MUNISI: Military Mathematics and Natural Sciences
e-ISSN: xxxx-xxxx; p-ISSN: xxxx-xxxx
Vol. 1, No. 1, June 2023, pp. 35-46
Available online at https://jurnal.idu.ac.id/index.php/munisi

Longitudinal Data Modeling on Physical Fitness, Organizational Activity, Gender, and its Effect on Student Grade Point Average (GPA)

Ardyan Lazuardy[1], Jihan Anavisha[1], Nawal Zuhdaniel Aflah[1], Ro'fah Nur Rachmawati[2]*
[1] Department of Chemistry, Faculty of Military Mathematics and Natural Sciences, Republic of Indonesia Defense University, Bogor, Indonesia
[2] Department of Mathematics, Faculty of Military Mathematics and Natural Sciences, Republic of Indonesia Defense University, Bogor, Indonesia
* Corresponding Author: rofah.rachmawati@idu.ac.id

Abstract

GPA is a characteristic measurement which indicates the academic achievements of students in completing the courses in every semester. The conducted research is to determine whether the impact of various factor contribute to the student GPA with different backgrounds. Sample students and data are collected from one study program: GPA, gender, physical fitness assessment, and their organization role. The data is then represented in longitudinal model with data analysis using RStudio. The results are: first, non-random effect model showed that time variable is not significant to the GPA although there is indication of reduction. Second, random effect model of ID (student) showed time is significant to GPA with negative effect. Third, with random effect parameter of gender, it was shown that time variable is not significant to the GPA with negative effect. Fourth, with random effect ID and gender, time is greatly impacted to the GPA with negative effect. Fifth, the model with covariate variable of physical fitness shown no significant effect to the GPA, nevertheless, there is a negative effect. Sixth, covariate variable model of student engaged in corps regiment showed that there is no significant effect to the GPA but, there is a decrease of it.

Keywords: *GPA; Physical Health; Organization; Gender.*

Introduction

Assessment in college is a process to standardize the measurement from various achievements and/or awards from each discipline. The standard measurement system that is being used is Grade Point Average (GPA) (Makkiyah et al., 2019). GPA is measurement of capabilities of student at certain period with the number of credits are taken (Sihite & Pratiwi, 2018). According to Sihite and Pratiwi (2018), the factors that impact the GPA are their salary, campus activity, age, learning hours, and organization participates. Organization is a place of groups with engaging teamwork rationally and systematically in managing its resources to achieve their goal (al Hairi, 2021). According to the research by Kurnia (2014) at Cokroaminoto University Yogyakarta, the activeness in the organization encouraged student to socialize, help each other, sharing opinions,

and engage in the social communities to apply their knowledge for leadership development and adulthood process (Kurnia, 2014).

Health is also considered the contribute factor of student's GPA. Their health condition is related to physical fitness each students. Hence, physical fitness is the ability to conduct daily activity (working) with not sustaining too much fatigue, hence the body still have sources to undertake additional workload (Darmawan, 2017). According to Alfarisi's research (2014) physical fitness affect the GPA of male and female medical students in Malahayati University of Bandar Lampung. In addition, the gender difference is not contributing significantly to the GPA, according to the research conducted by Makkiyah et.al., (2019). Students with good physical fitness obtained higher GPA because their healthy physical condition positively impacted their studying ability.

Cadets (College student) of Republic of Indonesia Defense University (RIDU) are advanced human source from 34 provinces in Indonesia that is being prepared to maintain the Nation's sovereignty and to defend against any threat dimension. Aside from attending courses, cadets also engaged with physical exercises and military basic training. They are being prepared to become the next generation of high intellectual nation's defender. They are highly intelligent and strong mentally and physically. Their daily activities are very tight scheduled. Some of them are course attendance, martial arts, physical training, drum band, orchestra, choir, and general sports. In addition, The Rector of RIDU has inaugurated the official organization called Corps Regiment of Student Cadet as a platform for implementing the leadership and organizing capabilities, encouraging the togetherness, and improving their diligence in studying. It also the place to guide and nurture the student cadet to shape attitudes, character, and behavior. It also helps students to obtain higher achievement. All of those activities both military and academic are carried out in order to balance reliable the academics and army officers.

The different living conditions and environment between student cadet and other student from different universities is due to the daily activities that is different especially when related to the organization and physical activity. Therefore, the conducted research is to find out how the influence of organizational activity, physical condition, and gender on the GPA of students.

