Introduction
This project seeks to gain insight into how various factors, ranging from economic statistics to host country, may affect the performance of a nation’s Olympic athletes. Although the majority of existing analyses judge performance by the total number of medals earned, this study will explore another metric of success.

Data Used

Olympics Data
This database provides location, medal and participation data by country, event, and year for the winter (1924-2006) and summer (1896-2008) Olympics.


World Economic Outlook
This source contains economic data for 184 countries (1980-2012), including Gross Domestic Product, national debts, revenues, and balances, import/export statistics, employment, and population. This data provided by The International Monetary Fund.


Data Statistics

- Total Rows: 3132
- Total Columns: 8
  - Country
  - Year of Games
  - Season of Games
  - Host Country
  - Number of Participants
  - Number of Medals Won
  - Population
  - GDP
Findings

The effect of a country's GDP on its participation/performance in the Olympics is negligible.

Is there a correlation between a country’s GDP and its performance at the Olympic Games? GDP and participation? Participation and performance? Surprisingly, it seems that GDP has little effect on these metrics. Figure 1 shows a parallel coordinate chart showing possible correlations between these three parameters. For GDP to correlate strongly with participation, each line between the two should be roughly horizontal, but that is clearly not the case here as many lines cross over each other.

However, it can be seen that there is a relatively strong correlation between the number of participating athletes from a country and the number of medals won. While not particularly surprising, it is nice to see such a definite confirmation of this in the data.

As further evidence, Figure 2 shows a heat map of the same data, sorted from top to bottom by highest participation. One can again see a strong correlation between participation and performance as the gradient of the number of participants is well-matched by that of the medal counts. The lack of correlation between these values and GDP is evidenced as well, with the heat map for GDP appearing to be random.
World rankings depend on point-of-view.

The New York Times visualization (http://london2012.nytimes.com/results) of Olympic performance directly compares countries by the number of medals won. Would normalizing medal counts by the number of participating athletes yield interesting insights?

Figure 3 is a side-by-side comparison of the total number of medals won by each country and the normalized medal count (i.e. number of medals won divided by participating athletes) for the 2012 Olympic Games. The countries on the X-axis are sorted according to the number of medals won. This is made obvious by the downward trend as one views the medal counts from left to right. However, the normalized medal counts do not follow this same trend. This implies that those countries which earn many medals do not necessarily have stronger athletes but may be winning through sheer numbers.

For example, the un-normalized plot ranks the United States, Russia, Great Britain, and China very highly in terms of medal counts. However, when normalized (medals/participants), Jamaica, the United States, the Netherlands, and Montenegro come out on top, having the highest ratio of medals per athlete.

Thus, depending on one's viewpoint, either Russia or Jamaica are among the top contending countries. After all, Russia has been very successful in earning medals in general while Jamaica has done an excellent job of making more with less. Regardless, it seems that the United States is a true world power when it comes to the Olympic Games.
Olympic hosts do not have "home court advantage."

Does being the host country affect performance?

At first glance, it would appear that being the host country has a great effect on the number of medals earned. Take, for example, Figure 4 below. This figure is a scatter plot of medal counts for each year. Each marker represents a country, with the host country highlighted by color and shape. Here it can be seen that the host countries tend to do much better than the other countries competing in each Olympic Games. This effect is even more pronounced than is evidenced in this chart, as each country that shares the same number of medals won overlaps, hiding a greater density below the host markers. Given all of this, it would seem that the host country enjoys a significant advantage in athletic prowess.

Figure 3. Comparison of raw medal counts to normalized medal counts.

Figure 4. Comparison of raw medal counts between the host country and all competing countries.
This, however, is not the entire story. The issue is similar to the one outlined above, namely that a country's performance may be better measured by the percentage of its athletes earning medals rather than by the total number of medals it has earned. Figure 5 normalizes the medal counts by the number of participants. One can see that the supposed advantage of the host countries all but disappears when this is taken into account. Thus, it would seem that host countries do earn more medals, but it is more likely due to the increased participation than by an increase in actual athletic performance. As described above, there is a strong correlation between participation and medals earned.

Finally, take Figure 6, which shows a definitive increase in participation for the host country at any particular Olympic Games, lending credence to the hypothesis that more medals are being earned by the host country simply because more athletes are being fielded.
World events have visible effect on Olympic Games.

Do certain world events create noticeable effects in the data? For example, the lack of Olympic Games due to world wars leaves an obvious gap in the data. Can other such historical markers be found without *a priori* knowledge of the events?

Figure 7 below graphs the number of participating athletes for any given Olympic Games. While a steady increase in participants is not unexpected (due to increasing worldwide prosperity, popularity of the games, etc.), notice that there are several prominent changes in participation. Many of these changes can be linked to world events:

- **1916** - The games are cancelled due to WWI.
- **1932** - Great Depression - "Held during the worldwide Great Depression, many nations and athletes were unable to pay for the trip to Los Angeles. Fewer than half returned to compete."[1] Additionally, LA was the only city which bid to host the games this year.
- **1936-1948** - Again the games are cancelled for war (WWII).
- **1956** - The first games to be held outside of Europe/North America as well as the first games in southern hemisphere...
- **1976** - Most African nations and a few other countries boycotted due to the International Olympic Committee's refusal to ban New Zealand after their mixed-race rugby team toured apartheid governed South Africa.
- **1980** - America leads a boycott in protest against the Soviet war in Afghanistan.
- **1992** - As the Cold War comes to an end, this is the first Olympic Games since 1972 that is free of boycotts.
Figure 7. Total participation (number of athletes) in each Olympic Games.

Spotfire Evaluation

Pros:

- Color schemes are maintained between pages, allowing for consistent presentation.
- Filter options are maintained between pages, allowing one to switch between multiple coordinated, full-screen views.
- The ability to quickly switch between saved filters. Also, the ability to assign a particular visualization (not just a page) to a filtering scheme, so that it never changes.

Cons:

- Can’t pull in data from multiple sources for basic chart types (e.g., size by data from one table, color from another). This is, however, possible with map data for some reason.
- When data is dense, crowding becomes an issue as there is no option for ordering of markers. The ability to ensure that certain data points are on top of others could be useful here.