Male gymnasts' performances during the Olympic cycle 2022-2024 through Visual Analytics

Giorgio Sisti - sisti.2134719@studenti.uniroma1.it

1 Introduction

Artistic Gymnastics competitions are characterized by complex performance evaluations, where each routine is assessed through multiple components such as Difficulty (D-score), Execution (E-score), and Final Score. Analyzing these performances across different competitions, years, and apparatuses is challenging due to the multidimensional nature of the data and the large number of athletes involved.

The goal of this project is to design an interactive Visual Analytics dashboard that supports the exploration and comparison of **Men's Artistic Gymnastics (MAG)** performances in World and European Championships from 2022 to 2024. The system aims to help users identify performance patterns, compare competitions and events, analyze consistency across apparatuses, and discover similarities between athletes through dimensionality reduction techniques. The dashboard integrates multiple coordinated visualizations and interactive filtering mechanisms, enabling both overview-level analysis and detailed inspection of individual performances.

2 Data Description

The dataset [4] contains individual performance records from World and European Championships held between 2022 and 2024.

Each record corresponds to a single athlete's routine and includes the following attributes:

- **Athlete**: name and surname of the athlete
- **Nation**: athlete's nationality
- **Event Date**: date of the competition in YYYY-MM-DD format
- **Competition**: World Championships or European Championships
- **Event**: official name of the competition
- **Year**: year in which the competition took place
- **Apparatus**: apparatus on which the routine was performed (FX, PH, SR, VT, PB, HB)
- Difficulty score (**D-score**): sum of the difficulty values of the performed elements, according to the Code of Points 2022–2024
- Execution score (**E-score**): score assigned by judges based on execution quality
- **Penalties**: deductions due to line faults or exercise construction errors
- **Final score**: sum of D-score, E-score, and penalties
- **Qualification** status: Qualified, Not Qualified, or Reserve
- **Rank**: qualification ranking among all athletes on the same apparatus

The dataset includes several thousand observations, allowing both statistical aggregation and fine-grained analysis.

Prior to the visual analysis, a preprocessing step was performed on the dataset. Records with incomplete information or with a significant number of missing values were

removed, as they could introduce noise or misleading aggregations in the visualizations. Additionally, attributes not relevant to the analytical goals of the project were discarded, retaining only features directly related to performance evaluation, competition context, and athlete identification.

The final AS value = 3622 tuples x 13 dimensions = 47086, bigger than the first one indicated due to the choice of adding two new dimensions and new tuples to the dataset.

3 Analytical Tasks

The visual analytics system is designed to support a set of analytical tasks aimed at exploring, comparing, and interpreting Men's Artistic Gymnastics performances across competitions, years, and apparatuses.

The intended users of the system include coaches, athletes, performance analysts, and gymnastics enthusiasts with an interest in competitive analysis. The dashboard is designed to be accessible to users with different levels of expertise, providing both high-level summaries for exploratory analysis and detailed views for expert inspection.

At an overview level, the system enables users to:

- Analyze average performance trends across years and competitions.
- Compare apparatus-level performance distributions.
- Identify differences between World and European Championships.

At an intermediate level, the dashboard supports comparative analysis by allowing users to:

- Compare Difficulty, Execution, and Final scores across athletes and events.

- Investigate how performance varies across apparatuses within the same competition or year.
- Assess consistency and variability of performances using aggregated views.

At a detailed level, the system allows users to:

- Inspect individual athlete performances and competition results.
- Analyze score distributions for specific subsets defined by filters or selections.
- Explore similarities between athletes by projecting multidimensional performance metrics into a lower-dimensional space using Principal Component Analysis (PCA).

Interactive mechanisms such as filtering, brushing, and selection enable users to dynamically refine the analysis, moving seamlessly from overview to details-on-demand and the coordination among multiple views ensures that interactions performed in one visualization are consistently reflected across the entire dashboard.

4 Visual Design and Interaction

The dashboard is designed following Visual Analytics principles, combining multiple coordinated views with interactive mechanisms to support progressive data exploration.

The layout is organized as a grid of four main visualizations, complemented by global filters and contextual details, ensuring a clear separation between overview, comparison, and detailed analysis tasks. A consistent visual encoding is adopted across all views.

- **Color** is used to distinguish competition types and qualification status, while position and scale are used to represent quantitative performance measures.

All visualizations share a coherent visual style to reduce cognitive load and facilitate cross-view comparisons.

The system provides a set of global filters (Year, Apparatus, Athlete, Nation, and Qualification status) that allow users to restrict the analysis to specific subsets of the data. These filters affect all visualizations simultaneously, enabling coordinated exploration across views. Interaction plays a central role in the dashboard.