Methods

The data used in this study is secondary data with longitudinal data types. Longitudinal data is data that is observed and measured repeatedly at a certain time interval (Islamiyati, 2014). The research data comes from a specific major/study program data set of student cadet GPA from semester 2 and semester 4, grades, gender, and information that the student cadet is a corps regiment administrator. The GPA for semesters 2 and 4 is used because the management of the corps regiment organization is known for 1 year in semesters 2 and 4. Variations in GPA for student cadets are shown in Figure 1. Differences in color on the lines indicate variations in each students (ID) or the number of student cadet samples. The graph shows a downward trend.

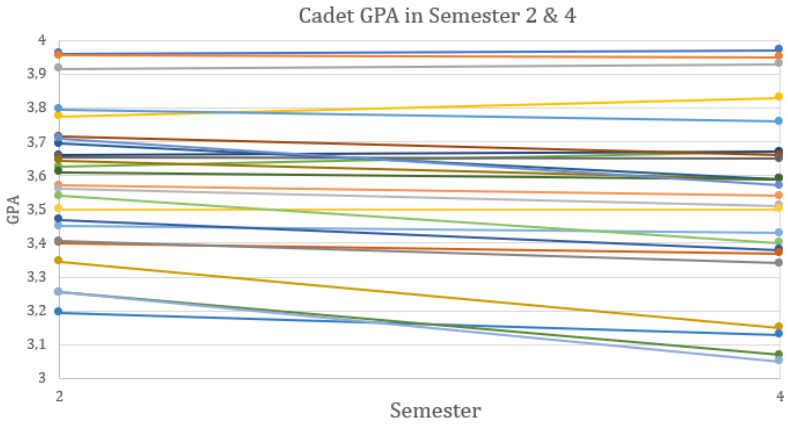

Figure 1. Variation of GPA samples.

The analytical method used in this study is longitudinal analysis with a linear mixed model approach. The first step is the formation of model specifications and then followed by parameter estimation. After obtaining the parameter values, then the results of the parameter values are included in the model specification so that a model is formed and the model results are interpreted. Furthermore, a hypothesis test was carried out to determine the effect of the independent variables on the response variable. In this study the data will be modeled into two general models, i.e. first model being a random intercept model and the second model being a random intercept and slop model. To find out the best model, the criteria for selecting the best model using AIC are used. Correlation between observations within the same individual in longitudinal data makes the usual statistical procedure, namely the General Linear Model (GLM) inapplicable. The proper method used for longitudinal data analysis is the Linear Mixed Model (LMM) (Nirmala et al., 2013). The linear mixed model is the result of the development of a linear model where the response variable (Y) is affected by fixed effects and random effects.

The linear mixed model is an extension of the linear model by adding random effects. This method is widely used in experimental designs for correlated data such as repeated measures experiments (Searle et al., 1992). The mixed model is a statistical model that contains both fixed and random effects. The mixed effect model can be represented as:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p + \alpha_i + e$$

where:

y : response variable

$x_1, x_2, \ldots, x_p$: predictor variable

β_0 : intercept, representing all expected value from y when all predictor variable are equal to zero.

$\beta_1, \beta_2, \ldots, \beta_p$: coefficient, representing all changes of y with respect of change in one unit of appropriate predictor variable.

e : error term, representing actual response error from predicted response due to random noise or other factors that are not obtained by the model.

alpha represents random effect, such as gender, student ID and others that can represents uniqueness of each and multilevel sample.

The coefficient β represents the effect of each predictor variable on the response variable. For example, if the beta is positive, then an increase in the first predictor variable is associated with an increase in the response variable. If the beta is negative, then an increase in the first predictor variable is associated with a decrease in the response. Linear mixed models can be used to estimate the effect of each predictor variable on responses, as well as to predict responses to new data. Coefficient estimation can be obtained using various techniques, such as least squares or maximum probability. From the description above, the stages of the data analysis method used are:

Data modelling

At this stage, data will be modelled in accordance to linear mixed model. Data is modelled as two specific model (Rizopoulus, 2012). The model equations are:

Model 1 Longitudinal Data
Random intercept model:

$$\begin{cases} & \beta_0 + \beta_1 t_k + a_i + \varepsilon_{ijk} \\ Y_{ijk} = & a_i \sim N(0, \sigma_{a_i}^2) \\ & \varepsilon_{ijk} \sim N(0, \sigma^2) \end{cases}$$

Model 2 Longitudinal Data
Intercept and random slot model:

$$\begin{cases} & \beta_0 + \beta_1 t_{ij} + a_i + a_{i1} t_{ij} + \varepsilon_{ijk} \\ Y_{ijk} = & a_i \sim N(0, G) \\ & \varepsilon_{ijk} \sim N(0, \sigma^2) \end{cases}$$

with,

$$G = \begin{pmatrix} \sigma_{00} & \sigma_{01} \\ \sigma_{10} & \sigma_{11} \end{pmatrix}$$

The next step is to define parameter value by parameter estimation.

