The investigation on how the mathematical equation from regression predicts the average point scored of an individual volleyball player in a game

Subject: Mathematics Higher Level

Word Count: 3983

Table of Contents

Introduction:

Statistics is a study of the collection, analyzed, presentation and interpreted a group of data. For decades, its knowledge and applications are widely used in many fields in our daily life. Statistics is applicable in both larger and smaller scale. In a larger scale, it can be applied to something like Big Data of a country – a set of data whose size is beyond the capability of the normal computer to interpret and analyze to produce a solution [1]. In which the result from using Big Data can be used to improve the country's logistics, economics, or even their education systems. In a smaller scale, it is applicable to something personal, such as tracking the number of coffee a person drinks in a day, or even the nutrients of a person. Thus, my objective for this paper is to connect the idea of statistics and mathematics to my interest in volleyball. I would achieve this objective by modeling a formula which I can use with volleyball.

Background information:

To understand the objective of the essay, it is essential to understand the concept and section of the Statistic this essay going to explore. In this case, the section of the statistics is the regression analysis.

Regression refers to a technique in statistics which is used for predictive modeling or data mining tasks. Its major tasks are to predict the outcome, and how the outcome is impacted by variables [2]. What produces the outcome for the regression is the relationship between the dependent variable, called the outcome variable, and the independent variable, called the predictor variable. There are many types of regression, in which the usage of each type depends on the circumstance and the field of study. The examples are Linear Regression or Logarithmic

Regression. In which, each type of regression usage depends on the type of data and the distribution [3].

This essay focuses on linear regression. It is the simplest form of regression, which the formula is in the form of a linear line equation. The equation is given by:

$$\hat{y} = mx + c \tag{1}$$

Where y is the outcome variable, m is the coefficient of the predictor variable, x is the predictor variable and c is a constant. There are also various types of linear regression, thus this essay features multi-linear regression, which deals with multiple predictor variables. The reason is because, in a real-life situation, there are more than one feature that affect that may contribute to the change in the outcome of the variable. The equation for the multi-linear equation is given by:

$$\hat{y} = \beta_n x_n + \cdots + \beta_1 x_1 + \beta_0 \tag{2}$$

Where β_n is the coefficient of the each predictor variable and β_0 is a constant.

While these techniques are usable to determine the outcome value, forecast the effect and trend forecasting, the multi-linear regression requires another branch of mathematics field of study. Examples are vectors and matrixes. More complex applications like partial differentiation or Lagrange multiplier could make the forecast and trend more accurate and realistic. Lagrange Multiplier is named after a French-Italian Mathematician Joseph-Louis Lagrange. It is a method of maximizing or minimizing $f(x_1, x_2 \ldots x_n)$ subjected to constraints $g(x_1, x_2 \ldots x_n) = C$. Moreover, this method is widely used to solve the optimization problems in many field of studies, such as: Economics or Physics. In which it can also be applied to regression with constraint. The equation of Lagrange Multiplier is given by:

$$\nabla f(x, y, z) = \lambda \nabla g(x, y, z) \tag{3}$$

$$g(x, y, z) = C \tag{4}$$

Where ∇f and ∇g are the vector gradient of each function, and "λ" is a scalar showing that these two vector gradients are parallel [4].

Minimize the sum of residual (Error):

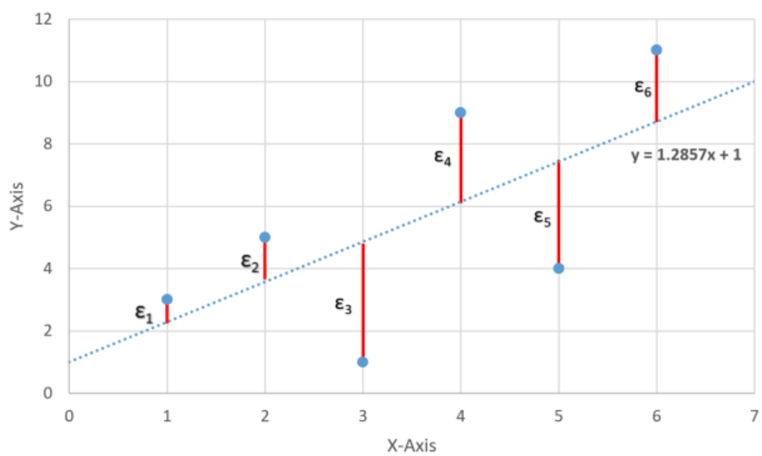

Figure 1. Simple Linear Regression showing residual

The line of best fit, trend line, is the most appropriate line passing through the data points. Though, it is unnecessary for the line to pass through all the data points. In addition, trend line best describes the relationship between the independent variables and the dependent variable. Thus, in Figure 1 – a snapshot from Microsoft Excel – shows 6 data points and its best-fit line as an example of the relationship between x-axis and y-axis. The red line on the figure represents the differences between the actual data point "y" and the predicted data point "$\hat{y}$" on the best-fit line, or "$y - \hat{y}$". Furthermore, it is represented by the symbol "ε", which is the error or the residual. Thus, the best-fit line must be a line which has the least total error between the "y" and "$\hat{y}$". Moreover, in Figure 1 includes the subscription on its residual, actual data or even predictive value, where the subscription represents the numbered value out of all the collected data. In other word, the best-fit line is the line of the least sum of residual.

Based on the aforementioned definition, the sum of the residual can be written as:

$$\sum_{i=1}^{n} \varepsilon_i = \varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \cdots\cdots + \varepsilon_k \tag{5}$$

Knowing that "$\varepsilon = y_i - \hat{y}_i$", which is mentioned above. Hence, by substituting "$y_i - \hat{y}_i$" for "ε", the sum of residual the equation can be written as such:

$$\sum_{i=1}^{n} \varepsilon_i = (y_1 - \hat{y}_1) + (y_2 - \hat{y}_2) + (y_3 - \hat{y}_3) + \cdots\cdots + (y_n - \hat{y}_n) \tag{6}$$

Through observation, "$y - \hat{y}$" yields a positive residual when the value of "$y \geq \hat{y}$", and a negative residual value when the value of "$y < \hat{y}$". With the reference to Figure 1, at $x = 3$ and $x = 5$, return negative residuals, while other points return positive residual. Ideally, all the residuals must be positive, because the best-fit line comes from the sum of residual. Hence, when adding all the residuals, if the negative residual added to the sum of residual, the total could sometimes be cancelled or reduced.

Absolute could be added to the residual, "$|\varepsilon_i|$" or "$|y_i - \hat{y}_i|$", thus the differences between the actual data and the predicted data, residual, will always be positive. Now, the sum of the residual will be appropriate to use to model the best-fit line.