- **Selections** performed in one visualization are propagated to the others, allowing users to analyze the same subset of data from different perspectives.
- **Hover interactions** provide details-on-demand through tooltips, while **click interactions** support explicit selection of athletes, events, or aggregated cells.
- **Brushing interactions** are used in the dimensionality reduction view to define subsets of athletes based on performance similarity. The brushed subset can be used to recompute the Principal Component Analysis, enabling interaction-driven analytics.

Additional contextual information is provided through lightweight informational elements integrated into the visual design. These elements offer concise explanations of each visualization without cluttering the interface, supporting interpretability for non-expert users.

Overall, the visual design and interaction model aim to balance expressiveness and simplicity, enabling users to move fluidly between high-level patterns and detailed performance analysis.

5 Plots description

This section describes the visualizations composing the dashboard and explains their role in supporting the analytical tasks outlined in the previous sections. Each plot is designed to highlight specific aspects of Men's Artistic Gymnastics performances, ranging from individual score components to aggregated trends and similarity analysis. The visualizations are tightly coordinated and share consistent visual encodings, allowing users to explore the data from multiple complementary perspectives.

5.1 D-Score vs E-Score Scatter plot

The D-score vs E-score scatter plot provides a two-dimensional overview of individual performances by mapping Difficulty score on the x-axis and Execution score on the y-axis. Each point represents a single athlete's routine, allowing users to visually assess the trade-off between difficulty and execution quality.

Color encoding is used to represent the qualification status (Qualified, Not Qualified, Reserve), enabling immediate identification of performance differences among qualification outcomes.

This view supports hover interactions to display detailed information about the athlete and the routine, while click-based selection allows users to focus on a specific athlete and propagate the selection to all coordinated views.

The scatter plot is particularly useful for identifying clusters of high-difficulty or high-execution performances, as well as outliers that deviate from typical scoring patterns.
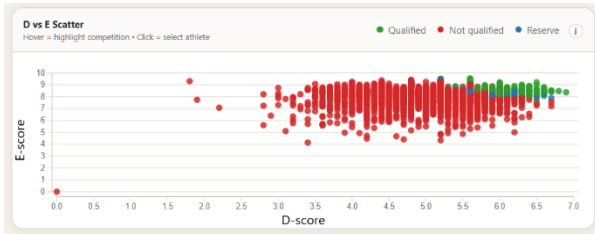
Figure 1: D-score vs E-score Scatter Plot

Figure 2: Performance Trend View (multiple years)

Figure 3: Dot Interval View (single year)

5.2 Performance Trend / Dot-interval View

The performance trend view visualizes Final Score variations across competitions and years.

- When multiple years are selected, the visualization adopts a dot–interval representation, where each event is represented by a central point indicating the mean Final Score and an interval showing the minimum and maximum values.
- When a single year is selected, the view switches to an event-level comparison, where performances from different competitions are displayed side by side. In this case, the dot–interval representation is preserved to highlight differences in central tendency and variability across events within the same year, rather than across time.

This adaptive design allows users to assess both temporal trends and event-level comparisons using consistent visual encoding, while color encoding distinguishes competition types, ensuring visual consistency with other views.

Hover interactions provide detailed statistics for each event, while selections performed in other views dynamically update the displayed data, enabling focused and context-aware analysis.
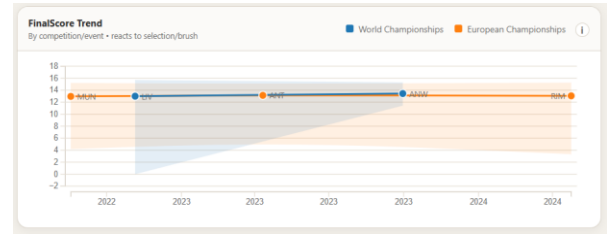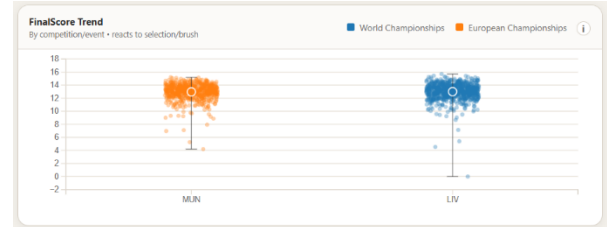
5.3 Apparatus x Year/Event Heatmap

The heatmap provides an aggregated overview of performance across apparatuses and competitions.

Rows represent apparatuses, while columns correspond to years and events, so each cell encodes the average Final Score for the corresponding apparatus–event combination using a sequential color scale. This visualization supports rapid identification of strong or weak apparatus-specific performances across competitions. Missing or unavailable aggregations are explicitly represented using a neutral color to avoid misleading interpretations.

