Joe monroy

Joe Monroy IB Math SL Type II Portfolio 1/12/13 Ms. Rozell Gold Medal Heights Objective: To consider the winning height for the men's high jump in the Olympic Games This table shows the heights achieved by the gold medal winners in the High Jump in numerous Summer Olympic Games throughout the 20th century 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 | (Note: Olympic Games were not held in 1940 and 1944.) Task One: Graph 1: This graph shows the correlation between the year and the height of the gold medal winners in the Olympic. The x-axis on this graph presents the years that the data was collect while the y-axis presents the height of the gold winner of that current year. This graph was graphed using the program Graphical Analysis. Constraints: During the years 1940 and 1944, the Olympic was cancelled due to World War II. Due to this, the data point does not continue according to its intervals of 4 years. It can be assumed that the high jumpers lost their practicing times during the World War, limiting their ability to improve on their skill, making the highest jumper in the year 1948 when the Olympic resumed lower than that of the previous year. Because of this the first two data points can be eliminated from the calculation of the equation due to the fact that the trend seems to start over at the year 1948. Task Two: The function that models the behavior of the graph the best is a linear function since the graph trends to correlate in a straight line as shown by the line of best fit. Calculation: To calculate the line of best fit, I split data the data points in half as shown by the gray line, and I then chose a point to be the median of both sides as shown below. Graph 2: Shows the midline splitting the data into two sections and the points that are chosen on both sides to represent the median. An equation

was then found that went through both points. This equation was then used as the line of best fit for the data. The equation y= mx+b was used to calculate the line of best fit since the trend of that data seemed to be linear. Graph 3: Shows the two points in comparison to the rest of the data By using the coordinates of the two points, the equation of the line of best fit was found as shown below. Slope=(y2- y1) / (x2- x1) Slope= 208. 84-228. 24(1953. 96-1973. 28) Slope=-19. 4(-19. 32) Slope= 1. 0 With the slope, I now find the y-intercept of the line of best fit y= 1. 0x+b 208. 84= 1. 0 1953. 96+b b= -1, 745. 12 y= 1. 0x-1, 745. 12 Task Three: Graph 4: Shows the line of best fit as found by using the two median points As seen in the graph shown above, the line of best fit goes through the middle of all the data. Since this line was calculated using the two points the line goes through both points. The first median point is an accurate representation of the left side of the data since the data points are in a straight line in correlation. However the second point was not quite accurate due to the fact that in years 1972 and 1976 there was not much improvement in the height of the high jumpers so the points that are used to calculate the median is quite off. This model does not work for the years after 1980 because there are limitations to how high a person can jump since in modern society we cannot yet overcome the forces of gravity. Since this model depicts that the years after 1980 there will be a steady increase in the height of the gold medal heights, this model is then false after about 10 years after 1980. This model also depicts that before the year 1952, there is a steady decline in the heights of the gold medal heights which is not true as seen in the years 1932 and 1936. This linear model is only valid in range of the years 1932 to about 1990 which is in the range of the data given. Task Four: Graph 5: Shows both models

(linear and logarithmic) in relative to each other The equation used to model the line on bottom is the natural logarithm model while the equation used to model the line above is the linear model. When comparing the two models, the difference between the two is that the y-intercept is lower for the natural logarithm model than the linear model, while the slope of the logarithmic function is greater than the linear. The two functions also intersect at the point (1962, 217). Task Five: Using the natural logarithm model as the function, the years 1940 and 1944 were plugged into the equation and the heights of both years are calculated as shown below. y= 1, 476(ln0. 0005904x) y= 1, 476(ln0. 00059041940) y= 200 The expected winning height in the 1940 Olympics using the logarithmic function is 200 cm. y= 1, 476(ln0. 00059041944) y= 203 The expected winning height in the 1944 Olympics using the logarithmic function is 203 cm. I then repeated the calculations but used the linear model as the function for this time. I used the linear model because I believe it is more accurate than the logarithmic model. y= 1. 0x-1, 745 y= 1. 01940-1, 745 y= 195 The expected winning height in the 1940 Olympics using the linear function is 195 cm. y= 1. 01944-1, 745 y= 199 The expected winning height in the 1944 Olympics using the linear function is 199 cm. I am going to use the results from the linear model function because it fits the trend of the data better. Due to the elimination of the first two points on the graph, the estimation of the years 1940 and 1944 will be lower than expected because the slope of the graph will be much higher. The graph does not take into account that the 1940 and 1944 Olympics didn't occur because of World War II. If the trend of the winning heights in 1932 and 1936 continued in the years of 1940 and 1944, they would be significantly higher than the values I calculated using my

