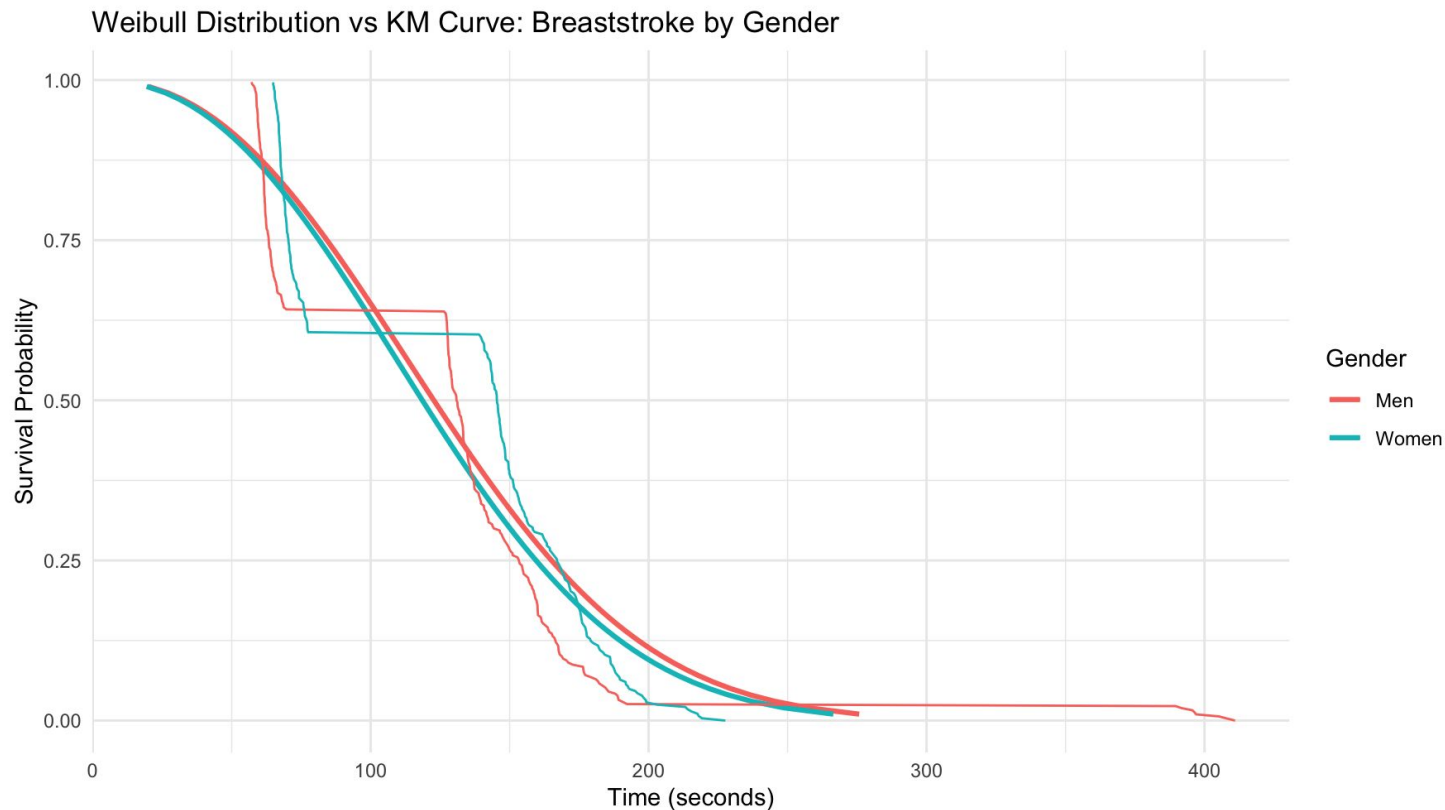
Olympic Breaststroke Survival Probabilities: Men vs Women (200m)



Weibull Model vs KM Curve: 200m Breaststroke



Why was it important to separate 100m and 200m?