The sum of the residual equation with absolute can be modified as given:

$$\sum_{i=1}^{n} \varepsilon_i = |y_1 - \hat{y}_1| + |y_2 - \hat{y}_2| + |y_3 - \hat{y}_3| + \cdots\cdots + |y_n - \hat{y}_n|$$

$$\sum_{i=1}^{n} \varepsilon_i = \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{7}$$

One approach to model the best-fit line, the line with the least sum of residual, is differentiating the sum of the residual equation – by differentiating the equation yields the minimum value of its derivative. However, it would be challenging because of the sigma notation and the absolute residual must be differentiated, where the absolute value will be very long and complicate to differentiate. An alternative approach can be taken.

An alternative approach of differentiating the sum of the residual equation can be divided into two separate sections. The first section is about differentiating the sigma notation, and the second section is a method called Least Square.

The first section of differentiating the sigma notation can be explained through an example given below:

$$\sum_{i=1}^{n} f(x)^i = f(x) + f(x)^2 + f(x)^3 + \cdots\cdots + f(x)^n$$

$$\frac{d\sum_{i=1}^{n} f(x)^i}{dx} = f'(x) + 2f(x)f'(x) + 3f(x)^2 f'(x) + \cdots\cdots + nf(x)^{n-1}f'(x)$$

$$\frac{d\sum_{i=1}^{n} f(x)^i}{dx} = \sum_{i=1}^{n} if(x)^{i-1} \tag{8}$$

From the example, it shows that the sigma notation is not affected by the derivation process. The reason is the sigma equation is equivalent to the summation of functions. Therefore, according to the additional rule of differentiation, each function in the summation can be differentiated independently.

The second section talks about the Least Square method. With a set of data, there is variation about the mean, "$\hat{y}$". Ideally, in order to have a good deviation about the mean of the actual data, the standard deviation "σ", must be taken into consideration. However, because of the problem with the cancellation of the residual, using absolute on the deviation or squaring it would help solve the problem [5]. Linking back to differentiating the sum of the residual equation with absolute residual, instead of using "$|\varepsilon|$", it can be substitute as "ε^2". Hence, the equation of the sum of residual now becomes:

$$\sum_{i=1}^{n} {\varepsilon_i}^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{9}$$

Now, the sum of the residual is differentiable. However, by using the Least Square to reduce the complexity in differentiation, the residual with larger will have more weight than the smaller residual value. On the other hand, the residual from the absolute are all weight equally.

Simple Linear Regression:

Simple linear regression is the simplest type of regression. It has only one predictor variable and one outcome variable. The equation of the simple linear regression best-fit line can be represented by the equation:

$$\hat{y}_i = \beta_0 + \beta_1 x_i \tag{10}$$

Where "$\hat{y}$" is the outcome variable, and "x" is the predictor variable. In this case, they are both an actual collected data, also called observed data. "β_0" is the intercept parameter which tells the status of the outcome variable when the predictor variable is 0, and "β_1" is the slope parameter which tells the amount of change in outcome variable when the predictor changes by 1 unit [6]. From the observation, "$\hat{y}$" in equation 10 is the same variable as "$\hat{y}$" in equation 9 of the sum of the residual equation. Thus, by substituting this equation into the sum of the residual equation, it yields another sum of the residual equation in term of two observed values and two coefficient:

$$\sum_{i=1}^{n} {\varepsilon_i}^2 = \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_i))^2 \tag{11}$$

Currently, to obtain the equation of the line of best fit is by solving for the two coefficients (β_0 and β_1) with the observed data (x_i and y_i). Therefore, the approach in finding the coefficient is through partial differentiation. While implicit differentiation might seem to be able to be used in differentiating multiple variables, but it actually is used for of one variable that are written implicitly [7]. With partial differentiation, the equation with multiple variables is being derived with respect to each variable of the equation, also, while assuming other variables are constant.

Hence, it can be applied to the sum of residual, which has multiple variables in the equation. With partial differentiation and equate it to zero, the minimum value of the coefficient of the line of best fit can be determined. The result of the partial differentiation when it is equated to 0 is similar to the normal differentiation, which yields either minimum, maximum or a saddle point. In fact, due to the principle of the Least Squared Method, the derived equation will always return the minimum value, because, logically, the maximum residual does not exist in this case – the maximum residual is when the slope is a vertical line where its value is undefined.

$$\frac{\partial \sum_{i=1}^{n} {\varepsilon_i}^2}{\partial \beta_0} = 0 \tag{12}$$

$$\frac{\partial \sum_{i=1}^{n} {\varepsilon_i}^2}{\partial \beta_1} = 0 \tag{13}$$

Partial differential derives the sum of the residual equation and produces 2 equations.

Derive the equation with respect to β_0:

$$\frac{\partial \sum_{i=1}^{n} {\varepsilon_i}^2}{\partial \beta_0} = \frac{\partial \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_i))^2}{\partial \beta_0}$$

$$\frac{\partial \sum_{i=1}^{n} {\varepsilon_i}^2}{\partial \beta_0} = \sum_{i=1}^{n} 2(y_i - \beta_0 - \beta_1 x_i)(-1)$$

$$0 = \sum_{i=1}^{n} 2(y_i - \beta_0 - \beta_1 x_i)(-1)$$

$$0 = -2\sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)$$

$$0 = \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \beta_0 - \sum_{i=1}^{n} \beta_1 x_i$$

$$0 = \sum_{i=1}^{n} y_i - n\beta_0 - \beta_1 \sum_{i=1}^{n} x_i$$

Derive the equation with respect to β_1:

$$\frac{\partial \sum_{i=1}^{n} \varepsilon_i{}^2}{\partial \beta_1} = \frac{\partial \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_i))^2}{\partial \beta_1}$$

$$\frac{\partial \sum_{i=1}^{n} \varepsilon_i{}^2}{\partial \beta_1} = \sum_{i=1}^{n} 2(y_i - \beta_0 - \beta_1 x_i)(-x_i)$$

$$0 = \sum_{i=1}^{n} 2(y_i - \beta_0 - \beta_1 x_i)(-x_i)$$

$$0 = \sum_{i=1}^{n} (x_i y_i - x_i \beta_0 - \beta_1 x_i{}^2)$$

$$0 = \sum_{i=1}^{n} x_i y_i - \beta_0 \sum_{i=1}^{n} x_i - \beta_1 \sum_{i=1}^{n} x_i{}^2$$