Users can interact with the heatmap by hovering over individual cells to inspect numerical values and by clicking on a cell to filter the dashboard according to the selected apparatus and event, enabling seamless transition from overview to focused analysis.
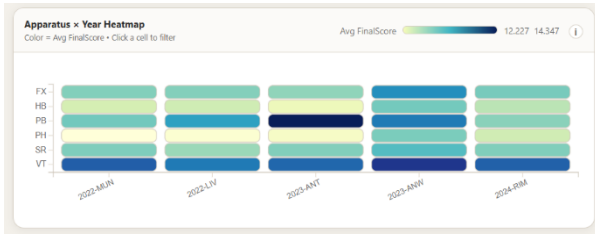
Figure 4: Apparatus x Year/Event Heatmap

5.4 Principal Component Analysis (PCA)

The Principal Component Analysis (PCA) view supports the exploration of similarities between athletes by projecting multidimensional performance metrics into a two-dimensional space.

The PCA is computed using Difficulty score, Execution score, Final score, and Penalties, allowing users to analyze overall performance profiles rather than individual metrics. Each point represents an athlete's performance instance, and the spatial proximity between points reflects similarity in the underlying performance features.

Brushing interactions allow users to define subsets of points based on their position in the projection space. Optionally, the PCA can be recomputed on the brushed subset to enable interaction-driven analysis of specific performance groups.

To ensure stability and interpretability, PCA recomputation is only performed when a valid subset of data is selected. Additional details, including explained variance and component loadings, are provided through an on-demand information panel integrated into the visualization.
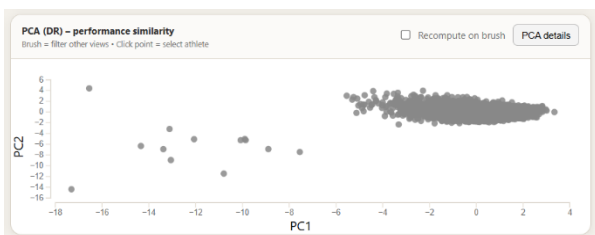

Figure 5: Principal Component Analysis (PCA)

6 Related Works

The use of visual analytics techniques for sports performance analysis has been widely investigated in domains such as football, basketball, and athletics, where rich spatio-temporal data and tracking information are available.

In contrast, relatively few academic studies explicitly focus on artistic gymnastics, and in particular on the visual analysis of competitive performance scores.

Most existing works related to gymnastics analytics primarily address biomechanical analysis [2, 3], injury prevention, or training optimization, often relying on numerical models or video-based motion analysis rather than interactive visual exploration of competition results. There is a lack of published visual analytics systems specifically designed to support the exploration and comparison of gymnastics competition scores across events, apparatuses, and years. Despite the limited presence in academic literature, the usefulness of performance analytics in gymnastics is widely acknowledged in practice.

National federations and professional organizations, such as British Gymnastics [1], emphasize the importance of systematic performance analysis to support athletes' development pathways, talent identification, and competition preparation. These approaches, however, are often implemented through internal tools or descriptive reports, rather than interactive, multi-view visual analytics systems.

This gap between practical needs and available academic solutions motivates the design of this dashboard.

The work aims to explore how coordinated visualizations, interactive filtering, and

dimensionality reduction techniques can support the analysis of Men's Artistic Gymnastics performances, providing an initial contribution toward structured visual analytics in this domain.

7 Insights

The interactive exploration supported by the dashboard enables the identification of several qualitative insights regarding Men's Artistic Gymnastics performances across competitions, years, and apparatuses.

At an aggregated level, the coordinated views reveal clear differences between World and European Championships. World Championship events generally exhibit higher average Final Scores and a wider performance variability, reflecting both a higher competitive level and a broader range of participating athletes. European Championships, in contrast, tend to show more compact score distributions, suggesting greater homogeneity among competitors.

The apparatus-level analysis highlights that performance variability is not uniform across events. Some apparatuses, such as Floor Exercise and High Bar, display larger score ranges and stronger dispersion, while others, such as Pommel Horse, show more concentrated score distributions. This suggests differing levels of technical risk and execution consistency depending on the apparatus.

The D vs E scatter plot allows users to observe distinct performance strategies adopted by athletes. Some competitors achieve high Final Scores primarily through high Difficulty values, while others rely on cleaner Execution scores. The visualization makes these trade-offs immediately apparent and supports comparisons between qualification statuses, revealing that qualified athletes tend to cluster in regions combining both high Difficulty and high Execution.