Parameter Test

In the parameter test, simultaneous and partial tests were carried out to determine the influence of independent variables on the response variables. The test was simultaneously carried out using a G-annotated likelihood ratio test and test statistics following the chi-square attribution. With the hypothesis are:

$H_0 = \beta_1 = \beta_2 = \cdots = \beta_p = 0$ (Independent variable simultaneously not significant towards response variable)

$H_1 = at\ least\ only\ one\ \beta_1 \neq 0$ with $i = 1,2, \ldots, p$ (I is the number of independent variable, at least only one independent variable that is simultaneously independent towards response variable)

The decision is denying the H_0 if G ≥ $X^2_{a;p}$ or p-values < a
The partial test is conducted by t test with the hypothesis:
H_0: $\beta_1 = 0$ (the parameter coefficient of I is not applicable)
H_1 : at least one $\beta_1 \neq 0$ with I = 1,2,...,p (I is the number of independent variable)
H_0 is denied if $t_{calculated} \geq t_{\frac{a}{2},(n-p-1)}$ or $t_{calculated} \leq -t_{\frac{a}{2},(n-p-1)}$ or if compared with p-value H_0 denied if p-value ≤ a.

Model Result Interpretation

The model that has been formed is interpretated its relationship between significant variables.

Selection of the best Model

The best model is determined from the smallest AIC (Akaike Information Criterion) value.

Result and Discussion

Data collection is carried out two times, that is on the semester 2 and semester 4. Using R Studio software, the results of this study is analyzed for each model. In model 1, we used a *Linear Mixed Model* (LMM) in the absence of random effects. The mathematical model is as follows:

$$Y_{ijk} = \beta_0 + \beta_1 t_k + \varepsilon_{ijk}$$
$$\varepsilon_{ijk} \sim N(0, \sigma^2_{residual})$$

The respondent variables are expressed by Y_{ijk} that is *Grade Point Average* (GPA) on the j-th person, the i-th identity, and the measurement at the k-th time, which is k=0 for the second semester GPA value and k=1 for the fourth semester GPA, and ε_{ijk} is residual. By using the R Studio application, the results in Table 1 are obtained.

Table 1. Model 1 Parameter Test Results Without Random Effects

	Estimation	Standard Error	p-value
Intercept (β_0)	3,5836	0,0472	0,000***
Time (β_1)	-0,0516	0,0668	0,4436

*** significant at $\alpha = 1\%$
** significant at $\alpha = 5\%$
* significant at $\alpha = 10\%$

Based on Table 1, the p-value obtained at the time variable is insignificant, which is more than 0.05, meaning that the distance between the GPA in semesters 2 and 4 does not affect the student's Grade Point Average (GPA). Then it will be further researched with model 2, we add a random effect. In model 2, a random effect will be added in the form of a student ID (identity). The addition of this random ID effect is carried out

considering the difference in identity of each individual or random student. The mathematical model is as follows:

$$Y_{ijk} = \beta_0 + \beta_1 t_k + a_i + \varepsilon_{ijk}$$
$$a_i \sim N(0, \sigma_{ID}^2)\ , \varepsilon_{ijk} \sim N(0, \sigma_{residual}^2)$$

The second model is being added a_i which is random effect of 25 identitiy a_i and ε_{ijk} which are not corelated. By using R studio, the results on table 2 are obtained.

