What factors drive success in Olympic gymnastics?

Abby Gopal | Sports Analytics: STAT 1341

Background

Olympic women's gymnastics is a sport where athletes compete in individual and team events on different apparatuses (Uneven Bars, Floor, Beam and Vault). Their scores are based on the combination of execution and difficulty for a specific event. Because the scoring is fairly complex, analyzing patterns can help identify which elements strongly predict success.

Gymnastics scoring is influenced by multiple factors. By analyzing patterns in scores, we can identify which elements most strongly predict the best scores, thus leading to improvement in performance.

I chose this topic as I think it is extremely underrated in the sports analytics world. Because its popularity typically rises only once every four years, this is understandable. Yet the amount of data and the different combinations of scores makes it a great source for analysis. I also like how it is much different from your typical football/basketball data as the scoring is complex in a different way.

Research Objective

Goal: To determine the key factors that drive success in Olympic gymnastics.
Hypothesis: We expect that execution scores likely affect the total score for a gymnast more than difficulty score. However, a combination of both, is what makes the well-rounded, best gymnast.
Outcome: Total score, whether that be for all around, team, individual, or specific events.
Key Predictors: Difficulty(D), Execution(E), Penalties all specific to each event, as well as individual and team performances.

In the end, I am trying to compare and rank the athletes based on their performance metrics, as well as explain correlations and predictors that are significant to impacting a total score.

Data Collection

I assumed it would be fairly difficult to find existing data. Upon my search, I went to USAgym.org, where I found pdfs of data from the Paris 2024 Olympics. These included data for multiple events and trials, including the men's and women's teams. The data was presented nicely, organized, and clean, however it was set up in a pdf form, and I scraped the web looking for any type of csv or excel files but couldn't find any. I tried pdf converters, but ultimately decided that if I wanted clean, precise data, I would have to create a script to do that for me.

In cursor, I wrote a script that took in 3 pdfs:
- **Women's Event Finals**
- **Women's Team Final**
- **Women's All-Around Final**

then converted all of these to clean csv files that I could then work with in Google Sheets.

The data was different for each different file, as certain sections only have leading gymnasts, certain countries etc. But overall, the data had about 30 gymnasts, and 10 different countries, with data from all 4 events in 2024.

Methodology

I conducted an exploratory data analysis (EDA) to identify patterns and relationships between difficulty scores, execution scores, and overall performance in Olympic gymnastics. I thought that this method would be appropriate to use for this project as this type of data analysis is used to understand a dataset's main characteristics and uncover patterns, relationships, and trends.

Prepping Data: There were a few new variables that needed to be created, as I was specifically focusing on difficulty and execution. This included each athlete's average difficulty score, average execution score, total all-around score, and score consistency (measured using the STDEV function across events).

All analysis and visualizations were completed in Google Sheets, using built-in functions and chart tools:
- AVERAGE(), CORREL(), STDEV() for summary statistics
- Scatterplots, column charts, and line charts for visual comparisons