Hence, with two equations and two unknown variables, simultaneous equation can be used to solve for the two missing variables, β_0 and β_1:

$$0 = \sum_{i=1}^{n} y_i - n\beta_0 - \beta_1 \sum_{i=1}^{n} x_i$$

$$\beta_0 = \frac{\sum_{i=1}^{n} y_i - \beta_1 \sum_{i=1}^{n} x_i}{n}$$

Substitute β_0 into second equation to find β_1

$$0 = \sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} y_i - \beta_1 \sum_{i=1}^{n} x_i}{n} \sum_{i=1}^{n} x_i - \beta_1 \sum_{i=1}^{n} x_i{}^2$$

$$0 = \sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i - \beta_1 (\sum_{i=1}^{n} x_i)^2}{n} - \beta_1 \sum_{i=1}^{n} x_i{}^2$$

$$0 = n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i + \beta_1 \left(\sum_{i=1}^{n} x_i \right)^2 - n\beta_1 \sum_{i=1}^{n} x_i{}^2$$

$$\beta_1 = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} {x_i}^2 - (\sum_{i=1}^{n} x_i)^2}$$

Now, substitute β_1 to find value for β_0:

$$\beta_0 = \frac{\sum_{i=1}^{n} y_i - \beta_1 \sum_{i=1}^{n} x_i}{n}$$

$$n\beta_0 = \sum_{i=1}^{n} y_i - (\frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} {x_i}^2 - (\sum_{i=1}^{n} x_i)^2}) \sum_{i=1}^{n} x_i$$

$$n\beta_0 = \frac{n\sum_{i=1}^{n} y_i \sum_{i=1}^{n} {x_i}^2 - (\sum_{i=1}^{n} x_i)^2 \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} {x_i}^2 - (\sum_{i=1}^{n} x_i)^2} - \left(\frac{n\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i - (\sum_{i=1}^{n} x_i)^2 \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} {x_i}^2 - (\sum_{i=1}^{n} x_i)^2}\right)$$

$$n\beta_0 = \frac{n\sum_{i=1}^{n} y_i \sum_{i=1}^{n} {x_i}^2 - (\sum_{i=1}^{n} x_i)^2 \sum_{i=1}^{n} y_i - n\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i + (\sum_{i=1}^{n} x_i)^2 \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} {x_i}^2 - (\sum_{i=1}^{n} x_i)^2}$$

$$\beta_0 = \frac{\sum_{i=1}^{n} y_i \sum_{i=1}^{n} {x_i}^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} {x_i}^2 - (\sum_{i=1}^{n} x_i)^2}$$

In the end, the coefficient variable for the linear regression where β_0 and β_1 are equal to:

$$\beta_0 = \frac{\sum_{i=1}^{n} y_i \sum_{i=1}^{n} {x_i}^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} {x_i}^2 - (\sum_{i=1}^{n} x_i)^2} \tag{14}$$

$$\beta_1 = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} {x_i}^2 - (\sum_{i=1}^{n} x_i)^2} \tag{15}$$

Multi-Linear Regression:

The concept and methods of simple linear regression can also apply to Multi-Linear Regression. The similarity between the simple linear regression and multi-linear regression lies in their beginning equation, the sum of the residual equation:

$$\sum_{i=1}^{n} {\varepsilon_i}^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{9}$$

For multi-linear regression the outcome variable "$\widehat{y_i}$" is similar linear line equation as the linear equation of the simple linear regression, equation 10; however, with multiple predictor variables. The multi-linear regression's equation for the best-fit line is given by:

$$\widehat{y_i} = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \cdots\cdots + \beta_k x_{ki} \tag{16}$$

Deduced from the equation, "x_{ki}" represents an independent variable or predictor value, from the best-fit line, in different axes, where "k" indicates the number of predictor variables and its coefficient, and "i" indicates the numbered of point out of "n" points, where "n" is the total number of actual data collected. Hence, by substituting equation 16 into the sum of residual, the equation becomes:

$$\sum_{i=1}^{n} {\varepsilon_i}^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i} - \beta_3 x_{3i} - \cdots\cdots - \beta_k x_{ki})^2 \tag{17}$$

Hence, by taking the same partial differentiation approach, the sum of residual for multi-linear regression will be derived with respect to "$k + 1$" variables which yields output "$k + 1$" equations.

Partially derive the equation with respect to β_0 and let the partial derivative equals to zero:

$$\frac{\partial \sum_{i=1}^{n} {\varepsilon_i}^2}{\partial \beta_0} = \frac{\partial \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i} - \beta_3 x_{3i} - \cdots\cdots - \beta_k x_{ki})^2}{\partial \beta_0}$$

$$\frac{\partial \sum_{i=1}^{n} {\varepsilon_i}^2}{\partial \beta_0} = 2 \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i - \beta_3 x_i - \cdots\cdots - \beta_k x_{ki})\,(-1)$$

$$0 = \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \beta_0 - \sum_{i=1}^{n} \beta_1 x_{1i} - \sum_{i=1}^{n} \beta_2 x_{2i} - \sum_{i=1}^{n} \beta_3 x_{3i} - \cdots\cdots - \sum_{i=1}^{n} \beta_k x_{ki}$$

$$\sum_{i=1}^{n} y_i = n\beta_0 + \beta_1 \sum_{i=1}^{n} x_{1i} + \beta_2 \sum_{i=1}^{n} x_{2i} + \beta_3 \sum_{i=1}^{n} x_{3i} + \cdots\cdots + \beta_k \sum_{i=1}^{n} x_{ki} \tag{18}$$

Derive the equation with respect to β_1 and let the derivative equals to zero:

$$\frac{\partial \sum_{i=1}^{n} \varepsilon_i^{\,2}}{\partial \beta_1} = \frac{\partial \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i} - \beta_3 x_{3i} - \cdots\cdots - \beta_k x_{ki})^2}{\partial \beta_1}$$

$$\frac{\partial \sum_{i=1}^{n} \varepsilon_i^{\,2}}{\partial \beta_0} = 2\sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i} - \beta_3 x_{3i} - \cdots\cdots - \beta_k x_{ki})\,(-x_{1i})$$

$$0 = \sum_{i=1}^{n} y_i x_{1i} - \sum_{i=1}^{n} \beta_0 x_{1i} - \sum_{i=1}^{n} \beta_1 (x_{1i})^2 - \sum_{i=1}^{n} \beta_2 x_{1i} x_{2i} - \sum_{i=1}^{n} \beta_3 x_{1i} x_{3i} - \cdots\cdots - \sum_{i=1}^{n} \beta_k x_{1i} x_{ki}$$

$$\sum_{i=1}^{n} y_i x_{1i} = \beta_0 \sum_{i=1}^{n} x_{1i} + \beta_1 \sum_{i=1}^{n} (x_{1i})^2 + \beta_2 \sum_{i=1}^{n} x_{1i} x_{2i} + \beta_3 \sum_{i=1}^{n} x_{1i} x_{3i} + \cdots\cdots + \beta_k \sum_{i=1}^{n} x_{1i} x_{ki} \tag{19}$$

