

Impact of Anthropometric Factor on Serve Velocities Evaluation Tennis Match at The 2019 US Open Grand Slam Tournament

Si Putu Pandit Dharma^a, Hari Setijono^b, Edy Mintarto^{c*}

^aPostgraduate Student at Sport Science and Education, Universitas Negeri Surabaya, Surabaya, 60213, Indonesia

^bProfessor at Sport Science and Education, Universitas Negeri Surabaya, Surabaya, 60213, Indonesia

^cLecturer at Sport Science and Education, Universitas Negeri Surabaya, Surabaya, 60213, Indonesia

Correspondence: panditdharma@yahoo.com

Received: 24 Sept 2019 **Accepted:** 01 Mei 2020 **Published:** 04 Mei 2020

Abstract

The purposes of this study are to analyze (a) anthropometric factor as multivariate regression model to predict the serves velocities of elite tennis player that compete in US Open Grand Slam 2019. The anthropometric factor was describe as, player height, player weight, age, and the Body Mass Index (BMI) of player. (b) this study also determind the significant level from the model to predict the serves speed. The data were collected from MATCH DETAILS BY IBM SLAMTRACKER. Results show that correlation between each independent variable (i.e, age, height, weight, BMI) to serves speed of the player respectively are -0.0094, 0.7457, 0.7135, and 0.1944. The data show us that variable Height and Weight have the most correlated predictors of tennis serves speed. More over the correlation level (R^2) from the model was 0.5767 or 57.67 % it indicates that the model can predict 57.67% the dependent variable, i.e serve speed and the other 42.33% was determind by the other factor not included in model. The present finding underlines the importance of player height and weight to determind the serves speed of the players that play in US Open Grand Slam Tournament 2019.

Keywords: Anthropometric factor; Serve velocities; IBM SLAMTRACKER; multivariate regression model.

1. Introduction

Tennis is a sport with no such certainty of time. This sport needs players to able to play with the intermittent game rhythm. So the predominant factor to predict the performance from the athletes is high stroke and serve velocities and it requiring high physical demands (Ulbricht, Fernandez-Fernandez, et al., 2015). The velocities of tennis stroke yet also a predominant factor determining the performance levels of elite tennis player (Kovacs & Ellenbecker, 2011). Recent study shows that serve is the most important stroke in tennis match (Elliott et al., 2013). Event serves has become a main factor in determining elite tennis player performance, but what is the factor to determining serves performance itself it's still a big discussion, and still very interesting field of testing and study (Ma et al., 2013) state that serves velocities are factors to determining player winning and losing in grand slams match such as Roland Garros. Anthropometric is also can become predictors factor to serves velocities (Ma et al., 2013). (Palmer et al., 2018) study that serve speed velocity is influenced by range of motion, strength, motor control, and power. And in these researches an anthropometric factor such as height, an age is under consideration. A Multivariate regressions model is used to predict one dependent variable. In this present studies anthropometric variables such as Age, Height, Weight, and BMI are used to predict the average velocities of serves from elite tennis athletes that compete in the round of 3rd Roland Garros Grand Slams 2017. (Ulbricht, Fernández-

Fernández, et al., 2015) in their research suggest that fitness characteristic is also hold an important role in determine serve performance. When using multivariant regression model there is 3 major pattern in describing the multivariant regression model (Irianto, 2010):

1. First, each of the independent variables are stand-alone when predicting the dependent variable, in this pattern, the independent variable has a low correlation between each other and the correlation coefficients (R^2) are the sum of each R^2 from the variables.
2. Each independent variable is not stand alone when predicting the dependent variable, in this pattern if the R^2 between each independent variable is big then its contribution to predicting the dependent variable is small.
3. When the independent variable has an endogenous or exogenous variable to predict the dependent variable.

Linear Regression model according to Santoso (2012) in (Mahardika, 2015) state that there are five assumption to follow: (1) Normality, (2) Multicollinearity, (3) Homocedasticity, (4) Autocorrelation, (5) and Linearity.

This present study used the second assumption of correlation to predict the dependent variable. In the second assumption, we need to calculate each predictor (independent variables) coefficient of correlation to the dependent variable, then we need to calculate the coefficients correlation from the model that includes all the independent variables.