Results

Execution scores had a much stronger correlation and made up a significant part of all around score than difficulty. Not really relevant to each event, however we do see a spike in execution on vault.
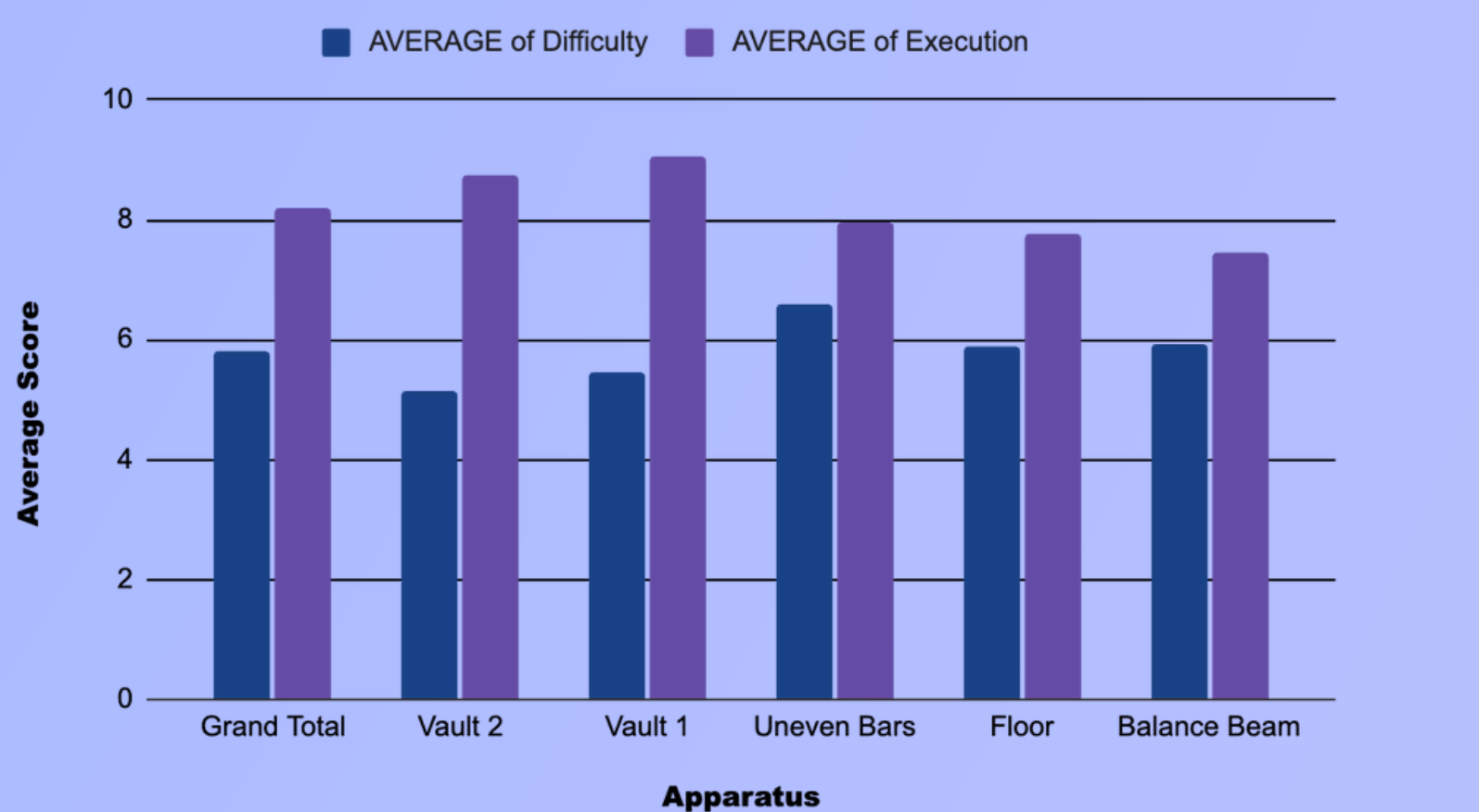

- The scatterplot points are closer to the trend line for execution, with an R² = 0.569, indicating a **moderately strong** linear relationship.
- The points are more scattered with R² = 0.22, showing a **weak** relationship between difficulty and overall success.