Similarly, derive the equation with respect to β_k and let the derivative equals to zero, where k is the last value:

$$\frac{\partial \sum_{i=1}^{n} \varepsilon_i^{\,2}}{\partial \beta_k} = \frac{\partial \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i} - \beta_3 x_{3i} - \cdots\cdots - \beta_k x_{ki})^2}{\partial \beta_k}$$

$$\frac{\partial \sum_{i=1}^{n} \varepsilon_i^{\,2}}{\partial \beta_k} = 2\sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i} - \beta_3 x_{3i} - \cdots\cdots - \beta_k x_{ki})\,(-x_{ki})$$

$$0 = \sum_{i=1}^{n} y_i x_{ki} - \sum_{i=1}^{n} \beta_0 x_{ki} - \sum_{i=1}^{n} \beta_1 x_{1i} x_{ki} - \sum_{i=1}^{n} \beta_2 x_{2i} x_{ki} - \sum_{i=1}^{n} \beta_3 x_{3i} x_{ki} - \cdots\cdots - \sum_{i=1}^{n} \beta_k (x_{ki})^2$$

$$\sum_{i=1}^{n} y_i x_{ki} = \beta_0 \sum_{i=1}^{n} x_{ki} + \beta_1 \sum_{i=1}^{n} x_{1i} x_{ki} + \beta_2 \sum_{i=1}^{n} x_{2i} x_{ki} + \beta_3 \sum_{i=1}^{n} x_{3i} x_{ki} + \cdots\cdots + \beta_k \sum_{i=1}^{n} (x_{ki})^2 \tag{20}$$

In simple linear regression, the value of the variable can be solved algebraically through the use of simultaneous equation. Solving simultaneous equation uses the substitution method, which this method is done by isolating one variable, substitute it into other equation and solve for the value of another variable. After that, substitute the obtained value into the first equation and solve for the variable used for substitution. The advantage of using simultaneous is the problem can be solved by hand, and quick. However, it would not be handy when it comes to many variables and equations. The reason is, it will take a long time to solve or to even isolate one variable to solve for others. It is not impossible to solve, but it is not efficient. Hence, for the multi-linear regression, where there are $k+1$ variables from $k+1$ equations, can be solve using an alternative method of matrix.

By using matrix as a technique to solve for multiple variables equations, the equations are represented as:

$$\begin{bmatrix} n & \sum_{i=1}^{n} x_{1i} & \sum_{i=1}^{n} x_{2i} & \cdots\cdots & \sum_{i=1}^{n} x_{ki} \\ \sum_{i=1}^{n} x_{1i} & \sum_{i=1}^{n} (x_{1i})^2 & \sum_{i=1}^{n} x_{1i}x_{2i} & \cdots\cdots & \sum_{i=1}^{n} x_{1i}x_{ki} \\ \sum_{i=1}^{n} x_{2i} & \sum_{i=1}^{n} x_{1i}x_{2i} & \sum_{i=1}^{n} (x_{2i})^2 & \cdots\cdots & \sum_{i=1}^{n} x_{2i}x_{ki} \\ \vdots & \ddots & ⋰ & ⋰ & \vdots \\ \vdots & ⋰ & \ddots & \ddots & \vdots \\ \sum_{i=1}^{n} x_{ki} & \sum_{i=1}^{n} x_{1i}x_{ki} & \sum_{i=1}^{n} x_{2i}x_{ki} & \cdots\cdots & \sum_{i=1}^{n} (x_{ki})^2 \end{bmatrix}_{(n+1,n+1)} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \vdots \\ \beta_k \end{bmatrix}_{(n+1,1)} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} y_i x_{1i} \\ \sum_{i=1}^{n} y_i x_{2i} \\ \vdots \\ \vdots \\ \sum_{i=1}^{n} y_i x_{ki} \end{bmatrix}_{(n+1,1)} \quad (21)$$

By letting "A" be the matrix of coefficients for each variable "β_0" to "β_k" in the matrix "ω", which is equals to the number in the matrix "B". The objective is to rearrange the equation and make matrix "ω" a subject. Hence, solve for each variable "β_0" to "β_k" in matrix "ω". The equation can be rewritten as:

$$A = \begin{bmatrix} n & \sum_{i=1}^{n} x_{1i} & \cdots & \sum_{i=1}^{n} x_{ki} \\ \sum_{i=1}^{n} x_{1i} & \sum_{i=1}^{k} (x_{1i})^2 & \cdots & \sum_{i=1}^{n} x_{1i}x_{ki} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} x_{ki} & \sum_{i=1}^{n} x_{1i}x_{ki} & \cdots & \sum_{i=1}^{n} (x_{ki})^2 \end{bmatrix}_{(n+1,n+1)} ,\omega = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}_{(n+1,1)} ,B = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} y_i x_{1i} \\ \vdots \\ \sum_{i=1}^{n} y_i x_{ki} \end{bmatrix}_{(n+1,1)}$$

$$A\omega = B \tag{22}$$

However, matrix "A" cannot move to divide matrix "B". Therefore, the inverse matrix must be used. The property of the inverse matrix is the product of the inverse matrix and the original matrix yields an identity matrix. Thus, by multiple inverse matrix of A on both side will give an equation with "ω" matrix as a subject.

$$A^{-1}A\omega = A^{-1}B$$

$$\omega = A^{-1}B \tag{23}$$

Thus, when rewriting the equation back with matrix notation, the equation becomes:

$$\begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^{n} x_{1i} & \cdots & \sum_{i=1}^{n} x_{ki} \\ \sum_{i=1}^{n} x_{1i} & \sum_{i=1}^{k} (x_{1i})^2 & \cdots & \sum_{i=1}^{n} x_{1i}x_{ki} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} x_{ki} & \sum_{i=1}^{n} x_{1i}x_{ki} & \cdots & \sum_{i=1}^{n} (x_{ki})^2 \end{bmatrix}^{-1} \cdot \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} y_i x_{1i} \\ \vdots \\ \sum_{i=1}^{n} y_i x_{ki} \end{bmatrix} \tag{24}$$

Now, technology can be used to find the value of each coefficient variable to get the equation of the multi-linear regression.