Comparing Goodness of Fit

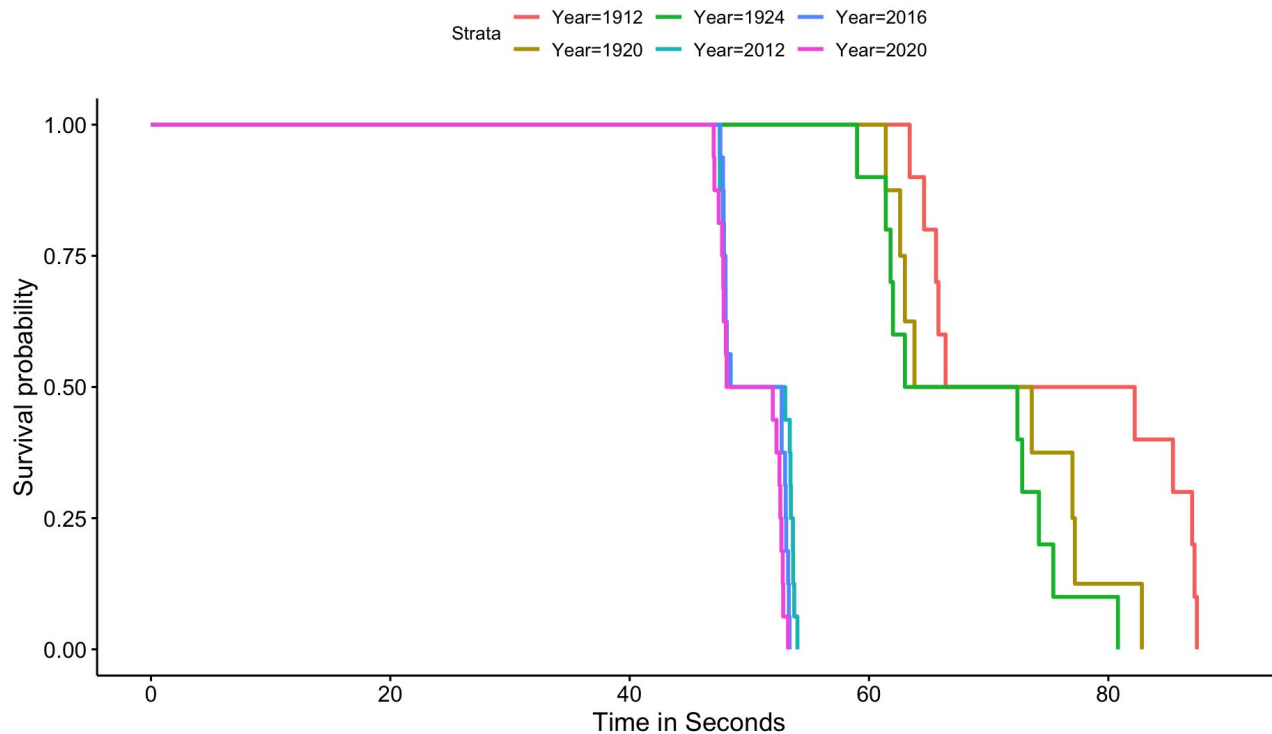
- 100m Model: 1190.82
- 200m Model: 3230.178
- Combo Model: 6393.7018

*Lower = Better

**Q2: How have finish times
changed over the years?**

100m Freestyle: 1910s vs 2010s (KM Curve)

Olympic 100m Freestyle Survival Probabilities: 1910s vs 2010s



$< 2 \times 10^{-16} \%$

or

[illegible]

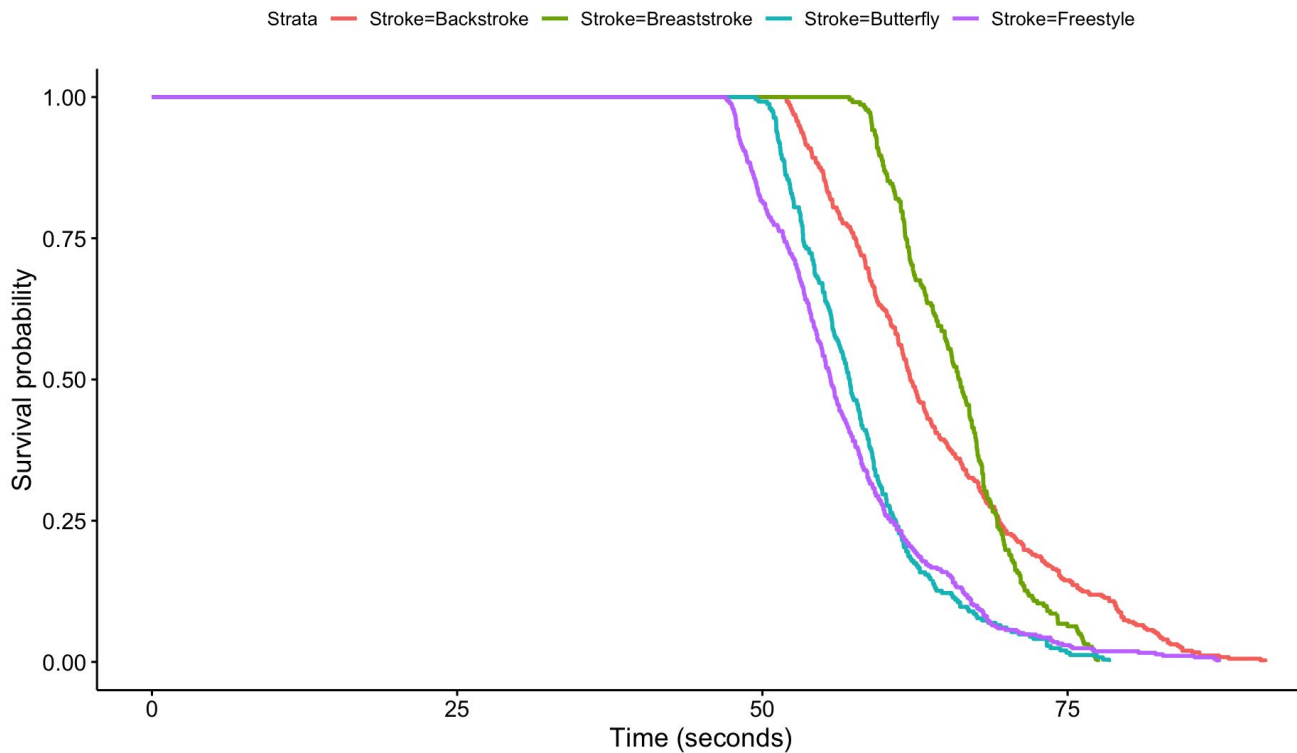
chance of seeing these differences in Years by Random Chance according to Log Rank Test