Execution scores are more strongly associated with total performance than difficulty scores. Medalists cluster near high execution values, suggesting that consistent, high-quality routines contribute more to success than attempting higher-difficulty skills alone.
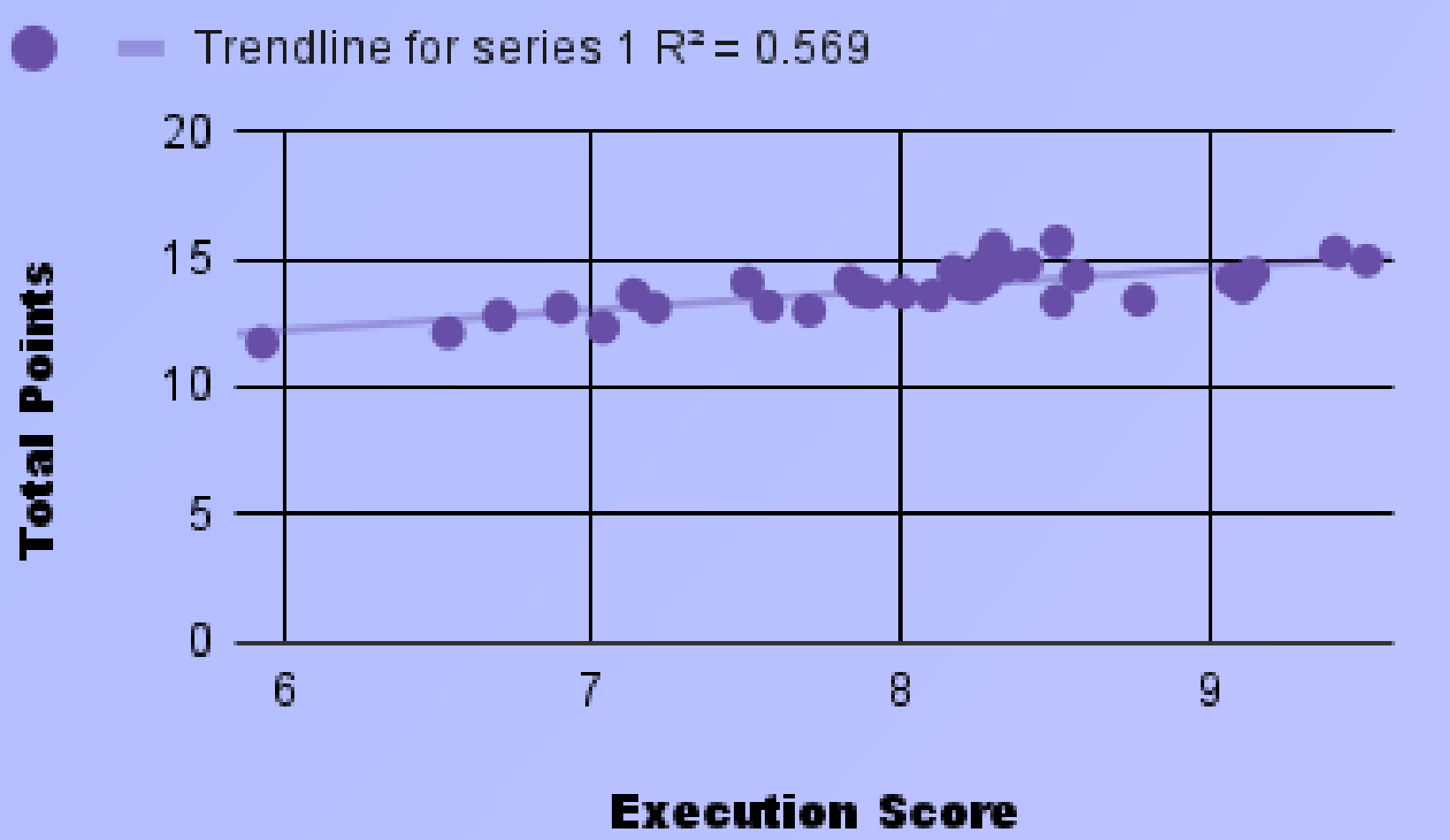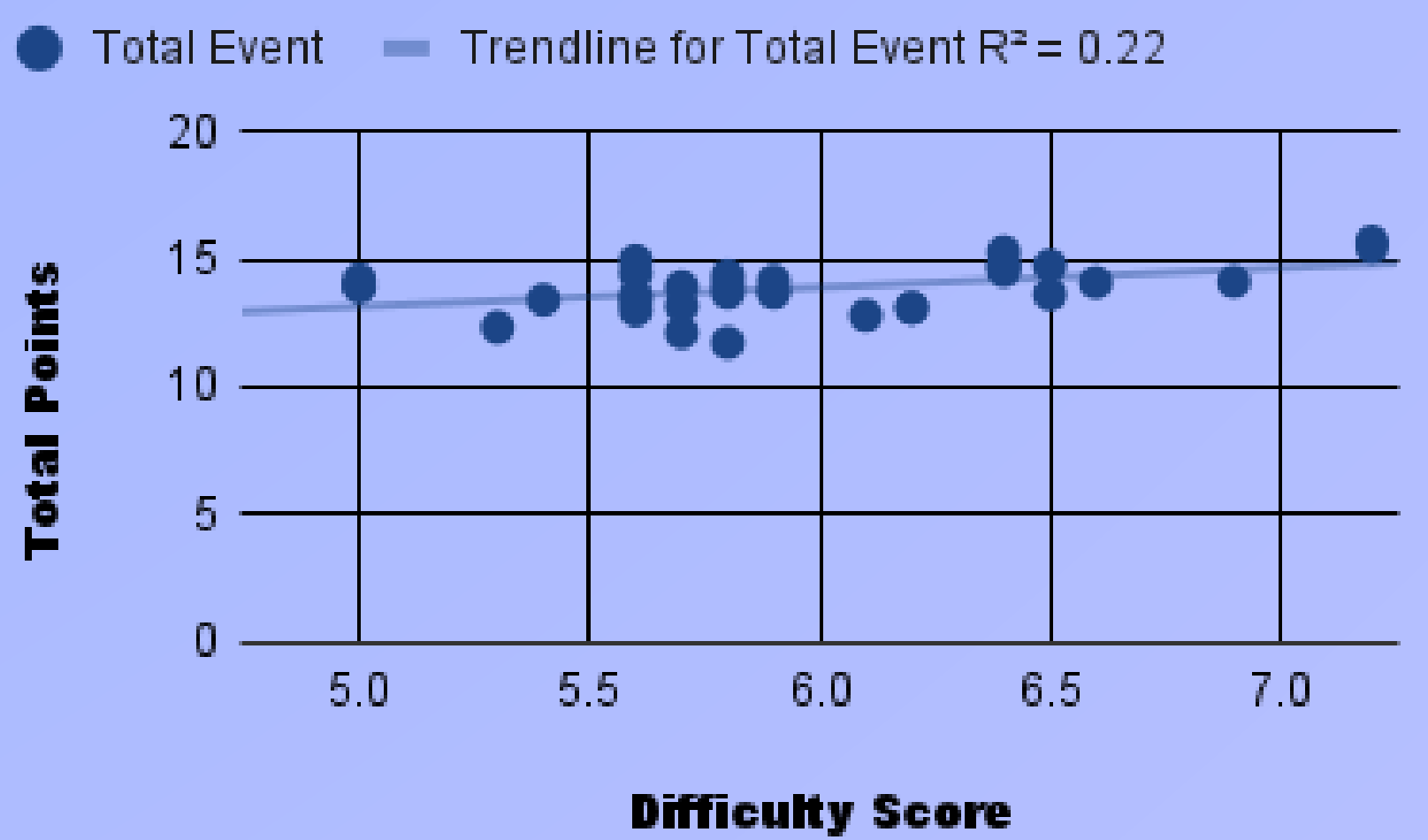
Simone Biles demonstrates that extremely high difficulty does not guarantee top scores when penalties occur. Her execution remains crucial, even with deductions, high-quality routines allow her to maintain medal-winning performance across all events.
