

Olympic Swimming Medals and Participation

Morgan Oscar Morel
UMSI HCI/PI
morelm@umich.edu

Rachael Leduc
UMSI IAR/HCI
rleduc@umich.edu

Anand Geetey
UMSI HCI
ageetey@umich.edu

Greg Cunningham
UMSI HCI
gcunning@umich.edu

The visualization is hosted at <http://www.gregory-cunningham.com/OlympicViz/OlympicViz.html>.

I. INTRODUCTION

The Olympic Games are an international sporting event in which every two years thousands of athletes from across the globe participate in various events. The summer and winter games alternate, which means that each seasonal Olympic happens every four years. The three individuals or teams that in each event are awarded medals. First place receives gold, second place receives silver, and third place receives bronze.

Since the first Modern Olympic Games, held in Greece in 1896, there have been 30 Summer Olympic Games held. Throughout this time games have grown in both the number of events and the number of participating countries. Data regarding which countries won which events in which years is freely available on the internet [1-28], and in fact there was a contest held in 2012 that offered a cash prize to the best visualization of this data. However, upon looking at the top three visualization contestants we found they all left something to be desired [29, 30, 31]. One reason for this is simply that data set is so vast it's incredibly difficult to get a detailed view on the data.

With this in mind, we decided to focus the scope of our project on only a subset of the data. It seemed to be far too big a feat to try and design a visualization of all the Olympic medal data in a way that is both elegant and informative, and we decided to focus only on the medals won in swimming in every Olympic Games since 1986. We chose swimming because it is one of the events that has happened in every year of the Olympic Games since the beginning. This offered the opportunity for us to see medaling and participation trends in throughout the entire existence of the modern games. Another reason we chose swimming is because the amount of events has changed throughout the years, though the total number of events is never incredibly high. This gave us the opportunity to see whether the distribution of medals across countries changes as more events are created, and made sure we would never have numbers too large to deal with.

Once we decided on the data we would be using, we began the design of our visualization by developing guiding questions about the kind of information that would be shown and the kind of relationships that it would enable. The first question we asked was *What are the trends in participation and medaling for swimming throughout the history of the Olympic Games?*

This question was very broad, but provided a context through which we decided to focus on temporal relationships. In addition, taking both medaling and participation into consideration meant that we would need to find a way to distinguish between not medaling and not participating. Next, we asked *Are there any trends in participation and medaling for specific time periods?* While somewhat similar to the first, in order to answer this question properly we needed to design a visualization that can accommodate time ranges, rather than discrete dates. While our final design included this feature, we were actually unable to implement it in the final application.

II. RELATED WORKS

As was previously mentioned, we found the winners of the 2012 London Olympics visualization somewhat lacking in the information they made available to the user. All of the top three visualizations only allowed the user to see medal data from a single year. While Christian Gross's visualization [29] allows the user to pick the data on a slider through interactivity, there is no way to easily see a country's performance in aggregate across a range of years. However, Gross's use of interaction helps allowed him to make the visual design elegant and uncluttered, while allowing the user to find further details if they so wish.

Ivo Afonso's visualization [30] is static, though the use of stacked bar charts elegantly illustrate the total medal count, along with the breakdown of gold, silver and bronze. In fact, Afonso's visualization demonstrates that users can easily decode the meaning of these types of bar charts in the context of Olympic medals, due to the fact that most users have enough background about the workings of the medal system to understand it as long as the gold, silver and bronze colors are clearly displayed. As with Gross's visualization however, Afonso's does not offer any aggregated data about a country's medal count over a range of years.

Another visualization that informed our design decisions was General Electric's World of Records visualization [31]. It visualizes athletic world records for summer games, and offers a number of interactions and views. The user can choose to view the data as a bar chart whose x-axis is time. In this view, the bars are made up of individual dots whose color indicates the sport the record was broken in. When hovering over a dot, the user is shown a dialog that provides more detailed information about that specific record. The user can also view the records by sport or by geographical location of the athlete, and there is a view in which the user can see how long the records went

before being broken. All of these views can then be filtered by gender, and by whether the record-breaking event took place during the Olympics. The filtering is visualized by dimming the events that do not apply to the specified filter. This allows the user to see the desired data in the context of the overall trends. By making the bars out of individual dots, this visualization is able to show aggregated trend throughout years, while simultaneously supporting more detailed inquiries about individual data points.

III. SYSTEM DESCRIPTION

Our visualization was made using Protovis to display the data and HTML/CSS to style the page. Since our group had extremely little knowledge of Protovis, our ultimate visualization was limited in comparison to our design sketches.