STATISTICALLY SIGNIFICANT

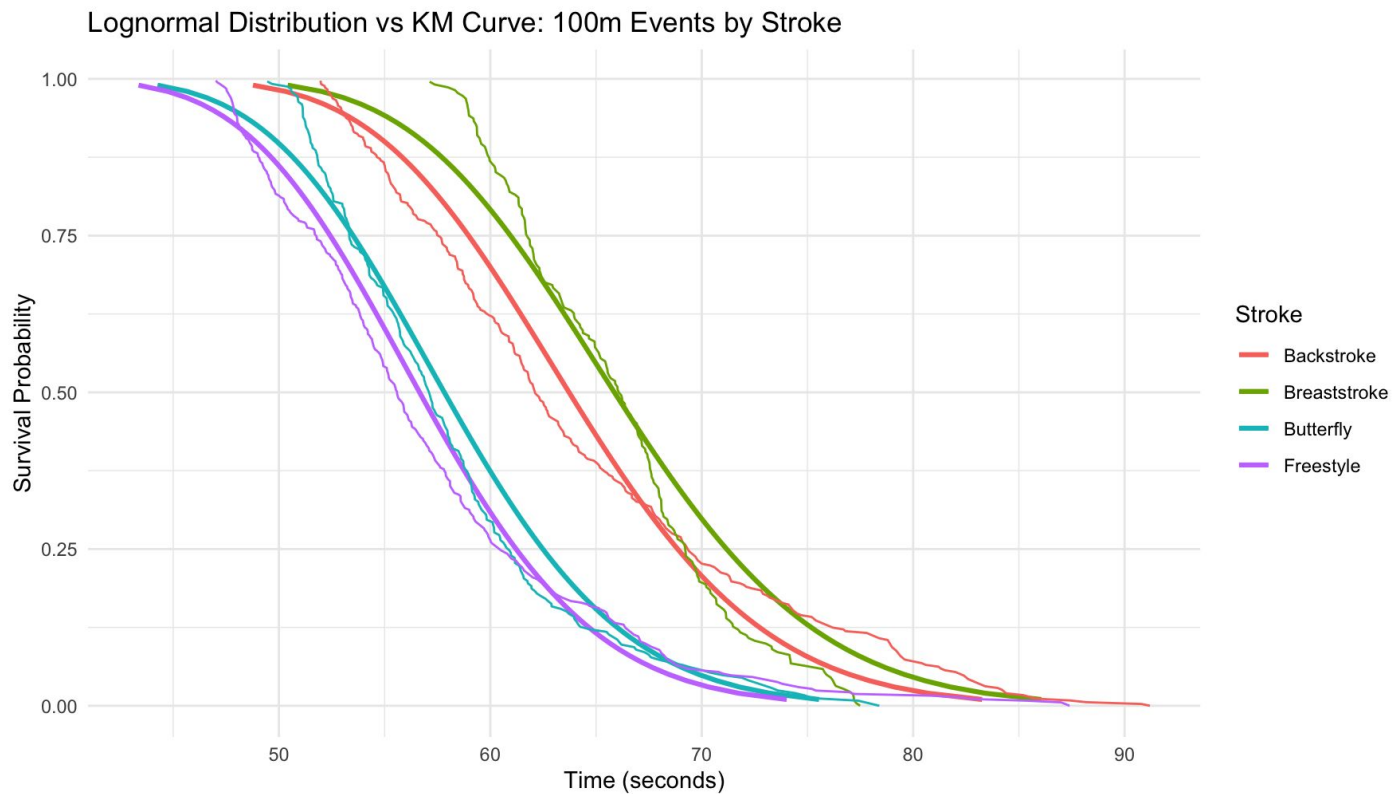
**Q3: How do finish times vary
across strokes?**