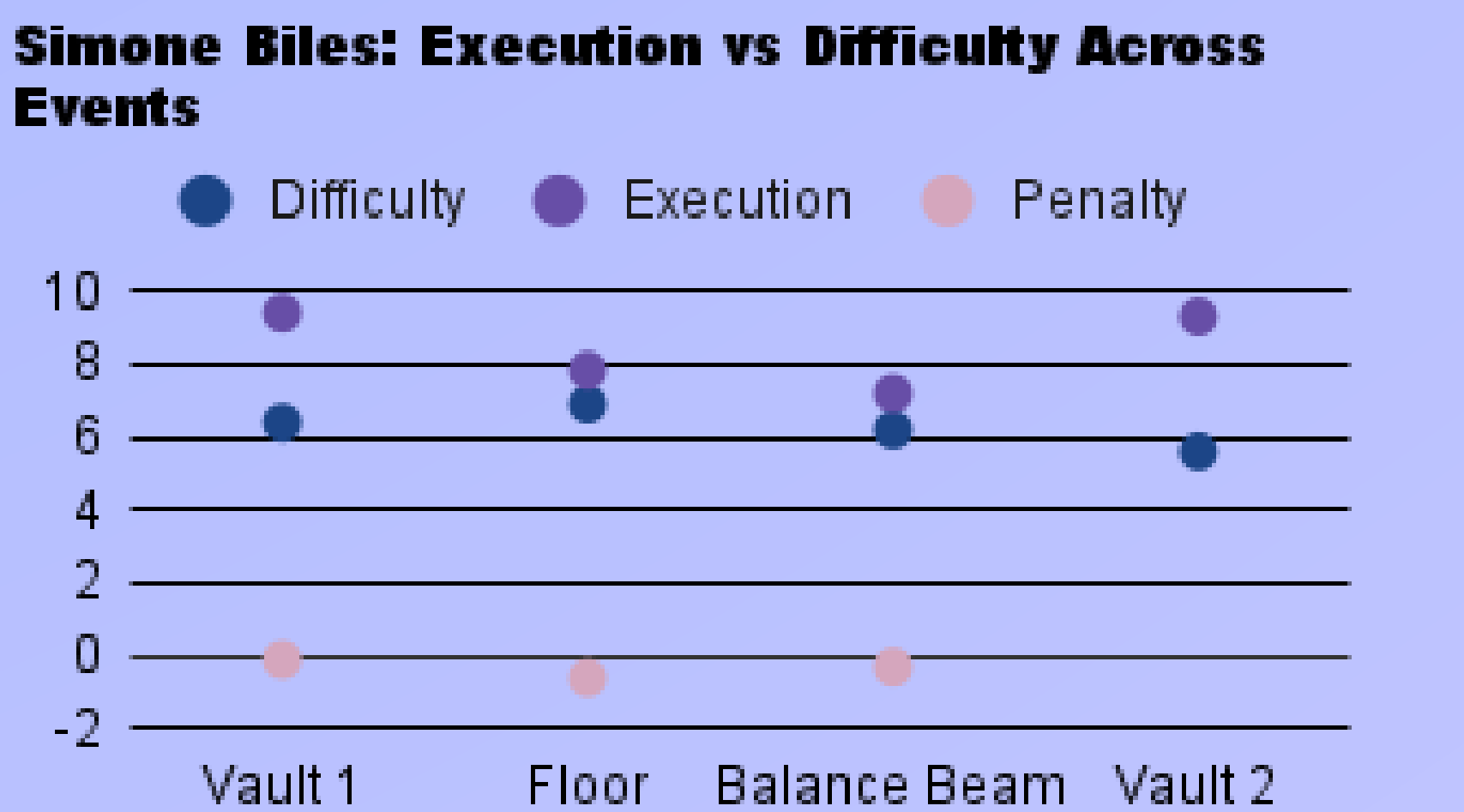
The pie chart clearly shows that execution (E) makes up a larger portion of Biles' total score than difficulty (D). This reinforces that while high difficulty sets the potential for a high score, it is execution that ultimately determines success.