Table 2. Model 2 Parameter Test Results with ID Random Effect

	Estimation	Standard Error	p-value
Intercept (β_0)	3,5836	0,0472	0,000***
Time (β_1)	-0,0516	0,0143	0,001***

*** significant at $\alpha = 1\%$
** significant at $\alpha = 5\%$
* significant at $\alpha = 10\%$

Based on Table 2, the p-value obtained at a significant time variable is smaller than 0.01, meaning that the distance between the GPA time in semesters 2 and 4 *affects the* student's Grade Point Average (GPA) with the random effect of each student's identity. Then it will be further researched with model 3, we change the random effect. In model 3, the random effect will be changed, which is in the form of a random effect of gender. This random gender selection was chosen because of the difference in gender in a class of 25 people. The mathematical model is as follows:

$$Y_{ijk} = \beta_0 + \beta_1 t_k + b_{j(i)} + \varepsilon_{ijk}$$
$$b_{j(i)} \sim N\left(0, \sigma_{jK}^2\right)\ , \varepsilon_{ijk} \sim N(0, \sigma_{residue}^2)$$

In this third model the random effect of ID or a_i is replaced with a random effect of gender or $b_{j(i)}$with $b_{j(i)}$ and ε_{ijk} uncorrelated. By using the R Studio, the results in Table 3 are obtained.

Table 3. Model 3 Parameter Test Results with Randomized Gender Effect

	Estimation	Standard Error	p-value
Intercept (β_0)	3,5836	0,0472	0,000***
Time (β_1)	-0,0516	0,0668	0,4374

*** significant at $\alpha = 1\%$
** significant at $\alpha = 5\%$
* significant at $\alpha = 10\%$

Based on Table 3, the p-value obtained on the insignificant time variable is more than 0.05, meaning that the distance between the GPA time in semesters 2 and 4 does not affect the student's Grade Point Average (GPA) with a random effect on each gender of the student. There are significant differences between model 2 and model 3, so

research will be conducted by combining the two random effects. In model 4 ID and Gender will be combined as a random effect. The mathematical model is as follows:

$$Y_{ijk} = \beta_0 + \beta_1 t_k + a_i + b_{j(i)} + \varepsilon_{ijk}$$

$$a_i \sim N(0, \sigma_{ID}^2)\ , b_{j(i)} \sim N\left(0, \sigma_{JK}^2\right), \varepsilon_{ijk} \sim N(0, \sigma_{residual}^2)$$

In this fourth model the random effect of ID or a_i is combined with the random effect of Gender or $b_{j(i)}$with a_i, $b_{j(i)}$, and ε_{ijk} that are uncorrelated. By using the R Studio, the results in Table 4 are obtained.

Table 4. Model 4 Parameter Test Results with Random ID and Gender Effects

	Estimation	**Standard Error**	**p-value**
Intercept (β_0)	3,5826	0,0472	0,000***
Time (β_1)	-0,0516	0,0143	0,001***

*** significant at $\alpha = 1\%$
** significant at $\alpha = 5\%$
* significant at $\alpha = 10\%$

Based on Table 4, the p-value obtained at a significant time variable of less than 0.01, meaning that the distance between the GPA in semesters 2 and 4 affects the student's Grade Point Average (GPA) with the random effect of each student's identity and the random effect of their gender. Then the coefficients of time and *intercept* from Table 1 to Table 4 are collected in Table 5.

Table 5. Result – Conclusion

Parameter	**Model 1 (s.e)**	**Model 2 (s.e)**	**Model 3 (s.e)**	**Model 4 (s.e)**
Fixed Effect				
Intercept (β_0)	3,5836 (0,0472)	3,5836 (0,0472)	3,5865 (0,0474)	3,5836 (0,0472)
Time (β_1)	-0,0516 (0,0668)	-0,0156 (0,0143)	-0,0547 (0,0665)	-0,0515 (0,0143)
Variance				
σ_{ID}^2	-	0,0532	-	0,05320
σ_{JK}^2	-	-	0,0479	0,00162
$\sigma_{residue}^2$	0,0557	0,00258	0,0079	0,00096
Total	0,0557	0,0557	0,00558	0,05578

Based on the four models that have been carried out, it can be seen that there is a slight influence of random effects on fixed effects and also does not have a large influence on *Standard Error*. It is noticed that there is no significant change in value to time in all four models, this is seen by the number of β_0 which always positive. But the

value of time remains negative and does not change. Then in this study, the correlation structure will be like in model 4.

$$\widehat{Corr}\left(Y_{ijk}, Y_{i*j*k*}\right) = 0$$

Research on model 4 with different identities, different genders, different times, i and i^*, $i \neq i^*$, the correlation is 0, As in the initial assumption of being mutually independent. Subsequent studies with the same identity, different gender, and different times had a correlation of 0.953.