Table 1. Coefficients Correlation Interpretation

Coefficients correlation interval	Interpretation
± 0.70 to 1.00	High
± 0.40 to 0.69	Moderate
± 0.20 to 0.39	Low
Less than 0.20	Neglected

Table 1. Young (1982) in (Mahardika, 2015):297 state that the interpretation of model ones should follow the terms:

- Sample must be taken in randomly order from a population that has a normal distribution
- Sample must be in normally distributed
- The dependent variable must have an interval or ratio scale.
- The independent and dependent variable must have theoretically correlation
- The regressions model must have linearity nature.

The multivariant regressions formula for two independent variables are

$$\hat{Y} = a + b_1X_1 + b_2X_2$$

next formula to calculate the coefficient regression:

$$\begin{aligned} \sum Y &= an + b_1 \sum X_1 + b_2 \sum X_2 \\ \sum X_1Y &= a \sum X_1 + b_1 \sum X_1^2 + b_2 \sum X_1X_2 \\ \sum X_2Y &= a \sum X_2 + b_1 \sum X_1X_2 + b_2 \sum X_2^2 \end{aligned}$$

Table 2. F_{test} Formula

Source of Variance	Df	SS	MS	F
Regression model b/a	k	$b_1 \sum X_1 Y + b_2 \sum X_2 Y$	$\frac{SS_{b/a}}{k}$	$\frac{MS_{b/a}}{MS_{residual}}$
Residual	N-k	$\sum (Y - \hat{Y})^2$	$\frac{SS_{residual}}{(n - k - 1)}$	

Furthermore this present study also tests if the regression model is significant as predictors, this use F_{test} according to (Irianto, 2010), see table 2. Then to go further to make sure that the predictors are valid and to understand the contribution of the predictor this study also calculates the coefficient correlation percentage in a formula.

$$R^2 = \frac{SS_{b/a}}{\sum Y^2}$$

After the coefficient of correlation was calculated then this present study also conducting a test for the significant of the correlation coefficient using the F_{test} as suggested by Irianto (2010).

$$F = \frac{R^2/k}{(1 - R^2)/(n - k - 1)}$$

Thus, this present study aimed was to understand the predictors model for the tennis serves velocities and how good is the model can predict the outcome. By using the anthropometric factors as predictors.

2. Research Method

The data were collected from MATCH DETAILS BY IBM SLAMTRACKER. The data give information about detailed players anthropometric detail such as age, height, and weight. BMI also calculated from the formula gives by weight divided with square height, from MATCH DETAILS BY IBM SLAMTRACKER this present studies also collected the average serves speed from each player's in each round until round 3. This present study also doing ANOVA test to get understand whether there's a difference between the mean of serves speed in each round for each of the players. This present study also conducted a normality test using two different types, first visual test by plotting the data, both for the dependent and independent variables. The second normality test was using the Shapiro Wilk test, because this test can accommodate *p-value* even though there are ties in the data. R Studio software was used to help with all of the calculations. Why uses R Studio to help the calculation, it is because R Studio is a open source software, so that we can avoid piracy issue on the research. R studio are widely used by professional in data science research (Hardley, 2020). (Marcous, 2020) in his article say that R studio can make the job a lot easier and can jump into the market in very short time. (Fox & Leanage, 2016) state in their research that R software is one of open source statistic software that can handle from easy to complicated data processing. But this present study also uses Microsoft Excel to make the calculation more accurate. And also manual calculation to ensure that the result was not contradicting.

3. Result

With the data presenting, this present studies first conducting normality test for the data, this test is important because it can determine what statistical uses to analyze the data later on whether it's going to use parametric statistic if the data are normally distributed, or use Nonparametric if the data not normally distributed. This present study uses a visual test using a plot diagram to plot each variable. The principal plot method is to compare the data with a factor that in normal distribution this present studies used R Studio to help plot the data. Figure 1. shows the normality plot from the Age of the athlete. Figure 2. Show normality plot from Height. Figure 3. Show normality plot from Weight. And then figure 4. Show normality plot from BMI.

