The Sky is the Limit

In this assignment I will be building a model for the relationship between the winning heights in men's high jump and the years that they took place.

The high jump event in the Olympics is a track and field athletics event. It is held every four years in the summer Olympics. In the event competitors must jump over a horizontal bar that is placed at various heights. The high jump has existed for centuries now and was popular in ancient Greece. Javier Sotomayor holds the current world record for men's high jump with a jump of 245 centimetres.

The table below gives the height (in centimetres) achieved by the gold medalists at various Olympic Games. Note: the Olympics were not held in 1940 and 1944 due to World War II.

Year	1932	1936	1948	1952	1956	1960	1964	1968	1972	1976	1980
Height (cm)	197	203	198	204	212	216	218	224	223	225	236

The independent variable is time; so let t years be the time. The dependant variable is height; so let h centimeters be the height. It is important to note that height cannot be negative as it is physically impossible to have a jump that is below 0 centimetres.

A constraint of plotting this data is that there are two missing points for the years 1940 and 1944, as there were no Olympics during these years. This makes the data slightly more difficult to analyze and could cause some inaccuracies later while plotting a function. Another constraint of the task is that there can easily be outliers in the data simply because of whoever is competing in the Olympics that year. A final constraint of this task is that there is a limited amount of data provided which could once again cause some inaccuracies.

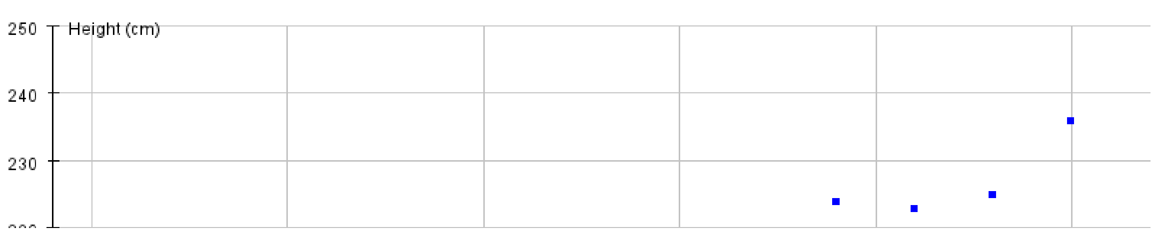
In this assignment I will be using two functions: the linear function and the square root function. I will define the parameters of each.

For the linear equation: $h = mt + b$
h is the winning height in centimetres, m is the slope of the function, t is the time in years, and b is the y-intercept.

For the square root function: $h = \sqrt{t}$
h is the winning height in centimetres and t is the time in years.

All calculations are rounded to the nearest number.

The graph below shows the two variables with the data plotted.



A linear function fits the graph well as it is effective in including the majority of the points. Other functions are likely to not pass through the first two years. Also this function increases at an appropriate amount. The heights of the high jumps should not increase or decrease significantly at each four-year interval similar to many other events in the Olympics. Although the graph may not continue to increase linearly in the future when the records become similar, it is effective in showing the steady increase during the past century.

Since this is a linear function the equation to model the data will be in the form of: $y = mx + b$. To find the gradient of this line use two points on the graph. I chose point number one at 1932, and point number nine at 1972 as these two points will fit the predicted line well.
So:

$$\frac{223\quad 197}{1972-1932}=\frac{13}{20}$$

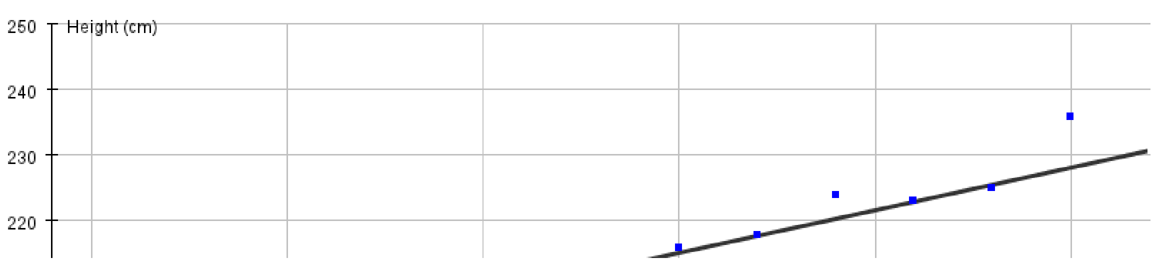
Now substitute values in for both x and y to find the equation of the line:

$$(197)=\frac{13}{20}(1932)+b$$

$$-1059=b$$

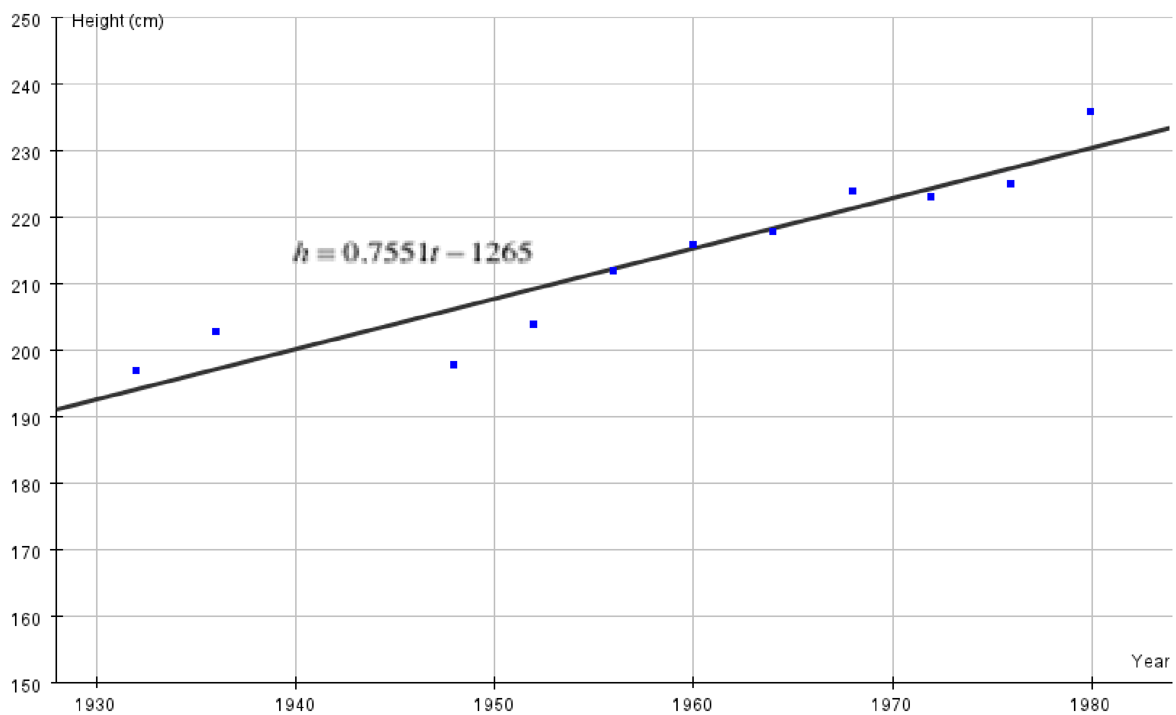
$$h=\frac{13}{20}t-1059$$

The graph below shows the function plotted to fit the original data. It fits fairly well.