$$\widehat{Corr}\left(Y_{ijk}, Y_{ij*k*}\right) = \frac{\hat{\sigma}^2_{ID}}{\hat{\sigma}^2_{ID} + \hat{\sigma}^2_{JK} + \hat{\sigma}^2_{residu}} = \frac{0{,}05320}{0{,}05578} = 0{,}953$$

Subsequent research with the same identity, same gender, different times, j and j^*, $j \neq j^*$ has a correlation of 0.982 which means it is very correlated.

$$\widehat{Corr}\left(Y_{ijk}, Y_{ijk*}\right) = \frac{\hat{\sigma}^2_{ID} + \hat{\sigma}^2_{JK}}{\hat{\sigma}^2_{ID} + \hat{\sigma}^2_{JK} + \hat{\sigma}^2_{residu}} = \frac{0{,}05320 + 0{,}00162}{0{,}05578} = 0{,}982$$

Additional covariates can also be added in research to explain diversity in different models, or to study the characteristics of each student's identity and gender that are related to time trends. In model 4 of this study, covariate variables will be added. Covariate variables are useful variables to reduce noise in data analysis caused by other variables other than the variables under study, this causes the effects of the variables studied could be seen more clearly. The added covariate variables are the results of the student's physical fitness assessment (NG), and status as an organizational manager or staff (ket). These covariate variables will be investigated for their respective influences or effects separately (Linden et al., 2006)

In model 5, the covariate variables resulting from the physical fitness scores of the students carried out at the end of the semester will be investigated for their effect on the previous model. Using the R Studio application obtained the results in Table 6.

Table 6. Model 4 Parameter Test Results with Physical Fitness Covariate Variables

	Estimation	**Standard Error**	**p-value**
Intercept (β_0)	3,685	0,128	0,000***
NG	-0,001	0,001	0,403
Time (β_1)	-0,127	0,132	0,347
NG: TIME (β_1)	0,001	0,001	0,565

*** significant at $\alpha = 1\%$
** significant at $\alpha = 5\%$
significant at $\alpha = 10\%$

Based on Table 6, a p-value was obtained on the covariate variable The result of the Physical fitness value is (NG) not significant, which is less than 0.05, meaning that the student physical fitness covariate variable does not have a significant effect on the student's Grade Point Average (GPA). Meanwhile, over time the results of physical fitness assessments have been negative. This means that Physical Fitness decreases over time but does not have a significant effect on changes in student GPA.

Table 7. Model 4 Parameter Test Results with Covariate Variables of Organizational Engagement Cadets

	Estimation	Standard Error	p-value
Intercept (β_0)	3,577	0,080	0,000***
ket	0,009	0,100	0,928
Time (β_1)	-0,045	0,024	0,075
ket: TIME (β_1)	-0,009	0,030	0,759

*** significant at $\alpha = 1\%$
**significant at $\alpha = 5\%$
* significant at $\alpha = 10\%$

Based on Table 7, a p-value is obtained on the covariate variable of the student's description as the administrator of the organization, which has no significant effect of more than 0.05, meaning that the covariate variable of the information as the administrator of the organization does not have a significant effect on the GPA of the student for 2 years or one school year. Meanwhile, the time score is still negative, meaning that there is a decrease in the GPA of students but it has no effect and significant effect on the status as a member of the organization.

Then, the coefficients of time and *intercept* from Table 5 to Table 7 are presented in Table 8.

Table 8. Conclusion of Test Results with Covariate Variables

Parameter	Model 5 Coefficient (s.e)	Model 6 Coefficient (s.e)	Model 7 Coefficient (s.e)
Fixed Effect			
Intercept (β_0)	3,685 (0,128)	3,577 (0,080)	3,569 (0,196)
Time (β_1)	-0,127 (0,132)	-0,045 (0,024)	-0,041 (0,178)
NG	-0,001 (0,001)	-	-0,000 (0,002)
NG: TIME	0,001 (0,002)	-	-0,003 (0,002)
ket	-	-0,045 (0,024)	-0,001 (0,108)
ket: TIME	-	-0,009 (0,030)	0,000 (0,035)
JKP	-	-	0,034 (0,111)
JKP : TIME	-	-	-0,041 (0,043)

Variance			
σ^2_{ID}	0,052	0,055	0,058
$\sigma^2_{residue}$	0,002	0,002	0,002
Total	0,054	0,057	0,06

Conclusion

The Grade Point Average (GPA) of RIDU students for one year (from semester 2 to semester 4) are decreased. However, this is not too significant regarding the academic achievements of them. Based on the research we conducted, there is a relationship between the decrease in GPA in students for one year with the random effect of student identity. In this case, it shows that the factors that affect the decline in the GPA of students can occur depending on each person. It can be from the different ways of learning of each individual or from the diligent nature and habits and lifestyles of each individual.