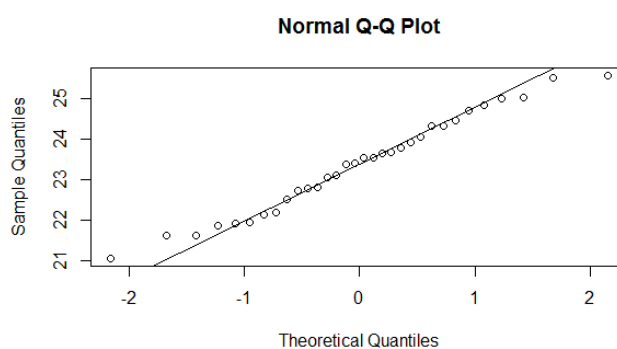


Figure 1. Normality Plot ~ Age

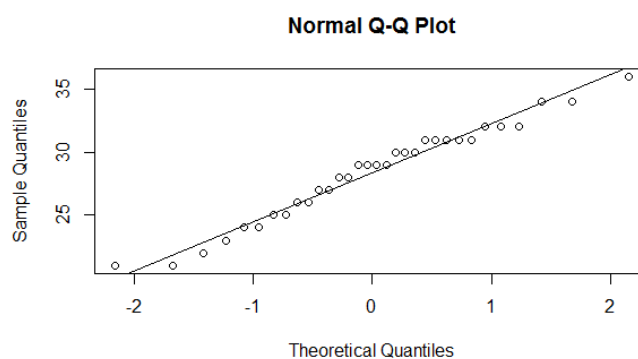


Figure 2. Normality Plot ~ Height

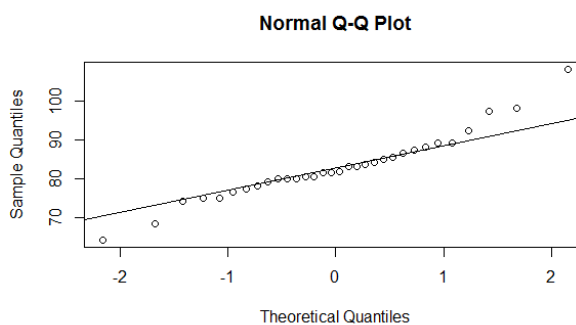


Figure 3. Normality Plot ~ Weight

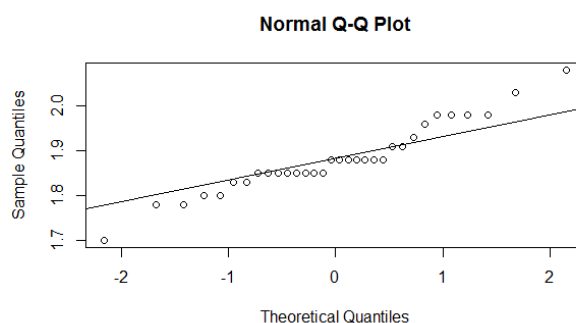


Figure 4. Normality Plot ~ BMI

Table 3. Anthropometrics Characteristics of US Open Grand Slams Tennis 2019

	Men		
	Mean	Standard deviation	Variance
Age	28.38	3.85	14.82
Heights	1.89	0.08	0.001
Weights	82.93	8.50	72.39
Body BMI	23.37	1.21	1.47

This present study also using ANOVA test for independent variables to see whether there is a difference in mean between serves speed in each round. The results indicated that a p-value of 0.9959 higher than $\alpha = 0.05$, this finding suggests accepting the null hypothesis which brings the conclusion to state that there is no difference in the mean in each round, therefore this present study uses the total average speed as the dependent variable. The conclusion also matches when we used F test, the F test result is 0.0041 is significantly lower than F table with df 93 which is 3.09. therefore accept the null hypothesis at the level significant 0.05. describing various elements of anthropometric from the elite players that play in the US Open Grand Slam 2019 Tournament. Table 3 presenting the descriptive statistic from the data such as mean, variance, and standard deviation. Also this present studies entering the BMI results and interpretation to the table And it shows complete anthropometric data from the descriptive characteristic. the results from the table show us that mean of the player heights is the lower compare to the mean of Weights, Age, and Body BMI which is 1.89 instead of respectively 82.93, 28.38, and 23.37. this can happen because of the presence of outliers data in other anthropometric characteristics except for heights. Because the calculation of the mean is also considering lower limit and upper limit from the data. In height for example with the mean value of 82.93, there are outliers in the data the players with the lower weights is SCHWARTZMAN, with weight in 64.10 kg then the heaviest player is ISNER with the weight of 108.2 kg this factor can make the mean of the data very large. The standard deviation from the mean tells us how tightly various of the sample are close to it's mean, so the lower the value of standard deviation shows that the closest the spreading of the data from it's mean. This present studies indicate that height had the smallest standard deviation that's indicating player's spread height is small then suggesting it is the same with the value of player's mean. While weights standard deviation is the higher among the data, it show's that the mean can not represent the average of the weight from players. It's consistent with the finding that variance as a dispersion of the data show's higher value among the data which is 72.39. (consistent with its standard deviation) and height also consistent with the standard deviation which is the lowest value in the data 0.001. while age and body BMI had also low in value in variance, respectively 14.87 and 1.40.

