

## PANEL DATA REGRESSION MODEL FOR PREDICTING ECONOMIC GROWTH BEFORE AND DURING THE COVID-19 PANDEMIC IN EAST JAVA PROVINCE

**Ronny Susetyoko<sup>1\*</sup>, Edi Satriyanto<sup>2</sup>, Alfi Fadliana<sup>3</sup>, Fitriah Maharani Humaira<sup>4</sup>**

<sup>1,2,3,4</sup>Department of Informatics and Computer Engineering, Surabaya State Electronics Polytechnic  
Raya Street ITS, Kampus PENS, Keputih Sukolilo Surabaya, 60111, Indonesia

Corresponding author's e-mail: \* [rony@pens.ac.id](mailto:rony@pens.ac.id)

### ABSTRACT

#### Article History:

Received: 26<sup>th</sup> June 2023

Revised: 28<sup>th</sup> September 2023

Accepted: 29<sup>th</sup> October 2023

#### Keywords:

Cross Section;

Economic Growth;

Fixed Effect Model;

GRDP;

Panel Data Regression.

Gross Regional Domestic Product (GRDP) is one of the key indicators to determine the economic condition of a region in a certain period. GRDP at constant prices has a positive and significant effect on economic growth. This study aims to predict economic growth in East Java before and during the Covid-19 pandemic based on the structural components of regional revenue and expenditure budget realization using panel data regression with data sources from the Directorate General of Fiscal Balance. The results of this study, the best model is the Fixed Effect Model (FEM) with R-Squared 0.99991 and Adj. R-Squared 0.99987. MSE, MAD and MAPE values on the training data are 338724.9919, 259.7182 and 0.6296 respectively. While the MSE, MAD and MAPE values in the testing data are 1716324.2736, 445.7959 and 1.0692 respectively. At the 95% confidence level, Locally Generated Revenue (LGR), Transfers to Regional and Village Funds (TRVF), and Other Revenue (OR) are not significant in the model or have little effect. But at the 99.9% confidence level, all factors (cross section) have a very significant effect. This can be interpreted that local wisdom, or the characteristics of each region/city has a major contribution to economic growth.



This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-ShareAlike 4.0 International License.

#### How to cite this article:

R. Susetyoko, E. Satriyanto, A. Fadliana, F. M. Humaira., "PANEL DATA REGRESSION MODEL FOR PREDICTING ECONOMIC GROWTH BEFORE AND DURING THE COVID-19 PANDEMIC IN EAST JAVA PROVINCE," *BAREKENG: J. Math. & App.*, vol. 17, iss. 4, pp. 2121-2134, December, 2023.

Copyright © 2023 Author(s)

Journal homepage: <https://ojs3.unpatti.ac.id/index.php/barekeng/>

Journal e-mail: [barekeng.math@yahoo.com](mailto:barekeng.math@yahoo.com); [barekeng.journal@mail.unpatti.ac.id](mailto:barekeng.journal@mail.unpatti.ac.id)

Research Article · Open Access

## 1. INTRODUCTION

Gross Regional Domestic Product (GRDP) is one of the main indicators to determine the economic condition of a region in a certain period [1][2], based on current prices or constant prices. According to data from the Central Statistics Agency (BPS), in 2021 the GDP contribution of the island of Java will be around 59% [1]. The basic sectors of regional economic activity can be analyzed using Location Quotient (LQ) with the Gross Regional Domestic Product (GRDP) variable as an indicator of regional growth [3]. The Covid-19 pandemic has had a significant impact on many sectors in most countries in Asia, America, Africa, Europe, and Australia [4]. Limited community activity resulted in a decrease in production and capital outflow which had an impact on the weakening of the rupiah exchange rate [5].

Panel data models, both fixed and random effects models, are highly reliable for multivariable longitudinal data sets. Panel data regression models are used for business forecasting and to determine the relationship between attributes in a data set [6]. Panel data regression is a combination of time series data and cross section data [7][8]. The panel data regression model is applied to determine the variables that influence economic growth in Asian countries. The results of this research show that taxes, government debt, household consumption, and credit interest rates have a significant effect on economic growth [7]. Panel data regression was also used to model the percentage of poor people by district/city in East Kalimantan. Attributes that are thought to influence include population growth rate, human development index, and adjusted per capita expenditure [8]. Panel data regression was also used to determine the attributes that influence the level of poverty in Madura [9].

This research aims to predict East Java's economic growth before the pandemic (2019) and during the Covid-19 pandemic (2020) based on Gross Regional Domestic Product (GRDP) as the basic sector of economic activity using panel data regression. Apart from making predictions, this model is also used to determine the influence of differences in entities (districts/cities) and the influence of differences in observation periods (years). The variables or attributes used in modeling are the structure of the realization of the Regional Revenue and Expenditure Budget (RREB) which includes Regional Original Income (LGR), Transfers to Regional and Village Funds (TRVF), Other Income (OR), Labor Expenditures (EE), Goods and Services Expenditures (GSE), Capital Expenditures (CE), Other Expenditures (OE), and Regional Fund Revenue (RFR). Panel data regression is a predictive analysis of time series data mining.