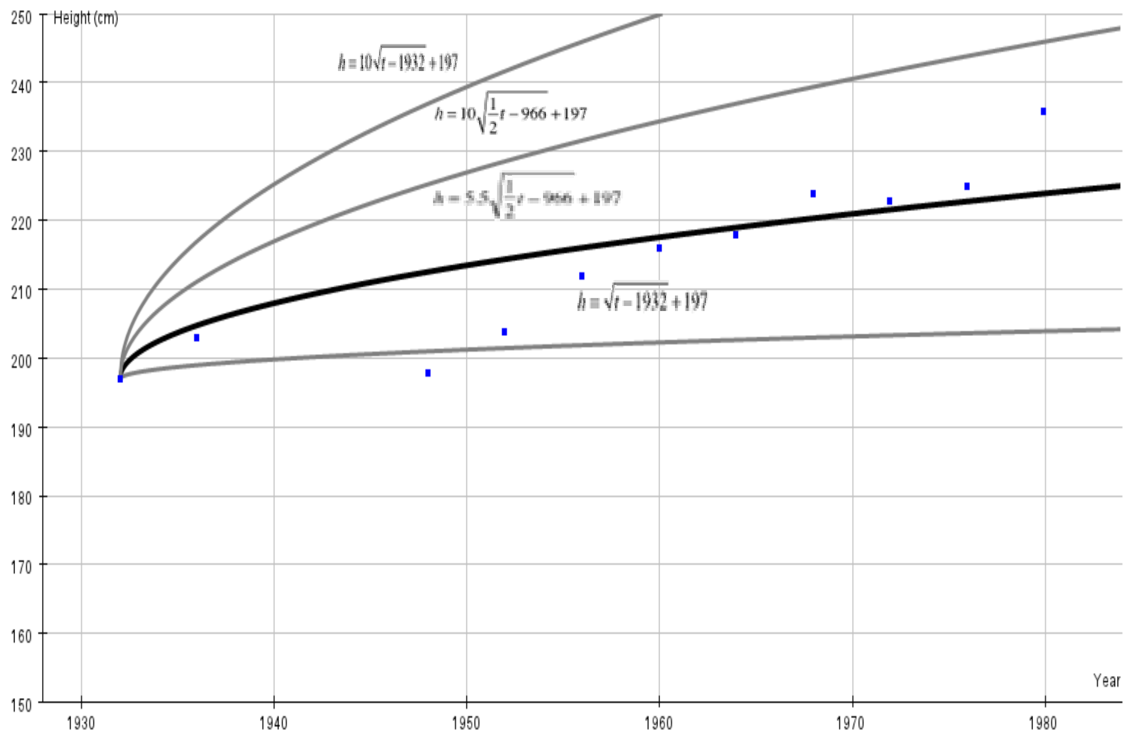
$$h = \frac{13}{20}t - 1059$$

Using the graphing program Autograph I found the line of regression for the data and the formula that goes along with it. I will use this equation, as it is more accurate. Shown below is this equation with its function. It varies slightly from the previous function.

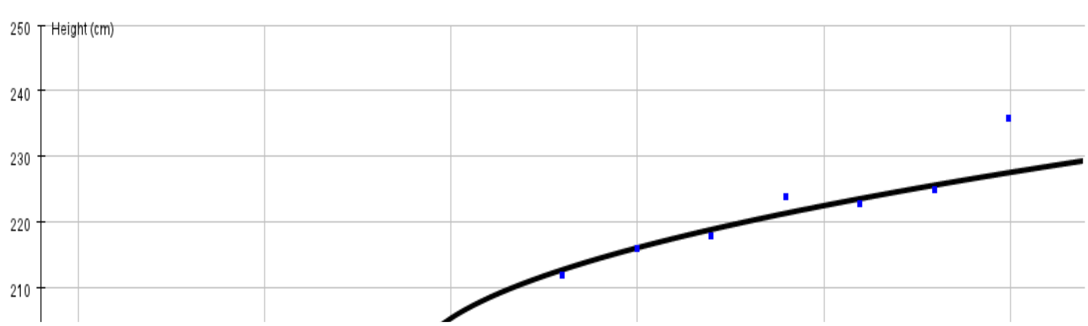


Now I will consider another function for the given data. Looking at the points plotted, it looks like the general shape of the square root function. This function is in the form: $h = \sqrt{t}$. Although the linear function fits the data well, the square root function would also be effective in graphing the data. Using the graphing program on my computer (Autograph) I used trial and error to find a suitable equation for this function. Shown below are a few of the equations that I tried, but

$h = 5.5\sqrt{\frac{1}{2}t - 966} + 197$ was the most suitable.