Table 4. Serve Velocities

	Mean	Variance	Standard deviation
Round 1	169.3	95.47	9.71
Round 2	170.1	86.70	9.32
Round 3	170.1	89.55	9.46
Average Serve Speed	170.0	82.61	9.08

Next table 4. shows the value of average speed from the player during 3 round matches in Roland Garros Grand Slams 2017. The data indicates that. Mean from the average speed from round 1, round 2, and round 3 respectively 169.3, 170.1, 170.0 in a glance it's not so different but because the present data from BAUTISTA-AGUT 154.67 km/hour and the most accelerated serve velocities from ISNER 187.33 km/hour. That's making the variance have a higher value on average is 82.61. even though the variance has a high value, but because in the real data the other data besides outliers had a very small dispersion. To create a valid regression model for a multivariant variable, first, we need to test the coefficient correlation between the independent variable, the result is presented in Table 5.

Table 5. Coefficients Correlation Between Each Independents Variable

	Age	Height	Weight	BMI
Age		0.0382	0.1056	0.1524
Height			0.8575	0.5778
Weight	0.1056			0.5778
BMI	0.1524	0.0804		

The result indicates there's a low correlation between age and weight with value 0.03, next correlation between age and weight is the low correlation with value 0.1056, next age and BMI the results low correlation next only variable height ≈ Weight had high correlation, other variable Height ≈ BMI, and Weight ≈ BMI indicate a moderate correlation.

Next table 6. shows the correlation between independent and dependent variables Because the coefficient correlation between age, body BMI and the average speed is relatively low -0.00949 and 0.1944 respectively, then these studies neglected those variables from the multivariant regression model. Only height and weight from the athlete are used as predictors to the dependent variable, 0.7456 and 0.7135 respectively.

Table 6. Coefficient Correlation between Each Independent Variable to Dependent Variable

Dependent variable	Independent variable			
	Age	Height	Weight	BMI
Average Serve Speed	0.00949	0.7456	0.7135	0.1944

4. Discussion

After understanding the correlation between each of the variables, both independent and dependent, these present studies attempt to use the independent variable that has a higher coefficient correlation value as predictors to the dependent variables. Based on table 6. data, present studies concluded that only height and weight are relevant as predictors to the dependent variables. Next using the formula,

$$\hat{Y} = a + b_1X_1 + b_2X_2$$

This present study calculated the formula, and find that the regression model is intercept which represents a with the value of 34.8845, and b_1 represents height variable 58.6185, and b_2 represents the weight is 0.2991. then the result also suggests that the model is following normal distribution this after present studies plotting residual value from the equation with the minimal value of -11.4394, 1Q -2.4265 median -0.4883, 3Q is 4.4467, and maximal value of 9.8524. Plot results from R Studio on figure 5. indicate that the formula residual plots look like a bell curve.

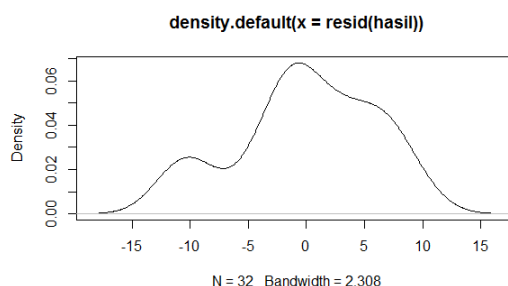


Figure 5. Residual Density Plot

In table 7. A Highlighted result from the regression model, furthermore the '*' star indicates that there is a low significant level in the variable heights therefore high relationship exists between height with the avespeed as the dependent variable. This is also matched with the suggestion of the R^2 which is 0.5767 or regression model can predict 57.67% the cause of the serve average speed.