Lagrange Multiplier

Lagrange Multiplier is a method in finding local maxima or minima of any function; however, with a constraint. Hence, it is a method in finding the minimized value of the coefficient of the multi-linear regression equation with restriction equation by using vector. The extrema exist in the $f(x_1, x_2 \ldots x_n)$ lies on the function $g(x_1, x_2 \ldots x_n)$, which is the constraint function. Note that this case focuses on only one constraint. The equation is given by:

$$\nabla f(x_1, x_2 \dots x_n) = \lambda \nabla g(x_1, x_2 \dots x_n) \tag{25}$$

$$g(x_1, x_2 \dots x_n) = C \tag{26}$$

Where $\nabla f(x_1, x_2 \dots x_n)$ is the gradient vector of $f((x_1, x_2 \dots x_n)$, where it is the derived through partial differentiation. Both $\nabla f(x_1, x_2 \dots x_n)$ and $\nabla g(x_1, x_2 \dots x_n)$ shares the same tangent line. "λ" is a scalar dummy variable which shows that both vector gradients are parallel. "C" is the constraint value of g function where the function f must passes through. Moreover, $\nabla g(x_1, x_2 \dots x_n)$ must not equate to 0 [8].

By equate both vector gradients, now it can be solved through technology by simultaneous equations method of $k + 1$ variables and $k + 1$ equations.

Application:

Now, all the methods required to create a model which predicts the point score in the game of a volleyball are explored.

From the investigation, in order to have an equation of the line of best first, the necessary material is the observed data (in this case is the data collected related to each volleyball player). Moreover, these variables must be independent of other variables. Therefore, the table below is the data table which contains the necessary data which going to be used in finding the line model of the line of best fit.

The data table has 6 columns which hold the variables that are required in creating the model. Notably, there are both quantitative and qualitative data. When using, the quantitative variables can be directly used; for instance, age, height, weight and the number of years played. Conversely, the qualitative data are unusable in this format unless the format changes, because regression only considers numbers. As a consequence, the qualitative data are represented in the *Method of Dummy Variables* instead.

The *Method of Dummy Variables* is the use of the integer 0 and 1 as a value for one new variable. In other word, a variable is used to either representing 0 or 1. Hence, the quantity of new variable depends on the option of the observed data. For instance, in volleyball, there are in total of five different positions – Left, Middle, Right, Setter and Libero. Thus, the position data needs 3 new variables to represent it. The reason is 3 variables for the binary system yields eight different options, while 2 variables are for four options.

Therefore, each variable from the collected data is represented by:

Table 1: This table shows the representation of the collected variable in term of the regression variables

Data Variable	Type of Variable	Choices	Regression Variable
Gender	Qualitative	Male → 0	x_{1i}
		Female →1	
Age	Quantitative	-	x_{2i}
Height	Quantitative	-	x_{3i}
Weight	Quantitative	-	x_{4i}
Position	Qualitative	Left → 0, 0, 0	x_{5i}, x_{6i}, x_{7i}
		Middle →0, 0, 1	
		Right → 0, 1, 0	
		Setter → 0, 1, 1	
		Libero → 1, 0, 0	
Years Played	Quantitative	-	x_{8i}

Average Points per Game	Quantitative	-	y_i

After stating all the variables necessary for creating the model, another step can be taken in order to find the value of all the coefficient of each variable. The equation becomes:

$$\sum_{i=1}^{80} {\varepsilon_i}^2 = \sum_{i=1}^{80} (y_i - \beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i} - \beta_3 x_{3i} - \cdots\cdots - \beta_8 x_{8i})^2 \tag{27}$$

Note that is the same as equation 17 where all the variables are substituted in. The variable "k" is substituted as 8, which is the number of the variable for the multi-linear regression. On the other hands, the variable "n" is substituted as 80, which indicates that there are in total of 80 actual collected data used to create the best-fit line model. These 80 data are collected from the volleyball players, worldwide, from the age of 16 to 30 years old, as well as, from the beginner to professional.

Now, without taking constraint into consideration, by using the matrix technique mentioned earlier to find the equation of the line of best fit.

$$\begin{bmatrix} 80 & \sum_{i=1}^{80} x_{1i} & \sum_{i=1}^{80} x_{2i} & \cdots\cdots & \sum_{i=1}^{80} x_{8i} \\ \sum_{i=1}^{80} x_{1i} & \sum_{i=1}^{80} (x_{1i})^2 & \sum_{i=1}^{80} x_{1i}x_{2i} & \cdots\cdots & \sum_{i=1}^{80} x_{1i}x_{8i} \\ \sum_{i=1}^{80} x_{2i} & \sum_{i=1}^{80} x_{1i}x_{2i} & \sum_{i=1}^{80} (x_{2i})^2 & \cdots\cdots & \sum_{i=1}^{80} x_{2i}x_{8i} \\ \vdots & \ddots & \iddots & \iddots & \vdots \\ \vdots & \iddots & \ddots & \ddots & \vdots \\ \sum_{i=1}^{80} x_{8i} & \sum_{i=1}^{80} x_{1i}x_{8i} & \sum_{i=1}^{80} x_{2i}x_{8i} & \cdots\cdots & \sum_{i=1}^{80} (x_{8i})^2 \end{bmatrix}_{(9,9)} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \vdots \\ \beta_8 \end{bmatrix}_{(9,1)} = \begin{bmatrix} \sum_{i=1}^{80} y_i \\ \sum_{i=1}^{80} y_i x_{1i} \\ \sum_{i=1}^{80} y_i x_{2i} \\ \vdots \\ \vdots \\ \sum_{i=1}^{80} y_i x_{8i} \end{bmatrix}_{(9,1)} \tag{28}$$

Hence, use the computer to calculate each element in the matrix:

$$\begin{bmatrix} 80 & 37 & 1775 & \cdots\cdots & 637 \\ 37 & 37 & 815 & \cdots\cdots & 308 \\ 1775 & 815 & 41839 & \cdots\cdots & 16335 \\ \vdots & \ddots & \iddots & \iddots & \vdots \\ \vdots & \iddots & \ddots & \ddots & \vdots \\ 637 & 308 & 16335 & \cdots\cdots & 7439 \end{bmatrix}_{(9,9)} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \vdots \\ \beta_8 \end{bmatrix}_{(9,1)} = \begin{bmatrix} 1171 \\ 492 \\ 27096 \\ \vdots \\ \vdots \\ 10278 \end{bmatrix}_{(9,1)} \tag{29}$$

The result from the computer shows that ω matrix, which contains $\beta_0, \beta_1, \beta_2, \dots, \beta_8$, as:

$$\omega = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \vdots \\ \beta_8 \end{bmatrix}_{(9,1)}$$

$$\omega = \begin{bmatrix} -9.72 \\ -0.112 \\ 0.0433 \\ \vdots \\ \vdots \\ 0.269 \end{bmatrix}_{(9,1)} \tag{30}$$

Where all of these elements in the matrix are the coefficients for the multi-linear regression which helps predict the point in which a player would score in a game. Hence, the equation is:

$$y = -9.72 - 0.112x_1 + 0.0433x_2 + 0.134x_3 + 0.00564x_4 - 13.8x_5 - 5.05x_6 - 0.291x_7 + 0.269x_8 \quad (31)$$

Hence, through substituting the actual value of a player, the average points scored in a game could be predicted. For example, by substituting the following set of x-value into the equation:

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
0	17	185	84.4	0	0	1	3

The outcome of the substitution yields an average of approximately 26 to 27 points scored in a game.