linear model. As said in previous explanations, the heights at the 1948 Olympics decreased in comparison to the 1936 Olympics because the jumpers had no time to practice or compete during the war. This made the graph develop a new trend, making the 1932 and 1936 Olympics outliers. Task Six: For predicting the winning heights in the 1984 and 2016 Olympics I will use my natural logarithmic model because it follows the trend of the later years better than in the earlier years. I came to the conclusion that my linear model is more accurate for the earlier years, however my logarithmic model is more accurate for the later years y= 1, 476(ln0. 0005904x) y= 1, 476(ln0. 00059041984) y= 233 The predicted winning height in the 1984 Olympics using the logarithmic function is 233 cm y= 1, 476(ln0. 00059042016) y= 257 The predicted winning height in the 2016 using the logarithmic function is 257 cm The height estimated by this model in the year 1984 seems credible since the increase in height does not seem to be out of reach of the human capability. However, in the year 2016, the height that was reached was 257 cm, which seems to be higher than a human can possibly jump. This is because the natural logarithm model depicts a straight line following the year 1980 and because a steady increase in the heights of the Olympic high jumpers does not seem possible for a human being, the further away from the year 1980 the lower the credibility of the answer. I conclude that the logarithmic model is a good model for determining the winning heights in future years, but since humans are physically unable to jump to heights of excess of 250 cm, the function is irrelevant past a certain amount of years. The following table shows the winning heights in every Olympic Games since 1986 Year | 1896 | 1904 | 1908 | 1912 | 1920 | 1928 | 1932 | 1936 | 1948 | 1952 | 1956 | Height(cm) | 190 | 180 | 191 | 193 | 193 |

194 | 197 | 203 | 198 | 204 | 212 | (Note: The 1916, 1940 and 1944 Olympics were not held due to war. The 1900 and 1924 Olympics are not shown.) Year | 1960 | 1964 | 1968 | 1972 | 1976 | 1980 | 1984 | 1988 | 1992 | 1996 | 2000 | 2004 | 2008 | Height(cm) | 216 | 218 | 224 | 223 | 225 | 236 | 235 | 238 | 234 | 239 | 235 | 236 | 236 | Task Seven: Graph 6: Shows the logarithm model with the additional points of data The model that was used to find the line of best fit in the last task does not fit the additional data. This is because between the years 1896 and 2008 there has been fluctuations in the height of the gold winners. The data starts out as a straight line with an outlier in the year 1904 then curves upwards in the years leading up to World War II (1928, 1932, 1936). During World War II the heights was assumed to have dropped due to the lost in practice times so the trend starts again in 1948 with a straight linear correlation all the way through to the year 1988. After that the data seems to level off into a horizontal correlation. It can be assumed that the data levels off because there is a limit to how high humans can physically jump. The modification that needs to be made is that all data points on the graph must be included to have a better sense of the median and range of the data to be more accurate as shown below. Graph 7: Shows the additional points and the newly modified logarithm model Furthermore, there are many other functions that can be used to model these data points such as the cubic model and the Gaussian model. As shown below the cubic model can model both the upward curve in the years leading to World War II and the leveling off during the last 4 to 5 years. However the model curves upward before the year 1896 and curves downward after the year 2008 which these two directions does not agree with the data given so this model can only be used between the years 1896 to the year 2008. Graph 8: Shows

the cubic model in relation to all the data points Graph 9: Shows the Gaussian model in relation to all the data points The Gaussian model can also be used to model this part of the data but it is different from the cubic model since the Gaussian model starts off with a level and horizontal line then curves up similarly to the cubic model. In addition to that the Gaussian model also models the leveling off as the years approach 2008 but has the same limitation as the cubic model as the graph slopes back downwards after the year 2008. The Gaussian model is a better representation of the data since at the beginning of the data range the slope does not curve downward before going back upwards like the cubic model. However, the disadvantage that the Gaussian model and the cubic model have is that they both do not show the fluctuation of the heights of the gold medalists during the year 1821 to the year 1896 (1821 being the year that the Olympics started).