100m Survival Probabilities by Stroke

Olympic 100m Swimming Survival Probabilities by Stroke



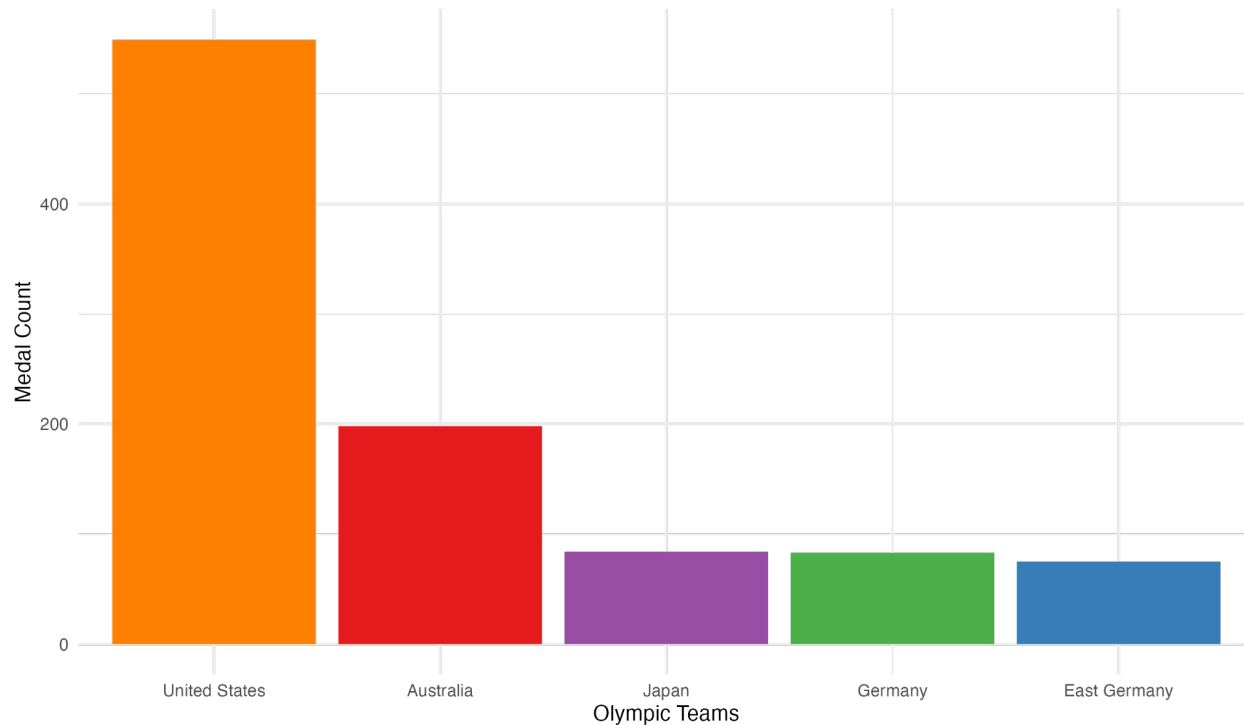
Lognormal Model vs KM Curve: 100m Events by Stroke