Thus, next will focus on the Multi-Linear Regression application with the constraint. Therefore, the first step in to set a constraint equation for the equation.

As mentioned that the objective of this investigation is to predict the performance of a volleyball player in a game. However, ideally, I designed this investigation to predict my own performance, so I can set my own goal in each game. Hence, the constraint I want to give to the equation I got from the Multiple Linear Regression is my personal statistic. They are all my current and actual information; for instance, my current age, height, and weight. Simply put, the equation must pass through this point. The constraint equation is given by:

$$g(\beta_0, \beta_1, \cdots, \beta_8) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7 + \beta_8 x_8 \quad (32)$$

Where,

Table 2. The coefficient of the constraint equation

y	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
25	0	17	185	84.4	0	0	1	3

Where "y" is the point scored in a game, "$x_1,\ x_2,\ x_3,\ x_4,\ x_5,\ x_6,\ x_7,\ x_8$" are the regression variables. Thus, the constraint equation means, I am able to score an average of 25 points when I 17 years old, 185 cm tall, weight 84.4 kg, and playing middle blocker for 3 years.

Now, use partial differentiation to derive the equation and put the resultant equation in vector notation. The first equation shown is the function of $\nabla g(\beta_0, \beta_1, \cdots, \beta_8)$, which is the constraint function:

$$\nabla g(\beta_0, \beta_1, \cdots, \beta_8) = \begin{pmatrix} 25 \\ 0 \\ 17 \\ 185 \\ 84.4 \\ 0 \\ 0 \\ 1 \\ 3 \end{pmatrix} \tag{33}$$

The second derived equation is the sum of the residual equation, which contains all the variables:

$$\nabla f\ (\beta_0, \beta_1, \cdots, \beta_8) = \begin{pmatrix} -2\sum_{i=1}^{80} y_i + 160\beta_0 + 2\beta_1 \sum_{i=1}^{80} x_{1i} + 2\beta_2 \sum_{i=1}^{80} x_{2i} + \cdots\cdots + 2\beta_8 \sum_{i=1}^{80} x_{8i} \\ -2\sum_{i=1}^{80} y_i x_{1i} + 2\beta_0 \sum_{i=1}^{80} x_{1i} + 2\beta_1 \sum_{i=1}^{80} (x_{1i})^2 + 2\beta_2 \sum_{i=1}^{80} x_{1i}x_{2i} + \cdots\cdots + 2\beta_8 \sum_{i=1}^{80} x_{1i}x_{8i} \\ \vdots \\ \vdots \\ -2\sum_{i=1}^{80} y_i x_{8i} + 2\beta_0 \sum_{i=1}^{80} x_{8i} + 2\beta_1 \sum_{i=1}^{80} x_{1i}x_{8i} + 2\beta_2 \sum_{i=1}^{80} x_{1i}x_{8i} + \cdots\cdots + 2\beta_8 \sum_{i=1}^{80} x_{1i}x_{8i} \end{pmatrix} \tag{34}$$

Hence, $\nabla f(\beta_0, \beta_1, \cdots, \beta_8) = \lambda \nabla g(\beta_0, \beta_1, \cdots, \beta_8)$ is:

$$\begin{pmatrix} -2\sum_{i=1}^{80} y_i + 160\beta_0 + 2\beta_1 \sum_{i=1}^{80} x_{1i} + 2\beta_2 \sum_{i=1}^{80} x_{2i} + \cdots\cdots + 2\beta_8 \sum_{i=1}^{80} x_{8i} \\ -2\sum_{i=1}^{80} y_i x_{1i} + 2\beta_0 \sum_{i=1}^{80} x_{1i} + 2\beta_1 \sum_{i=1}^{80} (x_{1i})^2 + 2\beta_2 \sum_{i=1}^{80} x_{1i}x_{2i} + \cdots\cdots + 2\beta_8 \sum_{i=1}^{80} x_{1i}x_{8i} \\ \vdots \\ \vdots \\ -2\sum_{i=1}^{80} y_i x_{8i} + 2\beta_0 \sum_{i=1}^{80} x_{8i} + 2\beta_1 \sum_{i=1}^{80} x_{1i}x_{8i} + 2\beta_2 \sum_{i=1}^{80} x_{1i}x_{8i} + \cdots\cdots + 2\beta_8 \sum_{i=1}^{80} x_{1i}x_{8i} \end{pmatrix} = \lambda \begin{pmatrix} 25 \\ 0 \\ 17 \\ 185 \\ 84.4 \\ 0 \\ 0 \\ 1 \\ 3 \end{pmatrix} \quad (35)$$

By using the actual data collected, the equation can be simplified into the following, where it consisted of the one vector equation and another constraint equation.

$$\begin{pmatrix} -2342 + 160\beta_0 + 74\beta_1 + 3550\beta_2 + 28390\beta_3 + \cdots\cdots + 1274\beta_8 \\ -984 + 74\beta_0 + 74\beta_1 + 1630\beta_2 + 12618\beta_3 + \cdots\cdots + 616\beta_8 \\ -54192 + 3550\beta_0 + 1630\beta_1 + 83678\beta_2 + 634608\beta_3 + \cdots\cdots + 32670\beta_8 \\ -423018 + 28390\beta_0 + 12618\beta_1 + 634608\beta_2 + 5058854\beta_3 + \cdots\cdots + 229934\beta_8 \\ -168658 + 11142\beta_0 + 4568\beta_1 + 252210\beta_2 + 1994970\beta_3 + \cdots\cdots + 92614\beta_8 \\ -30 + 14\beta_0 + 8\beta_1 + 310\beta_2 + 2312\beta_3 + \cdots\cdots + 124\beta_8 \\ -562 + 48\beta_0 + 24\beta_1 + 1030\beta_2 + 8342\beta_3 + \cdots\cdots + 364\beta_8 \\ -1076 + 66\beta_0 + 28\beta_1 + 1516\beta_2 + 11900\beta_3 + \cdots\cdots + 576\beta_8 \\ -20556 + 1274\beta_0 + 616\beta_1 + 32670\beta_2 + 229934\beta_3 + \cdots\cdots + 14878\beta_8 \end{pmatrix} = \lambda \begin{pmatrix} 25 \\ 0 \\ 17 \\ 185 \\ 84.4 \\ 0 \\ 0 \\ 1 \\ 3 \end{pmatrix} \quad (36)$$

$$25 = \beta_0 + 17\beta_2 + 185\beta_3 + 84.4\beta_4 + \beta_7 + 3\beta_8 \quad (37)$$

By using technology, the coefficient of each variable can be found as shown:

Table 3. The coefficient obtained from the multiple-linear regression with constraint (rounded to 4 decimal places)

Variables	Value	Variables	Value
β_0	141.2426	β_5	-31.5675
β_1	-9.0291	β_6	-13.0770
β_2	-0.4233	β_7	0.0138
β_3	-0.7169	β_8	1.3854
β_4	0.2299	λ	0.7100

From Table 3, each coefficient of the constrained score predicting function has its own meaning. The meaning is in such a way that if any x increases by 1 unit, the average point score per game will change by the value of the coefficient of x-variable. For instance, if my height increases by 1 cm, my average point scored in one game will decrease by 0.7169 points.