Table 7. Regression Formula Variable

Coefficients	Estimate	Standart E.	t value	Pr (> t)
(intercept)	34.8845	35.0929	0.994	0.3284
Height	58.6185	27.2310	2.153	0.0398 *
Weight	0.2991	0.2509	1.192	0.2429

Finding in these recent study indicate that an anthropometric factor is used to make better understanding in tennis serve for the coaches (Whiteside et al., 2015). (Ulbricht et al., 2016) fitness characteristic are hold an important role in tennis performace game. Anthropometric factor such as height, weight, and body mass index are highly involved to determine the athlete performance in tennis sport.

These study also support the finding from previous studies that an anthropometric factors heights and weights are both significant contributors to tennis serve speed velocity (Bonato et al., 2015). (Söğüt, 2017) find that age and gender are also part of conderation as contributors on tennis serve velocity. Other research find that body mass index and full body kinematics are significant contributors for the tennis serve velocity (Wong et al., 2014). (Baiget et al., 2016) tennis serve are the

most important shot in competitive tennis. And Age, Body height, body mass are the base to select participant during the research to determine tennis serve speed velocity.

The main finding of this study was that player heights and weights was strongly related with the tennis serve velocity. the multivariate model uses to predict the tennis serve velocity have showed that player height and weight contributed 57.6% to the tennis serve velocity. but there some limitation in this study. First in determining tennis serve velocity in would be necessary to introduce player skill to calculation (Kolman et al., 2019). Another factor are worth for consideration is the psychological factor, such as anxiety, emotion control, and contextual information (Cocks et al., 2016). Gender as one of anthropometric factor in different studies shows as contributor in serve speed velocity (Cross, 2014). Another factor that should be introduce in presence of tennis serve velocity is matchplay characteristic (Reid et al., 2016) these research find that sex-related specific training is needed to improve tennis stroke performace such as first serve and serve return.

5. Conclusion and Recommendation

There were not all the anthropometric variables had significant levels to the average serves speed of the elite tennis player. Our data suggest that age and body BMI only contribute -0.00949 and 0.1944 respectively, so only height and weight of the player had a high correlation with the average serves speed form the player with the value of 0.7456 and 0.7135. therefor the linear model neglected the age and body BMI from the equation. Our data also suggest that the significance of the coefficient correlation from the model is 57.67% can predict the cause of the serve speed, but the other 42.33% was caused by another factor outside the model. These studies also find very low in standard deviation and variance 0.08 and 0.0001 respectively in heights of the player participating in this US Open Grand Slams with the mean of 1.89 m. Present study indicate that beside antropometric factor there are other factor can be affect to serve velocity, such as player skill both technical and tactical (Kolman et al., 2019). Concerning other physical characteristics in tennis this finding suggests that player heights should be put into consideration for professional tennis athletes.

References

- Baiget, E., Corbi, F., Fuentes, J. P., & Fernández-Fernández, J. (2016). The Relationship between Maximum Isometric Strength and Ball Velocity in the Tennis Serve. *Journal of Human Kinetics*, 53(1), 63–71. <https://doi.org/10.1515/hukin-2016-0028>
- Bonato, M., Maggioni, M. A., Rossi, C., Rampichini, S., La Torre, A., & Merati, G. (2015). Relationship between anthropometric or functional characteristics and maximal serve velocity in professional tennis players. *Journal of Sports Medicine and Physical Fitness*.
- Cocks, A. J., Jackson, R. C., Bishop, D. T., & Williams, A. M. (2016). Anxiety, anticipation and contextual information: A test of attentional control theory. *Cognition and Emotion*, 30(6), 1037–1048. <https://doi.org/10.1080/02699931.2015.1044424>
- Cross, R. (2014). *Men's Tennis vs Women's Tennis*.
- Elliott, B., Reid, M., Crespo, M., & Kwon, Y. H. (2013). Editorial. *Sports Biomechanics*, 12(1), 1. <https://doi.org/10.1080/14763141.2013.772379>
- Fox, J., & Leverage, A. (2016). R and the Journal of Statistical Software. *Journal of Statistical Software*, 73(2). <https://doi.org/10.18637/jss.v073.i02>
- Hardley, M. (2020). *Making Research Data Visible at Oxford*. Rstudio.Com. <https://rstudio.com/about/customer-stories/university-of-oxford/>
- Irianto, A. (2010). *Statistik: Konsep Dasar Aplikasi dan Pengembangannya* (P. Media (ed.)). Jakarta Kencana.
- Kolman, N. S., Kramer, T., Elferink-Gemser, M. T., Huijgen, B. C. H., & Visscher, C. (2019). Technical and tactical skills related to performance levels in tennis: A systematic review. *Journal*