**Q4: What teams are most
successful?**

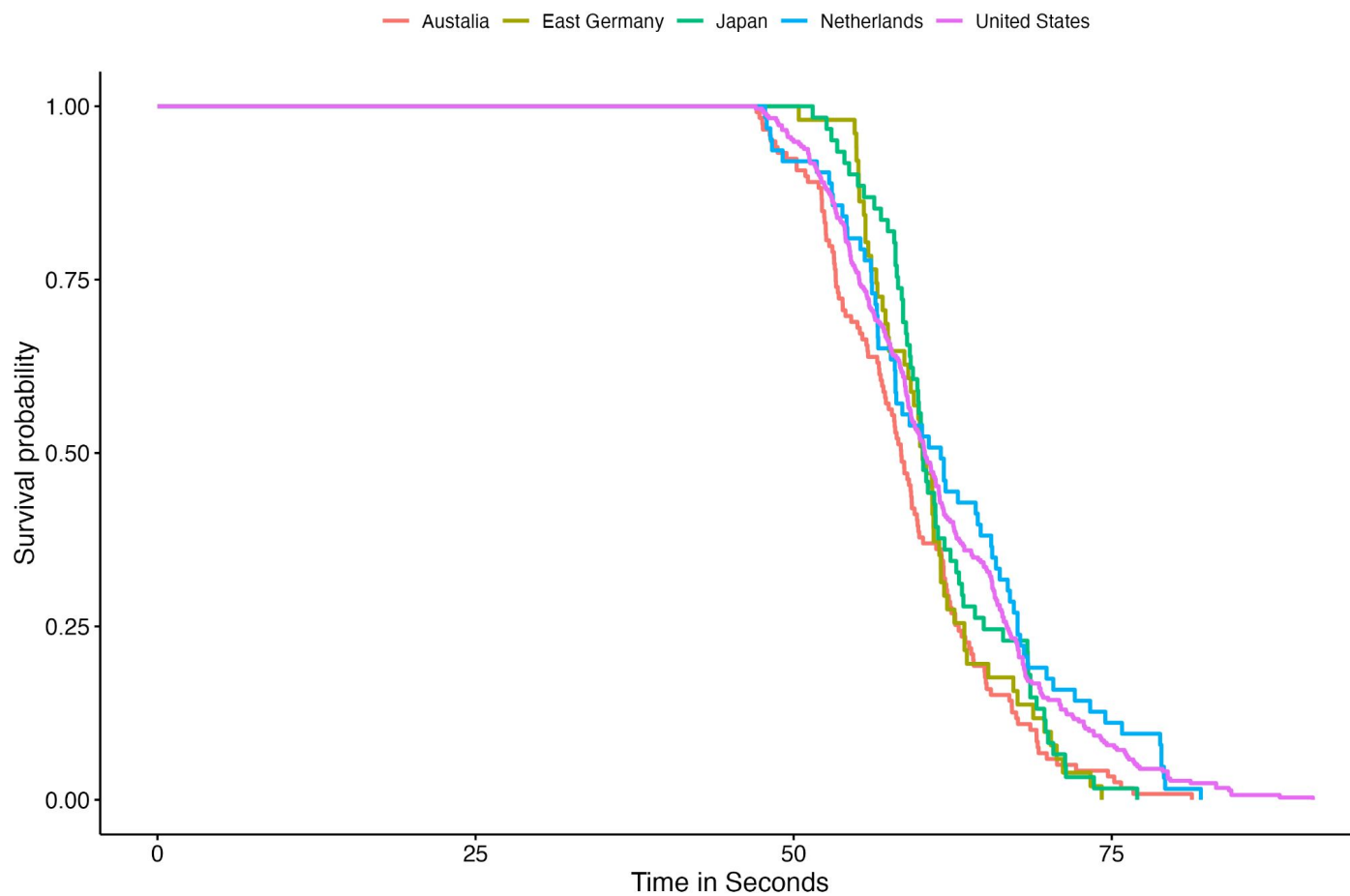
Triumphant Teams

Olympic Swimming Winners



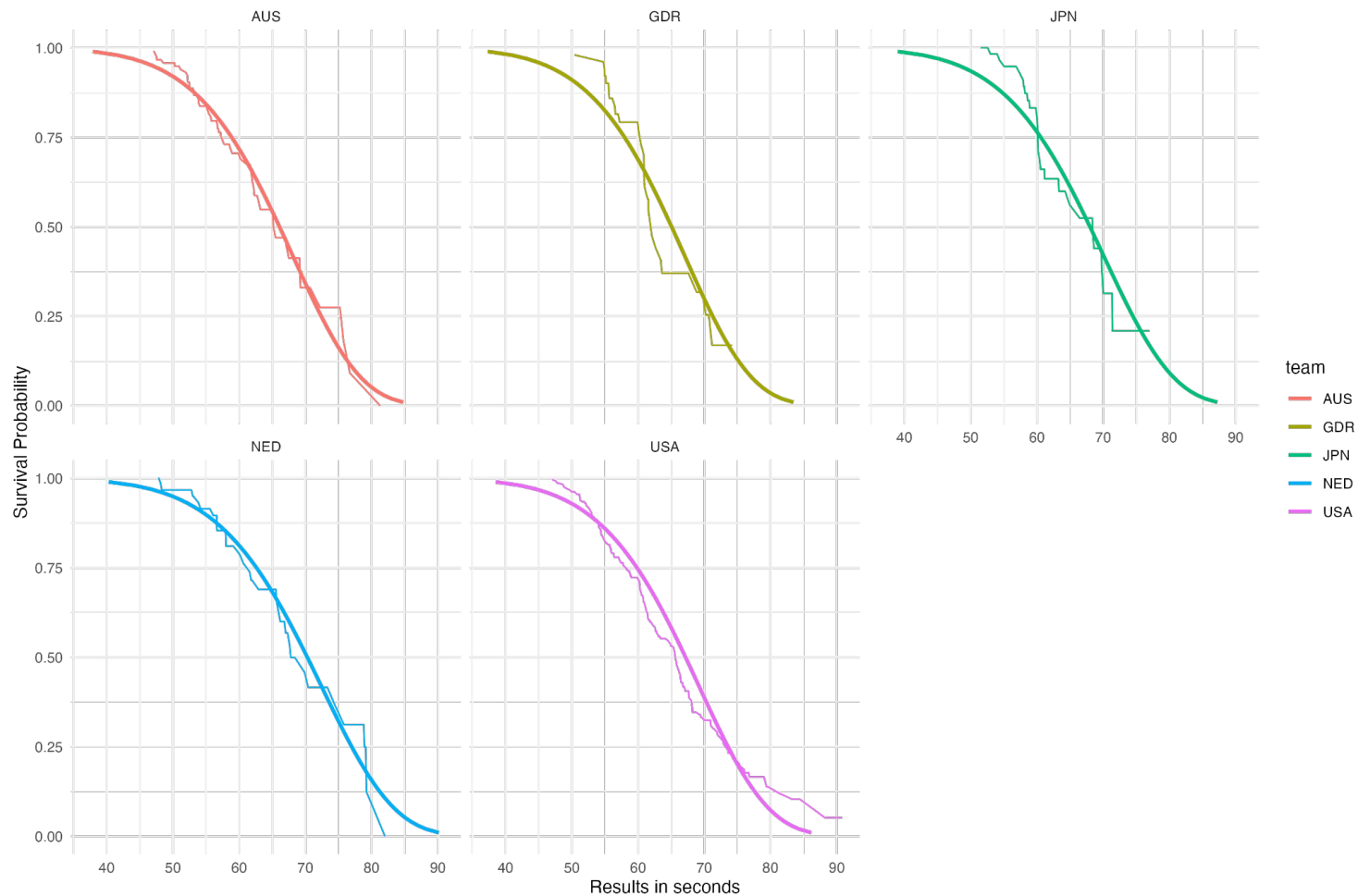
KM Curve

Top Teams in 100 Meter Events