The accuracy of the equation, obtained through all the previous methodologies, can be verified for the accuracy through substituting the constraint value. The reason is due to the nature of the constraint equation, where the function must pass through all the constraint points. Thus, the outcome of the substitution must yield the average point scored per game of the constraint.

$$y = 141.2426 + (-9.0291)(0) + (-0.4233)(17) + (-0.7169)(185) + \cdots + (0.7100)(3) \quad (38)$$

$$y = 24.99356$$

$$y \approx 25$$

From equation 38, the outcome of y is approximately 25, which is the equivalent to the value of constraint for average points scored in a game. The reason that it is approximated around 25 is that the coefficients are all rounded to 4 decimal places.

By substituting the same set of value into x-variable for the multiple linear regression equation, differences can be seen. In which the multiple linear regression equation yields,

$$y = -9.72 + (-0.112)(0) + (0.0433)(17) + (0.134)(185) + \cdots + (0.269)(3) \quad (39)$$

$$y = 16.798116$$

$$y \approx 17$$

The outcome from the Multi-Linear regression is approximately 17, while for the Multi-Linear regression with constraint yields approximately 25 for the same x-value input.

Conclusion:

Using all the method discussed above, the multi-linear regression is able to produce an equation with multiple independent to describe only the linear relationship between those predicted value and the actual value. It could be applied in the scenario, such as this equation that predicts the performance of a player in volleyball.

Even though this method is able to produce an equation, as a best-fit line, to describe the relationship between variable, but it might not be the most accurate best fit line which best describes the formula. Hence, the further approach to creating the most accurate equation that could, explicitly, used in the real-life situation is by acquiring more sample data. By doing this, the real data used to solve for the coefficient could be more accurate and based more on the trend of the sample. Notwithstanding the number of constraints used in the investigation and application might be too small. The reason is, in reality, and more professional world, there could be more factors which could impact the performance of a player than the one selected, such as, the play time or injury. In addition, to get the most accurate trend line, in case the relationship is non-linear, instead of using the multi-linear regression, the approach should be taken through the multi-nonlinear regression. However, this approach, in my opinion, is very challenging and will consume a lot more time. Still, a product of this approach sounds intriguing.

Appendix

1. Program used to calculate from MatLab

a.

```
>> syms k x y z
eqns = [100 == 2*y*z + 2*x*z + 2*x*y, x*y == k*(2*x + 2*y), x*z == k*(2*x + 2*z), y*z == k*(2*y + 2*z)];
vars = [k x y z];
[solk, solx, soly, solz] = solve(eqns, vars)
```

b.

```
solk =

 -(5*6^(1/2))/12
  (5*6^(1/2))/12
```

```
solx =

 -(5*6^(1/2))/3
  (5*6^(1/2))/3
```

```
soly =

 -(5*6^(1/2))/3
  (5*6^(1/2))/3
```

```
solz =

 -(5*6^(1/2))/3
  (5*6^(1/2))/3
```

2. Data Table

Player	Gender	Age	Height (cm)	Weight (kg)	Position(x 5)	Position (x6)	(Position x7)	Years Played	Average Point per game
1	0	17	184	86	0	0	1	4	25
2	0	17	180	65	0	0	0	3	14
3	1	18	149	44	1	0	0	3	2
4	1	17	167	60	0	0	1	2	22
5	0	17	173	70	0	1	0	3	15
6	0	17	179	58	0	0	1	3	10
7	0	16	184	93	0	1	1	3	5
8	0	15	173	64	0	0	0	1	8
9	0	17	174	63	0	0	0	4	20
10	1	18	167	60	0	0	1	3	11
11	0	17	182	65	0	0	0	2	13
12	0	17	170	60	0	0	1	2	18
13	1	17	157	53	0	0	0	6	10
14	1	17	155	47	0	0	0	7	13
15	0	17	168	55	0	1	1	1	2
16	0	27	194	101	0	0	0	15	15
17	0	30	183	73	0	1	1	18	21
18	0	24	199	89	0	0	1	10	29
19	0	28	188	85	1	0	0	16	3
20	0	16	170	63	1	0	0	2	2
21	1	19	165	60	0	1	1	8	7
22	0	19	175	69	0	1	0	3	16
23	0	29	204	87	0	0	0	15	28
24	1	19	164	78	0	1	0	6	14
25	1	23	190	70	0	0	1	10	25
26	1	35	180	67	0	0	1	19	22
27	1	19	174	65	0	1	0	12	16
28	1	33	169	57	0	1	1	20	8
29	1	29	171	62	1	0	0	17	3
30	0	17	185	60	0	0	1	3	22
31	1	16	153	50	0	1	0	3	13
32	0	18	170	74	0	1	0	2	17
33	1	28	180	71	0	1	1	15	9
34	0	31	194	89	0	0	0	19	24
35	0	25	201	92	0	0	1	10	26
36	0	23	187	85	0	0	1	13	18
37	1	20	179	57	0	0	0	8	17
38	1	14	161	45	1	0	0	5	1
39	1	16	176	63	0	1	0	6	12
40	0	19	190	67	0	0	0	2	14
41	1	10	154	43	0	0	0	1	11
42	0	32	187	86	0	1	0	17	20
43	1	27	194	73	0	0	1	14	15
44	0	26	188	82	0	0	0	6	14
45	1	25	181	72	0	0	0	7	21
46	1	28	184	69	0	0	1	13	22
47	0	22	193	77	0	0	1	8	17
48	1	23	178	59	0	1	0	9	14
49	1	27	176	60	0	0	0	12	20
50	0	22	183	76	0	0	0	8	18
51	0	18	179	72	0	0	0	4	15
52	1	16	172	50	0	0	1	5	18
53	1	19	177	65	0	1	0	2	7
54	1	18	158	54	0	1	1	3	6
55	1	17	167	69	0	0	1	2	8
56	0	33	184	70	0	0	1	18	18
57	1	29	170	69	0	0	0	11	17
58	0	30	190	88	0	0	0	10	23
59	0	28	187	74	0	1	0	15	20
60	0	24	179	89	0	1	1	8	10
61	0	25	165	80	1	0	0	9	2
61	0	25	165	80	1	0	0	9	2
62	0	18	182	64	0	0	0	2	12
63	1	28	174	68	0	0	1	13	22
64	0	27	184	74	0	0	1	11	24
65	1	24	170	65	0	1	0	9	13
66	1	26	178	77	0	0	0	10	17
67	1	30	181	73	0	0	0	15	15
68	1	25	152	47	1	0	0	10	2
69	0	27	176	80	0	1	1	13	9
70	0	18	186	83	0	0	0	2	16
71	0	26	193	89	0	0	1	13	19
72	0	28	188	73	0	0	0	12	18
73	1	25	175	81	0	0	0	10	20
74	0	18	183	78	0	0	1	2	17
75	0	17	176	81	0	1	0	2	7
76	1	19	173	59	0	0	0	3	10
77	0	18	189	78	0	1	1	1	9
78	1	18	163	61	0	1	0	3	11
79	0	30	187	80	0	0	1	14	26
80	1	23	175	61	0	0	1	6	18