A. Line Graph

The line graph (Fig. 1) on the top left of the screen shows total medal counts over time for selected countries. Countries are selected via drop down lists on the left, and each country is represented by a different colored line on the graph. Up to four countries can be selected at once in order to reduce confusion that too many lines might cause. To differentiate countries that have participated and did not medal from the countries that did not participate for that year, we have decided to draw the line for participation data slightly above zero, while non-participation data will be charted at zero. From the dropdown menu on the left, users can select a single year, which is marked by a vertical grey bar in the line chart. This selection also adjusts the other visualization components -- the bar chart and the pie chart.

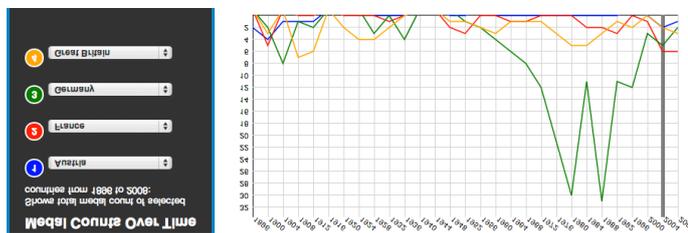


Fig. 1. Line graph showing medal counts over time for selected countries. Bar indicates selected year.

B. Bar Chart

The bar chart (Fig. 2) on the bottom left of the screen reflects the time selection from the line graph. For each country that medaled during that year or range of years, the bar shows a stack of medal counts--bronze stacked on top of silver, silver on top of gold. This way, the user can simultaneously see the total medal count as well as the breakdown of each medal type. Instead of a traditional solid bar, the bars are split into box units that represent medals. If the user would like to compare countries for a specific medal type (eg. just gold), the data can be further filtered on the side by selecting/unselecting each of the three medal types to appear on the chart. The bars can also

be sorted according to total medals, total gold, total silver, or total bronze.

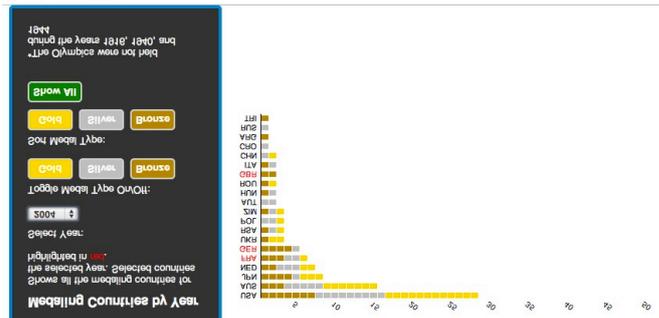


Fig. 2. Bar chart showing medal counts for selected year. Countries in red indicate countries displayed in line graph. Filter and sort controls on the left panel.

C. Pie Chart

The pie chart (Fig. 3) on the right of the page also reflects the years selection shown in the line graph, giving overall participation information for that period. The three wedges represent a count of the countries that medaled, that participated but did not medal, and that did not participate. The counts are printed within the pie chart itself, and the legend indicates which group the colored wedges represent. This chart is how users will be able to see the popularity (in terms of participating countries) of the sport in a given year.

- ◆ Medaled
- ◆ Participated
- (did not medal)
- ◆ Did not participate

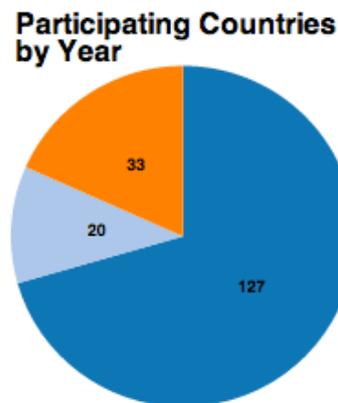


Fig. 3. Pie chart showing count of countries that medaled, participated, or did not participated for selected year.

IV. DISCUSSION

A. Unable to Implement

As mentioned earlier, our visualization was limited due to our ability in Protovis. We would have liked to offer the level of interactivity that was shown in our sketch (Fig. 4). One major component that we are missing is the ability to select a range of years instead of just one year for all visualizations. We ran into difficulty with this because the data showed the total

medal/participation count for each year, and it was not broken down enough. In our design, if a range of years is selected, the country is included in the medaling count if they have ever medaled in that range, or in the participation count if they have ever participated at any point in the range without medaling, even if it was only one year. However, this did not work because for example, if 8 countries participated one year, and ten countries participated the next year, it was not known if all 8 countries were part of the ten countries. We were disappointed that we could not support range selection since it would have been an improvement on the related works discussed previously.