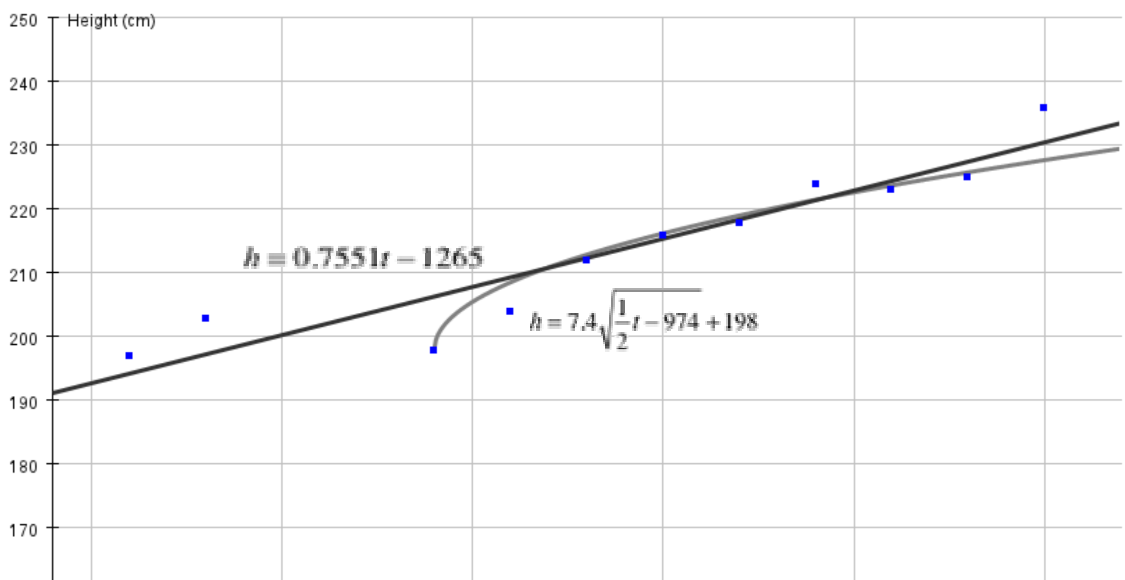


From looking at the graph above it seems as though a better fit for this function would be if the first point started at (1948, 198). This is partly because the Olympics were not held during World War II, which creates a gap in data. This war also meant that the athletes were unable to train so this is a reasonable assumption of why the two points after the war are almost identical to the points before the war. For this reasoning it would be acceptable to start the function at the point (1948, 198). To find the new equation for this function I used trial and error once again and found that the most suitable equation was: $h = 7.4\sqrt{\frac{1}{2}t - 974} + 198$



$$h = 7.4\sqrt{\frac{1}{2}t - 974} + 198$$

The model above fits the points better, however it ignores the first two data points. In both of the square root models created there will be outlier points, but in the second one there are less for the data given. For the predicted reason stated earlier, both of the square root functions that I created are acceptable to use. Also, the second model for the square root function is likely to be more accurate in predicting the winning heights for the future because it fits the data very well. Now I will plot both the linear and the square root function on the same graph to see the differences between the two.



Plotting both of the functions on the same graph is useful in showing the differences between each function. The linear function is more effective in going through or coming close to the majority of the points. In contrast, the square root function comes close to some of the points but it also leaves other points quite far off of the line. The square root function does seem more realistic for future data as the high jumps will start increasing by a lesser and lesser amount. The linear function however is effective for the current data, as the points do not seem to be increasing by smaller amounts quite yet. In the future the heights will continue to increase, but perhaps not at the same rate unless there is a new development such as new equipment or even a performance-enhancing drug that is legal.

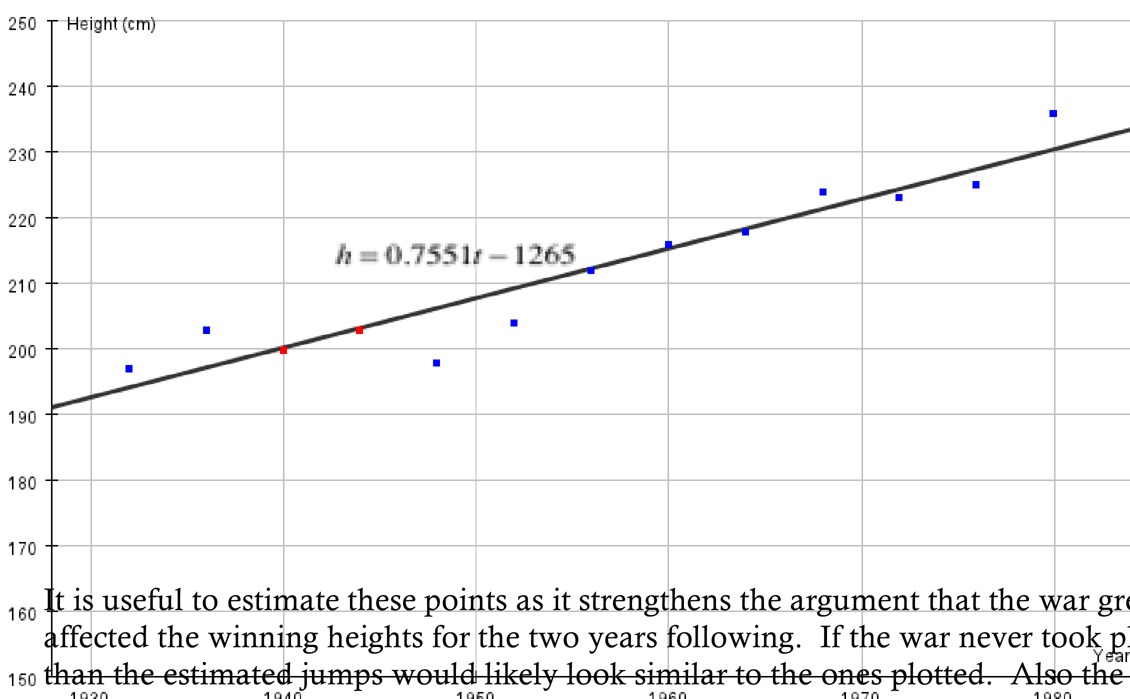
Had the games been held in 1940 and 1944, it is possible to predict what the winning heights would have been using the function created. I will use the linear equation, as it seems more accurate than the square root function for this prediction. I will place the dates in the equation and then isolate *h* to find the height at that date:

$$h = 0.7551(1940) - 1265$$
$$h = 200$$

Now when the year is 1944:

$$h = 0.7551(1944) - 1265$$
$$h = 203$$

Now when plotting these two points on the graph (in red), they seem quite accurate in estimating the winning heights for each of the given years. They both fit almost perfectly onto the line. The graph of this is shown below.



It is useful to estimate these points as it strengthens the argument that the war greatly affected the winning heights for the two years following. If the war never took place ~~than the estimated jumps would likely look similar to the ones plotted. Also the~~ two points following the war: (1948, 198) and (1952, 204) would likely have been higher jumps due to the fact that the athletes would be able to train and not focus on the war.

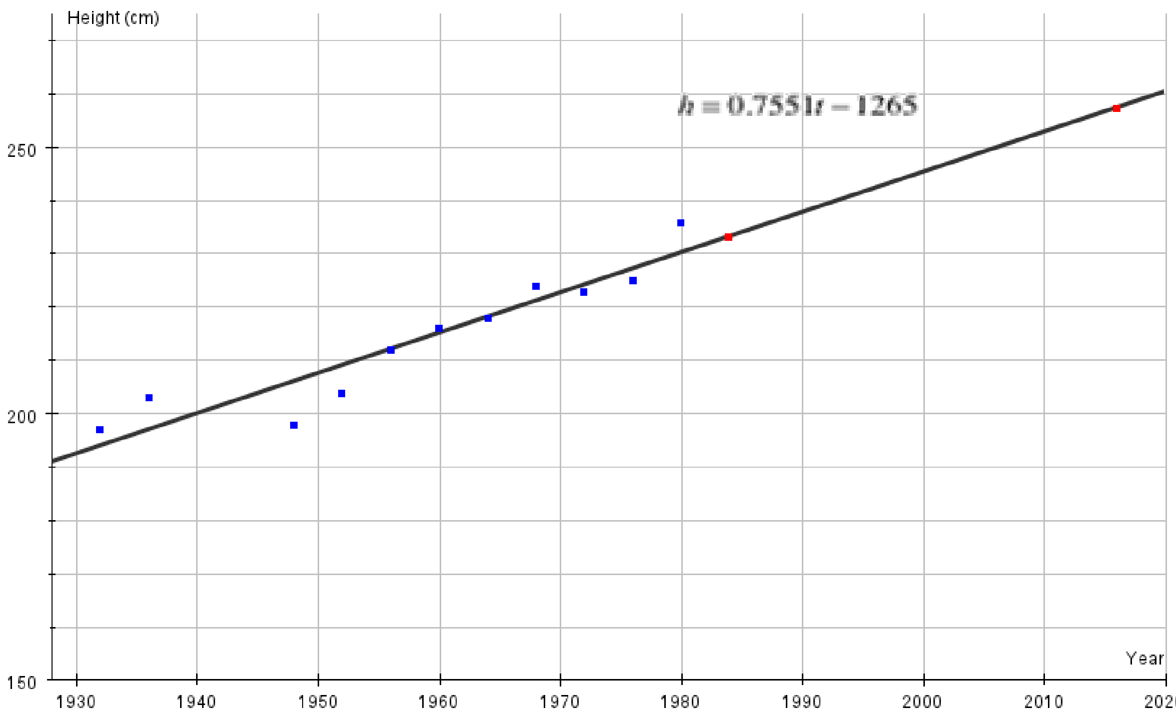
Using my model I will predict the winning heights for the years 1984 and 2016. I will use my equation for my linear function to find the most accurate estimate:

$$h = 0.7551(1984) - 1265$$
$$h = 233$$

Now when the year is 2016:

$$h = 0.7551(2016) - 1265$$
$$h = 257$$

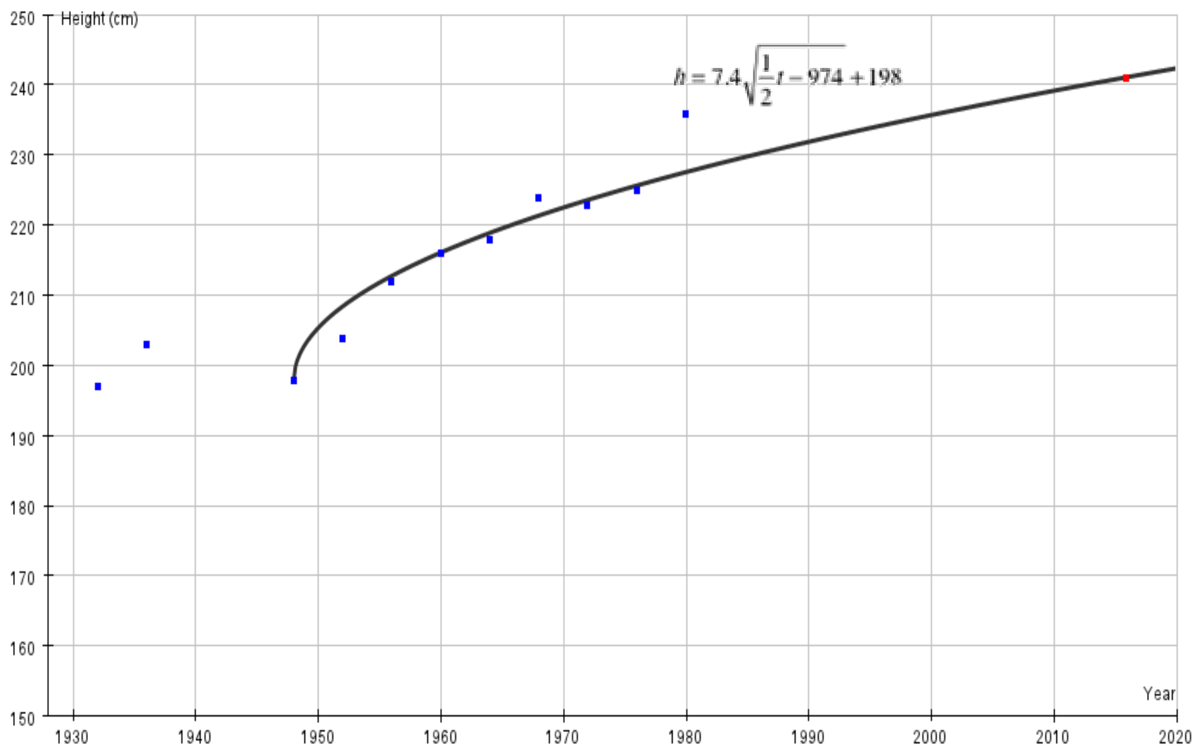
Now I will plot these points (in red) to see how accurate my estimation was on the graph below:



Looking at the graph it seems as though the point (1984, 233) is accurate, but the second point (2016, 257) is not logical in real life. Although the winning height of the long jump may increase in the future, it is doubtful that it will continue to increase at this linear slope. For this reason, winning heights in the future would be more appropriately modeled by using the square root function. The square root function slowly starts to increase by less and less, which makes perfect sense for this real world situation. So now I will find a more accurate estimation for the 2016 Olympics by using my square root function:

$$h = 7.4\sqrt{\frac{1}{2}(2016) - 974} + 198$$
$$h = 241$$

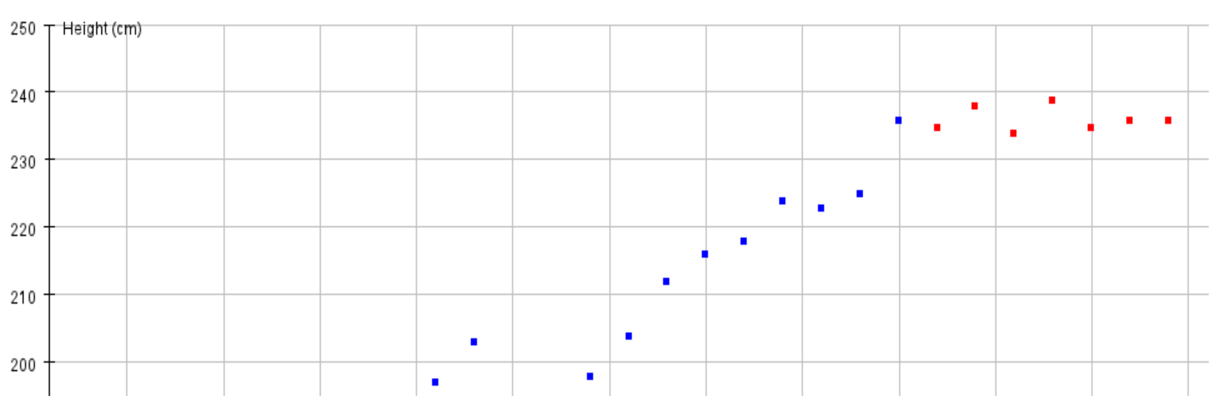
This height found from the square root function is a more realistic prediction. Like many other events, we can assume that the winning heights will not continue to increase at the same slope. Although the heights will likely continue to increase in the future, the rate at which they so do will probably drop. The graph below shows what I believe to be an accurate assumption of what the winning heights will look like in the future:



The following table gives the winning heights for all the other Olympic Games since 1896. Note: the Olympics were not held in 1900, 1916, and 1924.

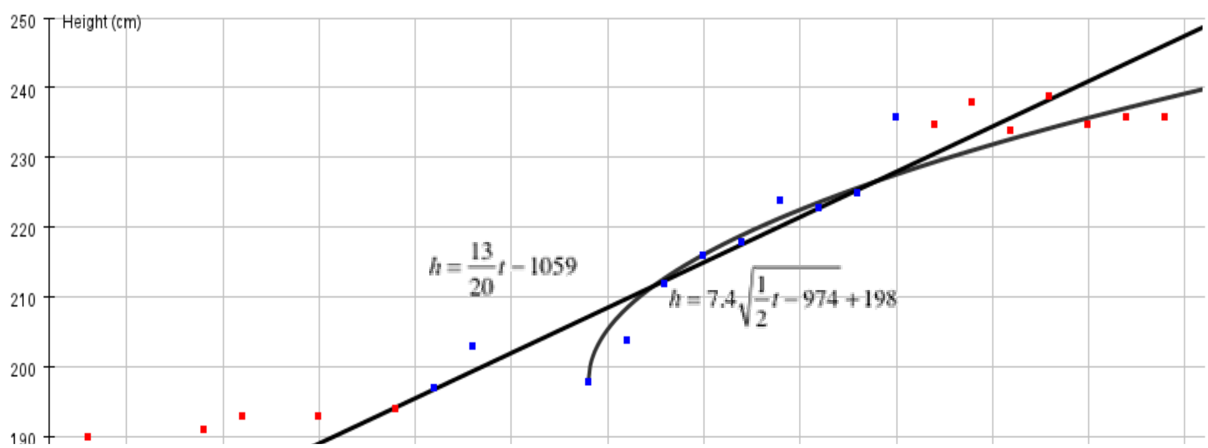
Year	1896	1904	1908	1912	1920	1928	1984	1988	1992	1996	2000	2004	2008
Height (cm)	190	180	191	193	193	194	235	238	234	239	235	236	236

The graph below is of all the additional data (in red) and the original data (in blue).