Then based on the research we have done using the *linear mixed model*, it was found that the covariate variables of physical fitness value and organizational engagements did not have a significant effect on changing the GPA for one year, with a positive coefficient. It can be concluded that the daily routine with tight schedule and student activity in organization does not affect the decrease in the GPA.

Further research can be carried out by developing the following aspects, 1) Random effects, the study is limited only to using random effects of identity and gender, the addition of other random effects could be interesting for subsequent research; 2) Covariate Variables, in this study only two variables were used in the model. Therefore, there are many other variables that can be used in the model to analyze GPA changes; 3) Area. This research is limited to only one department/study program, further research with other majors or within the scope of one larger faculty will likely obtain more details regarding the factor of GPA of students; 4) Intervariable interaction; The last model of this study only directly combines the two variables in one model, the interaction between two or three variables will be interesting for subsequent studies.

References

[1] Al Hairi, M. R. (2021). Budaya Organisasi dan Dampaknya Terhadap Lembaga Pendidikan. Adiba: Journal of Education, 1, 79–87.

[2] Alfarisi. R. (2014). Pengaruh Kebugaran Jasmani, Aktifitas Fisik, dan Indeks Massa Tubuh Terhadap Indeks Prestasi Kumulatif (IPK) Mahasiswa Kedokteran Universitas Malahayati Bandar Lampung. Jurnal Kedokteran dan Kesehatan, 1(1), 52. https://doi.org/10.33024/.v1i1.298

[3] Darmawan. I. (2017). Upaya Meningkatkan Kebugaran Jasmani Siswa Melalui Penjas. Jurnal Inspirasi Pendidikan, 7(2), 143–154. https://doi.org/10.21067/jip.v7i2.1700

[4] Kurnia. H. (2014). Pengaruh Keaktifan Berorganisasi Terhadap Indeks Prestasi Kumulatif Mahasiswa Universitas Cokroaminoto Yogyakarta. Academy of Educarion Journal, 5(2), 91. https://doi.org/10.47200/aoej.v5i2.120

[5] Linden, A., Trochim. W. M. K., & Adams, J. L. (2006). Evaluating Program Effectiveness Using the Regression Point Displacement Design. Evaluation and the Health Professions, 29(4). https://doi.org/10.1177/0163278706293402

[6] Makkiyah. F., Harfiani. E., & Anisah. (2019). Pengaruh Jenis Kelamin dalam Variasi Indeks Prestasi Kumulatif Mahasiswa Kedokteran di Universitas Pembangunan Nasional Veteran Jakarta. Jurnal Profesi Medika, 13(1), 35–39. https://doi.org/10.33533/jpm.v13i1.796

[7] Nirmala. F. (2013). Aplikasi General Linear Mixed Model (GLMM) pada Data Longitudinal Kadar Trombosit Demam Berdarah Dengue. Jurnal Biometrika dan Kependudukan, 2(2), 131-139.

[8] Rifai. D., Sihite. D. B., and Pratiwi. N. (2018). Analisis Jalur Terhadap Faktor Faktor yang Mempengaruhi Indeks Prestasi Kumulatif (IPK) Mahasiswa. Jurnal Statistika Industri dan Komputasi, 3(1), 31–39. https://doi.org/10.34151/statistika.v3i01.1083

[9] Searle. S. R., Casella. G., and McCulloch. C. E. (1992). Variance Components John Wiley and Sons. Inc. New York.

[10] Sihite. D. R. D. B., and Pratiwi. N. (2018). Analisis Jalur Terhadap Faktor Faktor yang Mempengaruhi Indeks Prestasi Kumulatif (IPK) Mahasiswa. Jurnal Statistika Industri Dan Komputasi, 3(1), 31–39. https://doi.org/10.34151/statistika.v3i01.1083

[11] Wirastuti. M. (2019). Analisis Longitudinal pada Data Pasien Diabetes Melitus. Jurnal Statistika: Jurnal Ilmiah Teori dan Aplikasi Statistika, 12(1), 13-19.