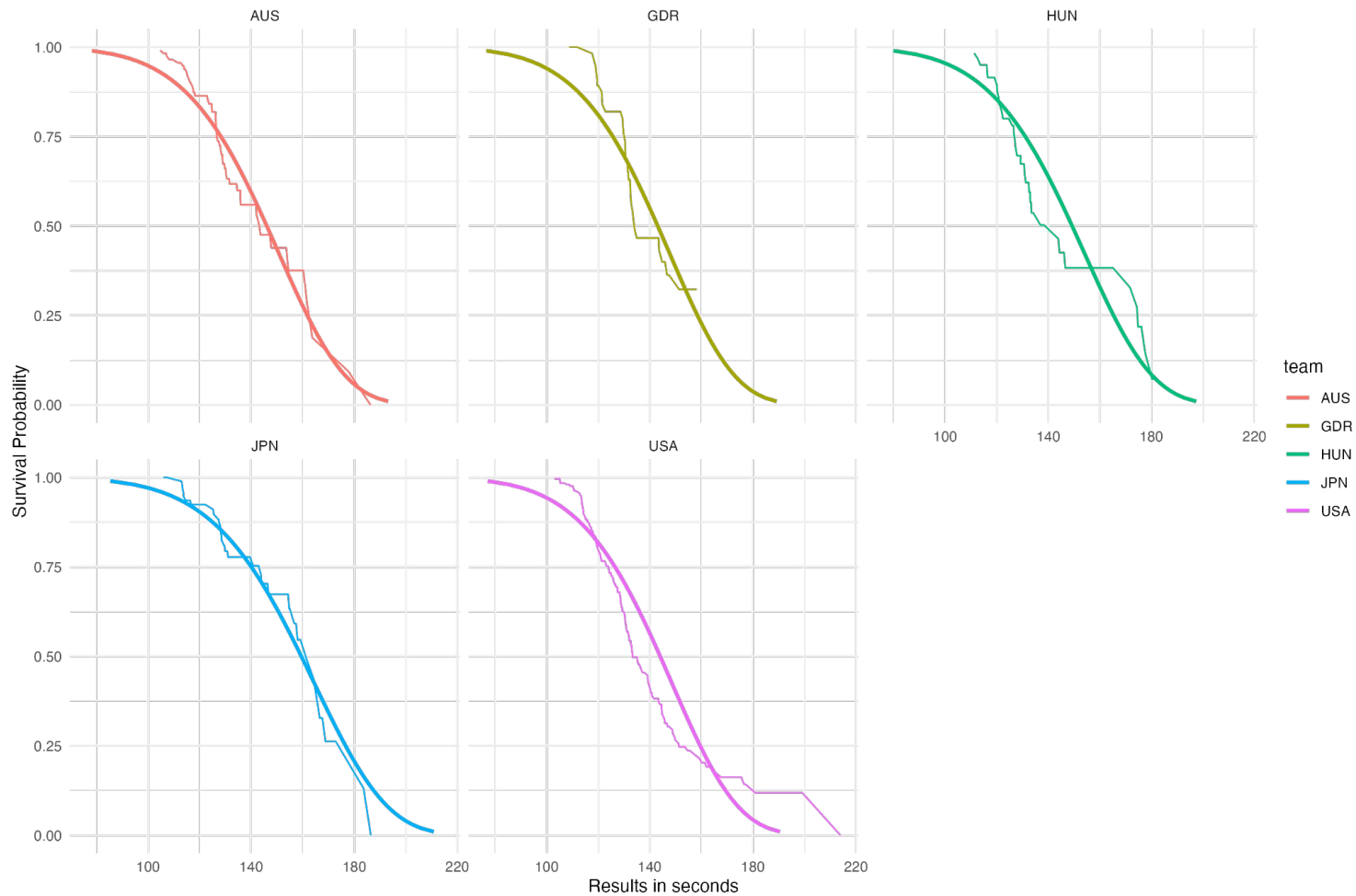
100 Meters

Champions of the 100 Meter Events



200 Meters

Champions of the 200 Meter Events

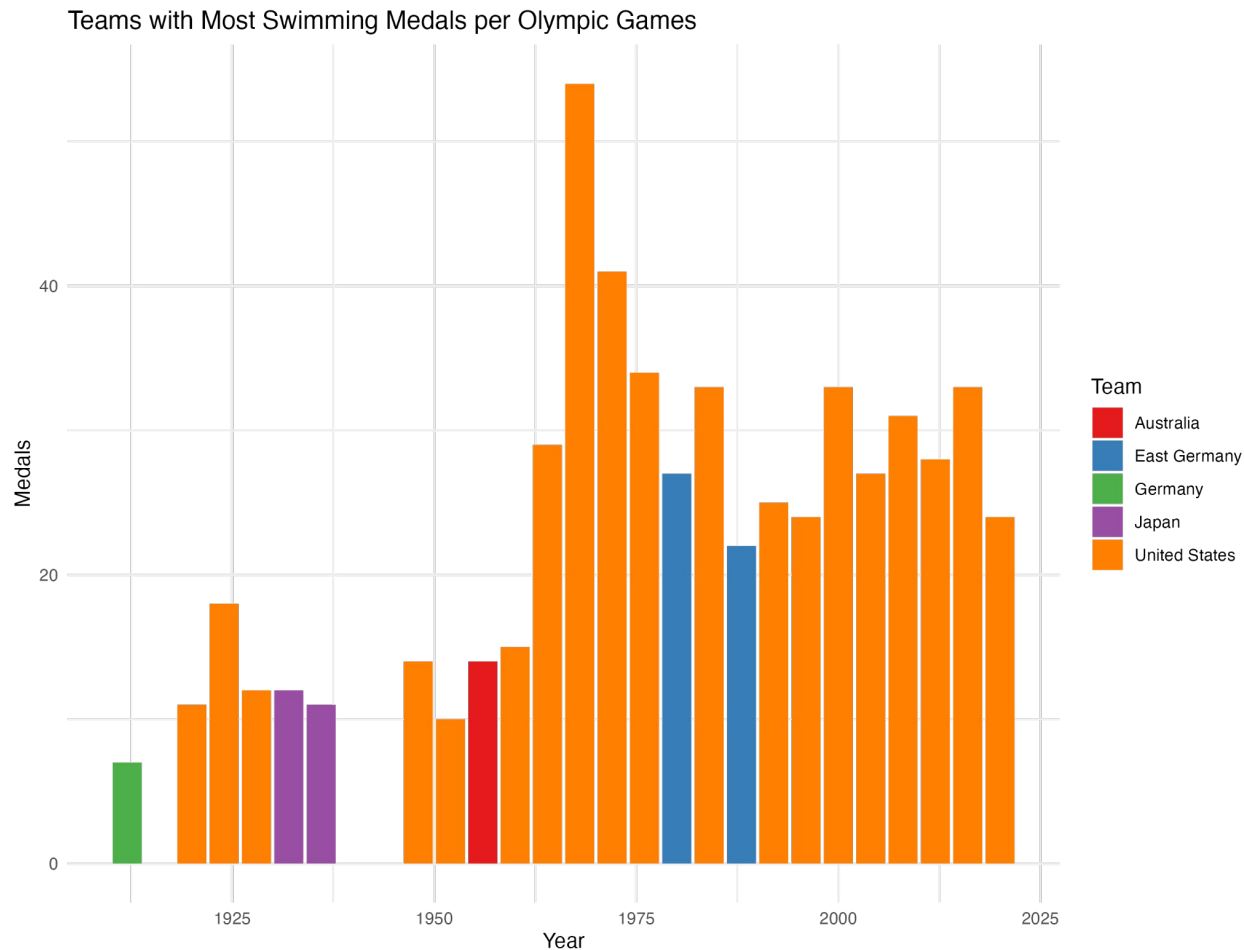


Are there other factors?

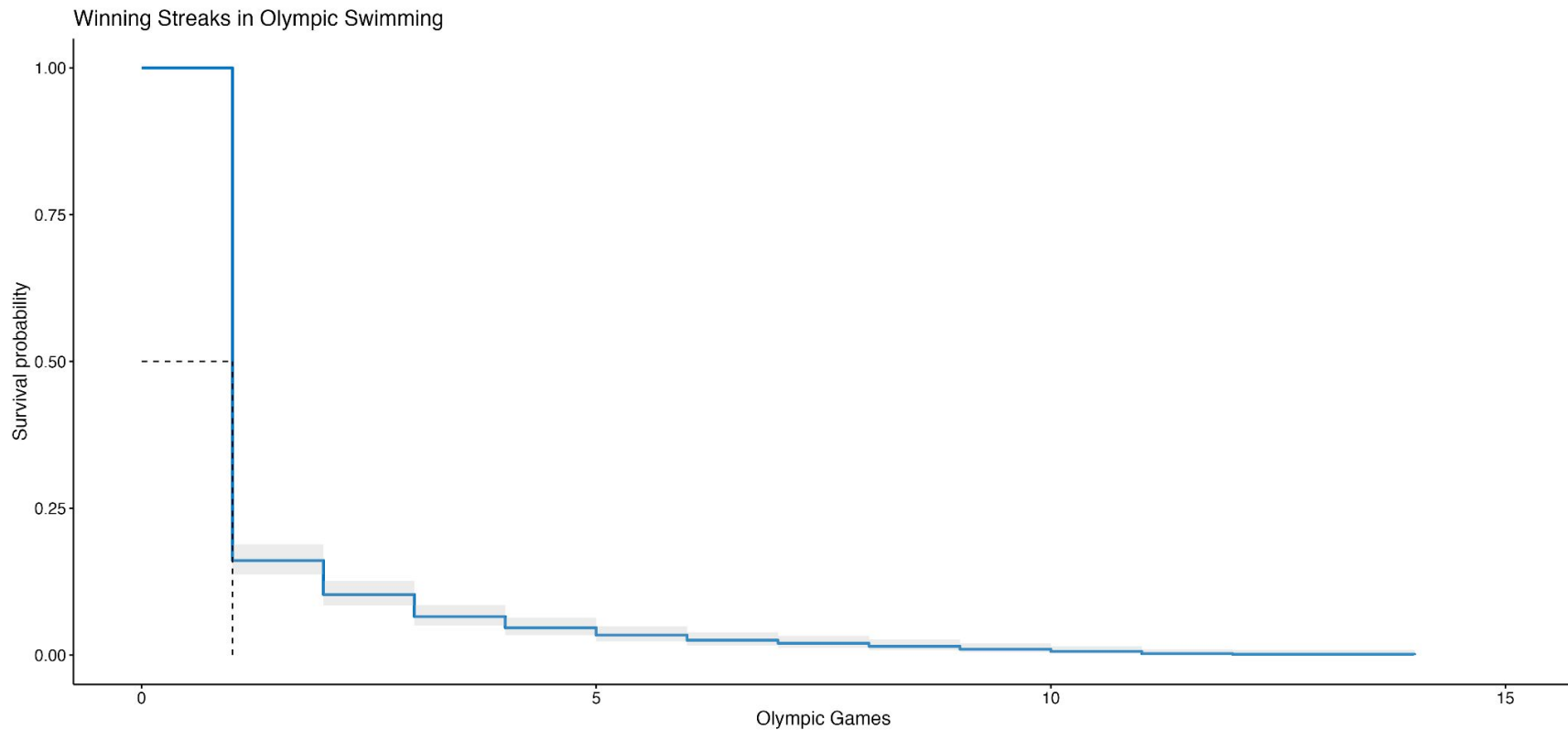
- A parametric model for the 100 meters events was ran with the two covariates– Team and Stroke.
- AIC for just Team as a covariate was 2558.457
- AIC for Team+Stroke was 4189.55
- No need to complicate it!