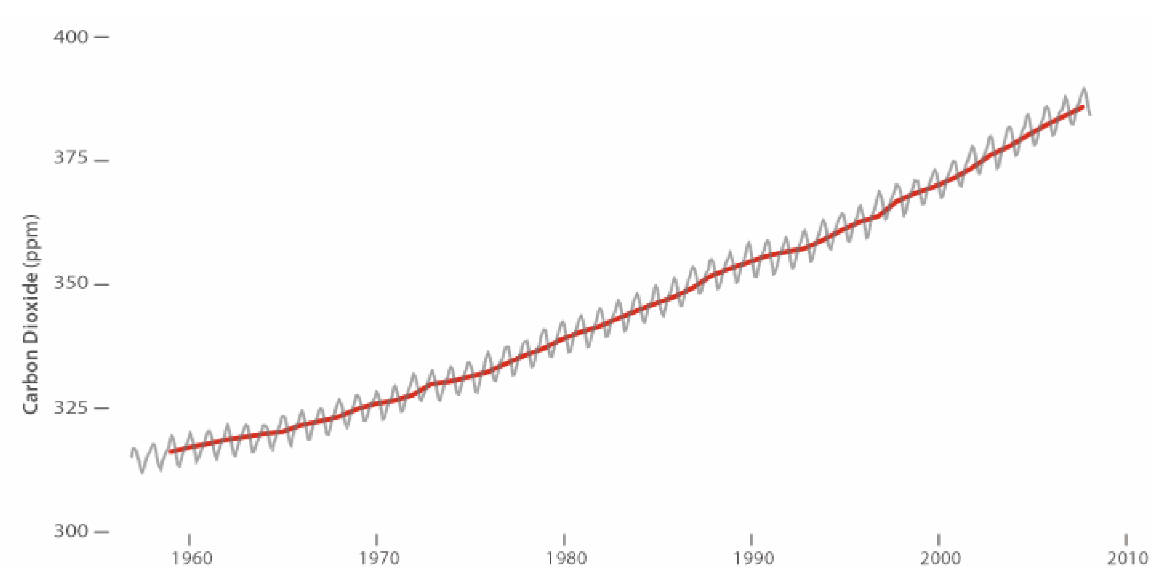
Now I plotted my original linear function and my square root function on the same graph to see how accurate my models were. I chose to use the original linear function that I found: $h = \frac{13}{20}t - 1059$ and not the line of regression that Autograph found. My reasoning for this was because the line of regression created an equation for the original data points plotted, so it did not include the additional data into this equation. Other than this, there were no modifications made to my model. The linear function fits the additional data decently, but it seems as though my prediction about the future that I stated earlier may be correct. The data points in the later years are increasing, but at a much slower rate. For this reason I then plotted my square root function which fits the additional data very well. I predict that the rate in which the winning heights will increase will continue to slow even more in the future.

The graph below is of both of the functions with the additional data plotted.



From 1896 to 2008 the winning heights have continued to increase. There are many things responsible for this such as better training techniques, better equipment, and possibly the use of performance enhancing drugs in certain cases. If I were to create a function for only the data given then I would choose the linear function as it fits the given data the most effectively. However, if I were to choose a function that modeled this data, but also the data in the future then I would choose the square root function. When creating a model for a real life situation it is important to make sure that the model makes logical sense. It is obvious that the square root function would be a more logical model to use than the linear function, but both of these functions were useful in visualizing the data.

My models can by compared to many other real-life situations. For example, the rate that humans are releasing carbon dioxide into the atmosphere has been increasing for many years. Our rate of releasing this carbon could be modeled by the linear function.



Kennedy, Caitlyn. "Climate Change: Atmospheric Carbon Dioxide." *National Oceanic and Atmospheric Administration*. September 25, 2011. <http://www.climatewatch.noaa.gov/article/2009/climate-change-atmospheric-carbon-dioxide>.

It continues to increase at a linear rate, but in the future this is likely to change. Since we have become aware of the problems that lie ahead due to this release of carbon, the model could change. Our release will most definitely not just come to a sudden stop, but it will hopefully increase at a slower rate. This gradual decline will take the shape of the square root function until hopefully our carbon release will start to level out and maybe even decrease. This example follows the exact same pattern as the winning heights for high jump in the Olympics. During the first years of the Olympics the winning heights continued to increase at a linear rate. Now these winning heights are becoming increasingly harder to beat so the rate of increase is slowing down. This will change the linear model to a square root model that will likely continue into the future.

My models created were effective for visualizing the data and also to predict winning heights for the future. When creating models like the functions in this assignment it is important to think logically and realistically. For example, simple things in this assignment were to realize that the height cannot be negative or that there will always be some outlier points because the winning heights will not increase consistently. The logical portion of creating models is perhaps even more important. If both of these aspects are thought about carefully than effective models will be created.