Temporal analysis through the performance trend view enables the inspection of score stability across years and competitions. When multiple years are selected, the dot–interval representation reveals how performance consistency varies across events, highlighting competitions with stable outcomes versus those characterized by larger fluctuations. In the single-year case, the event-level comparison facilitates direct inspection of relative performance across competitions within the same season.

The PCA view provides a complementary perspective by projecting multidimensional performance metrics into a low-dimensional space. Athletes with similar overall performance profiles tend to appear close to each other, allowing the identification of groups with comparable scoring patterns. Although the PCA is not intended for precise clustering, it supports exploratory analysis and hypothesis generation regarding similarities between athletes across different apparatuses and competitions.

Overall, the integration of multiple coordinated views and interactive mechanisms enables users to move fluidly from overview to details-on-demand, uncovering patterns that would be difficult to detect using tabular data or static visualizations alone.

8 Conclusion and future work

This project presented an interactive visual analytics dashboard for exploring Men's Artistic Gymnastics performance data from World and European Championships between 2022 and 2024.

Future work could extend the analysis to additional years or competition phases, enrich the dataset with routine-level or biomechanical features, and integrate more advanced analytical methods to complement the PCA view.

A user-centered evaluation involving coaches or analysts could further assess the effectiveness and usability of the system.

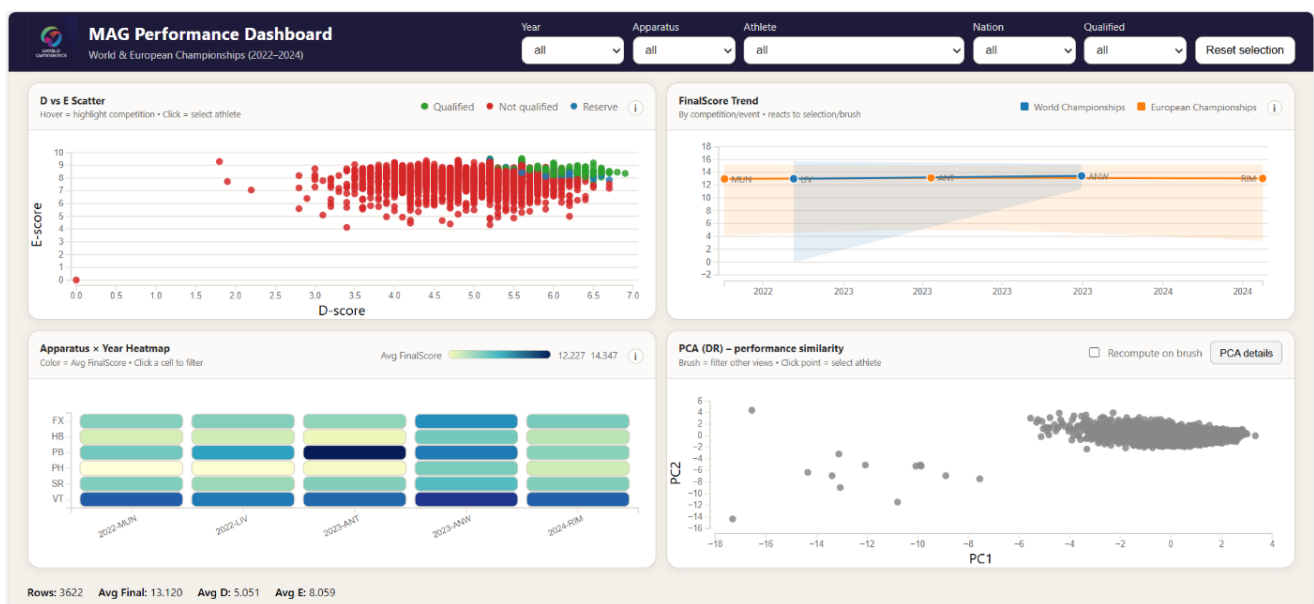Vision and Image Understanding, Volume 159, 2017, Pages 154-163, ISSN 1077-3142, https://doi.org/10.1016/j.cviu.2016.11.006.

[4] World Gymnastics - Results https://www.gymnastics.sport/site/events/searchresults.php#filter

Figure 6: Dashboard Overview

References

[1] Performance Analysis with Gymnastics –

UK Sports Institute https://uksportsinstitute.co.uk/resource/performance-analysis-with-gymnastics/

[2] Zhu, D., Zhang, Z., Chen, M. *et al.* A Perspective on Rhythmic Gymnastics Performance Analysis Powered by Intelligent Fabric. *Adv. Fiber Mater.* **5**, 1–11 (2023). https://doi.org/10.1007/s42765-022-00197-w

[3] Brian Reily, Hao Zhang, William Hoff, Real-time gymnast detection and performance analysis with a portable 3D camera, Computer