Q5: Is success consistent?

Winning over
the years

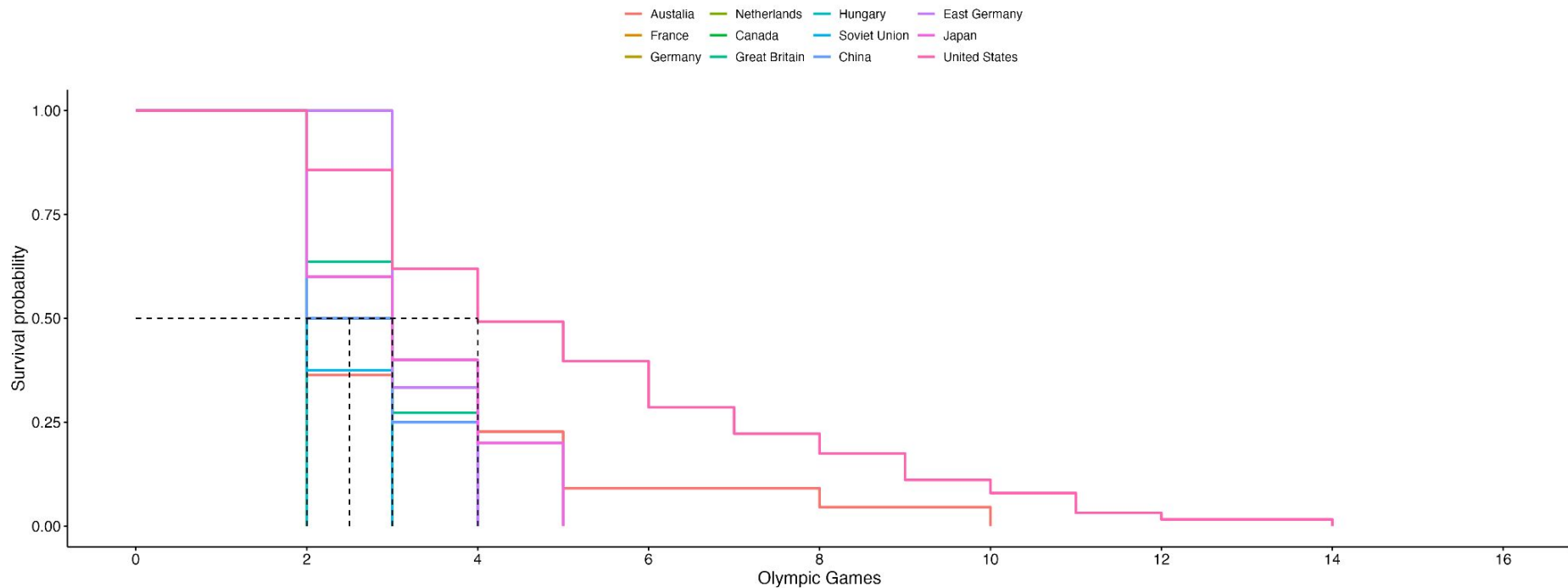


Winning streaks



Winning streaks for teams

Teams with Winning Streaks in Olympic Swimming



Conclusions

Takeaways

- On average, Male swimmers are faster than Female swimmers
 - Separating events by Stroke & Distance creates more accurate models than just separating by Stroke
- Swimmers have gotten significantly faster over the past 100 years
- Fastest event to slowest: Freestyle, Butterfly, Backstroke, Breaststroke
- The United States and Australia are the most successful teams
 - Team is more indicative of faster speeds than Stroke
- Most teams do not win an event twice in a row, however, the United States is the exception to that
 - They much more likely to win an event multiple times in a row

Limitations

- There are four missing Olympic years from the data (1896, 1900, 1904, and 1908) which means we are missing key data for our analyses across years.
- Since we are focusing on results times, it's important to note that the majority of world records in swimming have not been broken at the Olympic Games so this is not a complete representation of international competitive swimming.
- Some events have been in the Olympics for a longer period of time than others, which affects the analysis as events that have been added more recently are typically more competitive overall

Future Steps

- Our analysis focused on the individual events, but there are multiple relay events that can be analyzed in the future.
- Looking into individual athletes and how their successes influence the overall outcomes since there are athletes who have overwhelming wins.
- Marathon swimming is an event that is not included in the data that could be interesting to explore since it is so new.
- Additionally, other international events like the World Aquatics Championships would be very interesting to analyze.