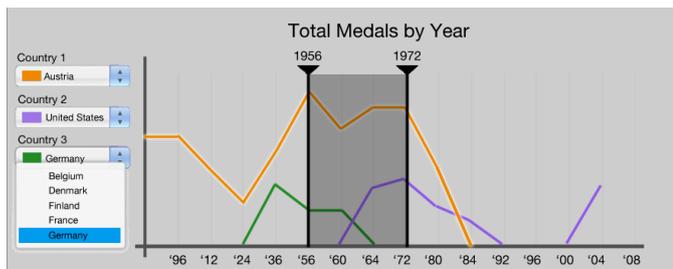


Fig. 4. Line chart design with range selector.

We also had difficulty implementing clicking interaction, so we limited ourselves to form elements. Therefore we were unable to make the lines in the line graph clickable. In our design we intended users to be able to click a specific line, which highlights the line to stand out from the other lines in order to allow for users to distinguish lines that fall at the same data points.

We split the bars into box units that represent medals. However, we wanted to go a step further in our design to show one box for each medal that could have been won during the selected time period, but only the boxes that represent medals won would be colored in (Fig. 5). Although we were not able to do this, the purpose was to normalize the comparison since the medals earned will be shown in relation to the medals available.

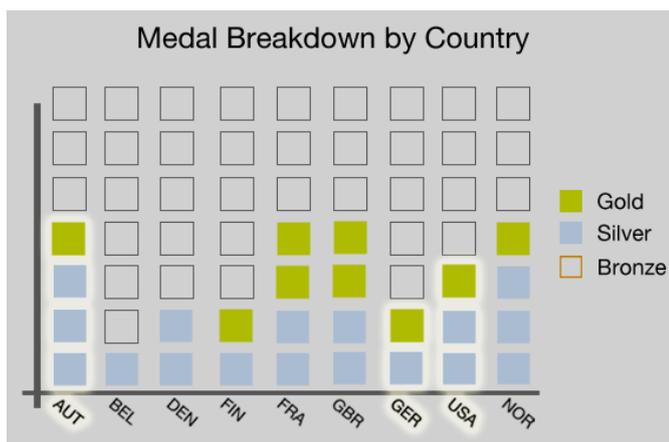


Fig. 5. Bar chart with total possible medals.

B. Other Design Decisions

We were motivated to look for limitations in our data by Fishkin and Stone’s work with dynamic queries. Their visualization took missing data into consideration, encoding those points in a distinct way [32]. Missing data can give just as much information as traditional data points can. One design decision we made from the beginning was to somehow make a distinction between countries that participated and did not medal vs. countries that did not participate in a given year, even though the medal count for both would be zero. As mentioned earlier, we decided to show those who participated but did not medal as slightly above zero on the line graph.

We were concerned that the lines in the line graph would be too jumbled or would overlap too much. First we considered small multiples to visualize individual lines, and then we decided to allow for the interactive scrubbing of a date or date range in the line graph. This was informed by a paper by Müller and Schumann, which states that “if the main task of a user is to compare multiple time-dependent variables, it makes sense to provide interaction techniques that allow navigating the time axis or brushing certain data values in different views” [33]. Our planned interactions were designed with the intent of allowing the user to have freedom to collect and view data or specific areas of the time. While we were unable to fully integrate this type of interaction, we used the drop-down date selection to a similar means, giving the user the ability to explore detailed views of time.

We also found out during the round robin critique that some people were worried about using a line graph for non-continuous data. With this in mind we considered perhaps using a scatter plot instead, as the discrete characteristics of the dots reflects the discrete nature of the data. However, we came to the conclusion that a scatterplot would be too difficult for users to make comparisons, and the interval suggested by a line graph would visualize the medaling trends of countries. We also looked at the General Electric visualization, but we decided the connected bars would be too complicated to implement. Ultimately, we decided to keep the line graph since it nicely showed trends in medal earning over time.

Upon finding out that the highest medal count for one country in one year was almost double what our y-axis maximum was for the bar chart, we decided to rotate the chart. Transitioning from a vertical bar chart to a horizontal chart allowed for much more room for expansion across the page, as well as limiting the effect of so much white-space.

Unfortunately, we did not have time to do any user testing at all. We would be very interested in seeing the results, especially in how users understand the navigation and flow of the visualization.