3. Matrix Elements

	x0	x1	x2	x3	x4	x5	x6	x7	x8	y1
x0	80	37	1775	14195	5571	7	24	33	637	1171
x1	37	37	815	6309	2284	4	12	14	308	492
x2	1775	815	41839	317304	126105	155	515	758	16335	27096
x3	14195	6309	317304	2529427	997485	1156	4174	5950	114967	211509
x4	5571	2284	126105	997485	400785	426	1670	2355	46307	84329
x5	7	4	155	1156	426	7	0	0	62	15
x6	24	12	515	4174	1670	0	24	10	182	281
x7	33	14	758	5950	2355	0	10	33	288	538
x8	637	308	16335	114967	46307	62	182	288	7439	10278

4. Coefficient List

	Coefficients
Intercept	-9.724937852
Gender	-0.111921414
Age	0.043339368
Height (cm)	0.133921262
Weight (kg)	0.005636726
Position(x5)	-13.87304679
Position (x6)	-5.045839573
(Position x7)	-0.290646251
Years Played	0.269383101

5. Vector Elements

	y	x0	x1	x2	x3	x4	x5	x6	x7	x8
x0	-2342	160	74	3550	28390	11142	14	48	66	1274
x1	-984	74	74	1630	12618	4568	8	24	28	616
x2	-54192	3550	1630	83678	634608	252210	310	1030	1516	32670
x3	-423018	28390	12618	634608	5058854	1994970	2312	8342	11900	229934
x4	-168658	11142	4568	252210	1994970	801570	852	3340	4710	92614
x5	-30	14	8	310	2312	852	14	0	0	124
x6	-562	48	24	1030	8342	3340	0	48	20	364
x7	-1076	66	28	1516	11900	4710	0	20	66	576
x8	-20556	1274	616	32670	229934	92614	124	364	576	14878

6. Formula and Solution for Lagrange Multiplier

```
syms a b c d e f g h i z
eqns = [25 == a + 17*c + 185*d + 84.4*e + h + 3*i, 25 == -2342 + 160*a*z + 74*b*z + 3550*c*z + 28390*d*z + 11142*e*z +
14*f*z + 48*g*z + 66*h*z + 1274*i*z, 0 == -984 + 74*a*z + 74*b*z + 1630*c*z + 12618*d*z + 4568*e*z + 8*f*z + 24*g*z +
28*h*z + 616*i*z, 17 == -54192 + 3550*a*z + 1630*b*z + 83678*c*z + 634608*d*z + 252210*e*z + 310*f*z + 1030*g*z +
1516*h*z + 32670*i*z, 185 == -423018 + 28390*a*z + 12618*b*z + 634608*c*z + 5058854*d*z + 1994970*e*z + 2312*f*z +
8342*g*z + 11900*h*z + 229934*i*z, 84.4 == -168658 + 11142*a*z + 4568*b*z + 252210*c*z + 1994970*d*z + 801570*e*z +
852*f*z + 3340*g*z + 4710*h*z + 92614*i*z, 0 == -30 + 14*a*z + 8*b*z + 310*c*z + 2312*d*z + 852*e*z + 14*f*z + 0*g*z +
0*h*z + 124*i*z, 0 == -562 + 48*a*z + 24*b*z + 1030*c*z + 8342*d*z + 3340*e*z + 0*f*z + 48*g*z + 20*h*z + 364*i*z, 1 ==
-1076 + 66*a*z + 28*b*z + 1516*c*z + 11900*d*z + 4710*e*z + 0*f*z + 20*g*z + 66*h*z + 576*i*z, 3 == -20556 + 1274*a*z +
616*b*z + 32670*c*z + 229934*d*z + 92614*e*z + 124*f*z + 364*g*z + 576*h*z + 14878*i*z];
vars = [a b c d e f g h i z];
[solva, solvb, solvc, solvd, solve, solvf, solvg, solvh, solvi, solvz] = solve(eqns, vars);
dsol = double([solva, solvb, solvc, solvd, solve, solvf, solvg, solvh, solvi, solvz])
```

Work Cited

1. *Big Data Analytics*. IBM, 11 Mar. 2012, https://www.ibm.com/analytics/hadoop/big-data-analytics.
2. *Lesson 1: Simple Linear Regression | STAT 501*. https://newonlinecourses.science.psu.edu/stat501/node/250/. Accessed 22 Feb. 2019.
3. "What Is Linear Regression?" *Statistics Solutions*, https://www.statisticssolutions.com/what-is-linear-regression/. Accessed 22 Feb. 2019.
4. James, Stewart. *Calculus*. Eight, 2015.
5. Steven J. Miller. *The Method of Least Squares*. Mathematics Department Brown University, 12 May 2013, https://web.williams.edu/Mathematics/sjmiller/public_html/BrownClasses/54/handouts/MethodLeastSquares.pdf.
6. Dan Campbell , and Sherlock Campbell . *Introduction to Regression and Data Analysis* . Statlab Workshop , 28 Oct. 2008, http://statlab.stat.yale.edu/workshops/IntroRegression/StatLab-IntroRegressionFa08.pdf.
7. *What Is the Difference between Implicit Differentiation and Partial Differentiation? - Quora*. https://www.quora.com/What-is-the-difference-between-implicit-differentiation-and-partial-differentiation. Accessed 22 Feb. 2019.
8. *The Method of Lagrange Multipliers - Mathonline*. http://mathonline.wikidot.com/the-method-of-lagrange-multipliers. Accessed 22 Feb. 2019.