- of Sports Sciences*, 37(1), 108–121. <https://doi.org/10.1080/02640414.2018.1483699>
- Kovacs, M. S., & Ellenbecker, T. S. (2011). A performance evaluation of the tennis serve: Implications for strength, speed, power, and flexibility training. *Strength and Conditioning Journal*, 33(4), 22–30. <https://doi.org/10.1519/SSC.0b013e318225d59a>
- Ma, S. M., Liu, C. C., Tan, Y., & Ma, S. C. (2013). Winning matches in Grand Slam men's singles: An analysis of player performance-related variables from 1991 to 2008. *Journal of Sports Sciences*, 31(11), 1147–1155. <https://doi.org/10.1080/02640414.2013.775472>
- Mahardika, I. M. S. (2015). *Metodologi Penelitian* (UNESA University Press Surabaya (ed.); Edisi 1). UNESA University Press Surabaya.
- Marcous, D. (2020). *Outsmarting Traffic with R & Shiny*. Rstudio.Com. <https://rstudio.com/about/customer-stories/waze/>
- Palmer, K., Jones, D., Morgan, C., & Zeppieri, G. (2018). Relationship Between Range of Motion, Strength, Motor Control, Power, and the Tennis Serve in Competitive-Level Tennis Players: A Pilot Study. *Sports Health*, 10(5), 462–467. <https://doi.org/10.1177/1941738118785348>
- Reid, M., Morgan, S., & Whiteside, D. (2016). Matchplay characteristics of Grand Slam tennis: implications for training and conditioning. *Journal of Sports Sciences*, 34(19), 1791–1798. <https://doi.org/10.1080/02640414.2016.1139161>
- Söğüt, M. (2017). A comparison of serve speed and motor coordination between elite and club level tennis players. *Journal of Human Kinetics*, 55(1), 171–176. <https://doi.org/10.1515/hukin-2017-0015>
- Ulbricht, A., Fernandez-Fernandez, J., Mendez-Villanueva, A., & Ferrauti, A. (2015). The relative age effect and physical fitness characteristics in German male tennis players. *Journal of Sports Science and Medicine*, 14(3), 634–642.
- Ulbricht, A., Fernandez-Fernandez, J., Mendez-Villanueva, A., & Ferrauti, A. (2016). Impact of Fitness Characteristics on Tennis Performance in Elite Junior Tennis Players. *Journal of Strength and Conditioning Research*. <https://doi.org/10.1519/JSC.0000000000001267>
- Ulbricht, A., Fernández-Fernández, J., Mendez-Villanueva, A., & Ferrauti, A. (2015). Impact Of Fitness Characteristic On Tennis Performance in Elite Junior Tennis Player. *Journal of Strength and Conditioning Research*, 30(4), 989–998. <https://doi.org/doi:10.1519/jsc.0000000000001267>
- Whiteside, D., Elliott, B. C., Lay, B., & Reid, M. (2015). Coordination and variability in the elite female tennis serve. *Journal of Sports Sciences*. <https://doi.org/10.1080/02640414.2014.962569>
- Wong, F. K. H., Keung, J. H. K., Lau, N. M. L., Ng, D. K. S., Chung, J. W. Y., & Chow, D. H. K. (2014). Effects of body mass index and full body kinematics on tennis serve speed. *Journal of Human Kinetics*, 40(1), 21–28. <https://doi.org/10.2478/hukin-2014-0003>