V. FUTURE WORK

One relationship we were unable to visualize was that of the breakdown of medals for a single country for different dates. In order to support this we would have needed to allow the user to chose two distinct dates, as well as chose two distinct date ranges, which would have greatly complicated the interactions available. Furthermore, we had difficulty

developing a way to display the medal breakdown of two different time periods on the same bar chart. We eventually decided that if we visualized this, it would have been by using grouped bar charts, where the bars are still grouped by country. In order to tell which bar represents which time range, we would use a visual indicator, such as matching the color of the markers used for different ranges to a color marker for the bars.

Another feature we left out was giving the user more detailed information about the medals. By separating the medal breakdown into discrete chunks, like what was done in the World of Records visualization, we were given the opportunity to give the user detailed info about whichever medal they hovered over. This information could include many different things, such as the name of the athlete, the name of the event, the time, and whether or not a world record was broken. If we decide to work further on this visualization, this is a feature we would like to include.

In the future, we think the most exciting application would be to scale the project up by supporting data from all Olympic sports, not just swimming. We envision a dropdown list at the top of the page in which the user could select a sport, and the three visualization components would change to reflect that sport.

REFERENCES

- [1] S. Rogers. (2012, August 13). "If Michael Phelps were a country, how would his medal haul compare?" *The Guardian*.
<http://www.guardian.co.uk/sport/datablog/2012/aug/01/if-michael-phelps-were-a-country#data> (see end of article for data file)
- [2] http://en.wikipedia.org/wiki/Swimming_at_the_1896_Summer_Olympics
- [3] http://en.wikipedia.org/wiki/Swimming_at_the_1900_Summer_Olympics
- [4] http://en.wikipedia.org/wiki/Swimming_at_the_1904_Summer_Olympics
- [5] http://en.wikipedia.org/wiki/Swimming_at_the_1908_Summer_Olympics
- [6] http://en.wikipedia.org/wiki/Swimming_at_the_1912_Summer_Olympics
- [7] http://en.wikipedia.org/wiki/Swimming_at_the_1916_Summer_Olympics
- [8] http://en.wikipedia.org/wiki/Swimming_at_the_1920_Summer_Olympics
- [9] http://en.wikipedia.org/wiki/Swimming_at_the_1924_Summer_Olympics
- [10] http://en.wikipedia.org/wiki/Swimming_at_the_1928_Summer_Olympics
- [11] <http://la84foundation.org/6oic/OfficialReports/1932/1932spart3.pdf>
- [12] <http://la84foundation.org/6oic/OfficialReports/1936/1936spart5.pdf>
- [13] <http://www.swimrankings.net/index.php?page=meetDetail&meetId=562163>
- [14] <http://www.swimrankings.net/index.php?page=meetDetail&meetId=562162>
- [15] <http://www.swimrankings.net/index.php?page=meetDetail&meetId=562161>
- [16] <http://www.swimrankings.net/index.php?page=meetDetail&meetId=562160>
- [17] <http://www.swimrankings.net/index.php?page=meetDetail&meetId=562159>
- [18] <http://www.swimrankings.net/index.php?page=meetDetail&meetId=562158>
- [19] <http://www.swimrankings.net/index.php?page=meetDetail&meetId=562157>
- [20] <http://www.swimrankings.net/index.php?page=meetDetail&meetId=562033>
- [21] <http://www.swimrankings.net/index.php?page=meetDetail&meetId=561955>
- [22] <http://www.swimrankings.net/index.php?page=meetDetail&meetId=561954>
- [23] <http://www.swimrankings.net/index.php?page=meetDetail&meetId=561947>
- [24] <http://www.swimrankings.net/index.php?page=meetDetail&meetId=542807>
- [25] <http://www.swimrankings.net/index.php?page=meetDetail&meetId=524072>
- [26] <http://www.swimrankings.net/index.php?page=meetDetail&meetId=515568>
- [27] <http://www.swimrankings.net/index.php?page=meetDetail&meetId=503656>
- [28] <http://www.swimrankings.net/index.php?page=meetDetail&meetId=524069>
- [29] C. Gross. "The Rising of Olympic Mountains."
<http://visualizing.org/visualizations/rising-olympic-mountains>
- [30] I. Afonso. "Olympic Medal Winners."
<http://visualizing.org/full-screen/41351>
- [31] General Electric. "World of Records."
<http://visualization.geblogs.com/visualization/records/>
- [32] Fishkin + Stone, "Enhanced Dynamic Queries via Movable Filters," CHI'95.
- [33] W. Aigner, S. Miksch, W. Muller, H. Schumann, and C. Tominski, "Visual Methods for Analyzing Time-Oriented Data," IEEE Trans. on Visualization and Computer Graphics, Vol. 14, No.1, Jan-Feb. 2009, pp. 47-60.