Simone Biles: Execution vs Difficulty Across Events

Biles D vs E
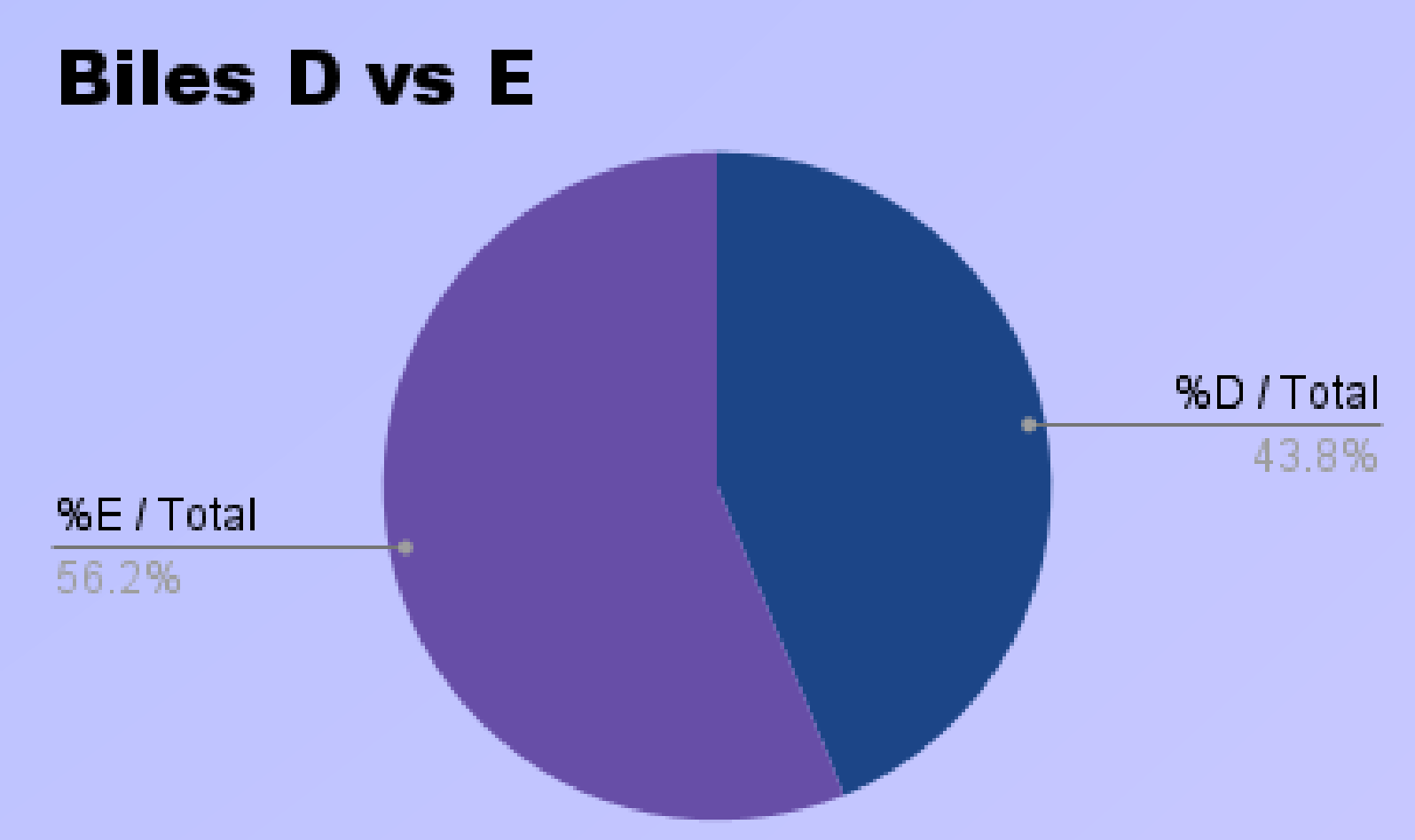
Interpretation

The results of my analysis show that execution scores (E) have a stronger impact on total performance in women's Olympic gymnastics than difficulty scores (D). This aligns with the hypothesis that while difficulty contributes to the potential maximum score, execution ultimately determines success.

The Simone Biles Effect
Simone Biles exemplifies this pattern. She consistently performs extremely high-difficulty routines, yet what sets her apart is her ability to execute them at an exceptional level. When comparing all other athletes, their difficulty scores are often similar, so competitions ultimately come down to execution. This demonstrates that while Biles can attempt "crazy" skills and still succeed, for most athletes, improving execution is the key differentiator.

Implications

These findings suggest that analysts, coaches, and gymnasts could prioritize execution quality over solely increasing difficulty when planning routines. Again with the Simone Biles example, her dominance shows that exceptional difficulty can provide an edge, but for the majority of gymnasts, success depends on performing slightly less difficult routines flawlessly. This supports my hypothesis, emphasizing E over D.

Limitations

I think my biggest limitation in this project is the small sample size, with only 30 gymnasts across 10 countries. However, because gymnastics is so consistent, and scoring follows such a set formula, it does not affect the overall result of my project. As with any sport, there are many confounding variables that aren't included. The dataset did not include factors such as injury history, experience level, or psychological factors that may influence performance.

Future Research

Future research could expand this analysis by incorporating a larger dataset, such as results from multiple Olympic cycles going years back where there were different rules, and much less advanced skills. I would also be interested in following one specific gymnast and showing how the adaptations to execution vs difficulty progresses. Additional variables could also be included, such as landing scores, execution deductions, event-specific difficulty composition, etc.

Additional Info

I was also interested in whether the order of competition influenced scores, thinking that early or late rotations might give athletes an advantage or disadvantage. However, the data showed no clear effect, scores were driven almost entirely by execution and difficulty.