Panel data regression models are also applied in several cases, including: to predict the relationship between tourism and economic development in Southeast Asian countries [10], analysis of the influence of population aging on public debt [11], analysis of the influence of several attributes of bad credit, loans banks, inflation, and CO2 emissions on the economic growth of several countries [12], modeling consumption by considering the cross-section of each type of product category and household [13], and analyzing the increase in infant mortality rates sequentially to obtain several recommendations that support reducing infant mortality rates in East Java [14]. This method is also applied to modeling the depth of poverty index in districts/cities in Papua Province. The Fixed Effect Model (FEM) was chosen as the best model with an R-square value of 82.5% [15]. Panel data regression models are also used to analyze the impact of several e-commerce attributes on economic growth in Indonesia. The results of this research show that the e-commerce attribute that has a very significant influence is computer users [16]. Panel data is also used to model income inequality based on economic growth attributes. The results of this study show that the attribute of economic growth has a significant influence on income inequality [17]. Similar research was also carried out to model the percentage of poor people based on several attributes, namely: economic growth rate, average years of schooling, dependency ratio, and life expectancy. The results of this research show that the attributes that have a significant influence are the average number of years of schooling, dependency ratio, and life expectancy [18].

The advantages of panel data regression are the greater number of observations, reducing the occurrence of multicollinearity, prediction results tend to be more effective, and obtaining more information. [19]. The panel data model is also very good at predicting poverty in Bengkulu Province with the Fixed Effect Model (FEM) as the best model. The significant attribute in this model is GRDP, with a Mean Absolute Percentage Error (MAPE) value of 6.59% and a Mean Square Error (MSE) of 5.48 [20].

Gross Regional Domestic Product (GRDP) is one of the main indicators to determine the economic condition of a region in a certain period [1][2], based on the current price or constant price. According to data from Badan Pusat Statistik (BPS), in 2021 the GRDP contribution of the island of Java will be around 59% [1]. GRDP has a significant effect on economic growth [3]. The Covid-19 pandemic has had a significant impact on many sectors in most Asian, American, African, European and Australian countries [4]. Limited community activities resulted in decreased production and capital outflows, which resulted in a weakening of the rupiah exchange rate [5].

This research aims to predict East Java's economic growth before the pandemic (2019) and during the Covid-19 pandemic (2020) based on Gross Regional Domestic Product (GRDP) using panel data regression. Apart from predicting, it is also used to analyze the influence of differences in entities (districts/cities) and/or the influence of differences in observation periods (years). The variables or attributes used in modeling are the structure of the realization of the Regional Revenue and Expenditure Budget (RREB) which includes Original Regional Income (LGR), Transfers to Regional and Village Funds (TRVF), Other Income (OR), Labor Expenditures. (EE), Goods and Services Expenditures (GSE), Capital Expenditures (CE), Other Expenditures (OE), and Regional Fund Revenue (RFR).

Panel data regression is predictive analytics from time series data mining. Panel data models, both fixed and random effects models are very reliable for multivariable longitudinal datasets. This model is used for business forecasting as well as to find out the relationship between attributes in a data set [6]. Panel data regression is a combination of time series data and cross section data [7][8].

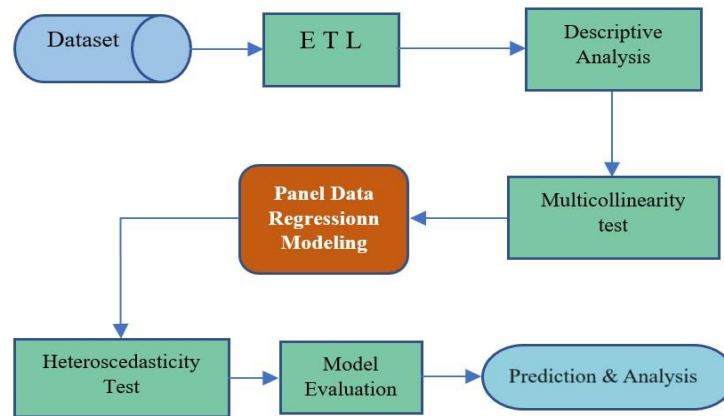
The panel data regression model is applied to determine the attributes that influence economic growth in Asian countries [7]. This regression model is also used to model the percentage of poor people by district/city in East Kalimantan. Attributes that are thought to influence include population growth rate, human development index, and adjusted per capita expenditure [8]. Panel data regression was also used to determine the attributes that influence the level of poverty in Madura [9].

Some other applications of panel data regression models include: to predict the relationship between tourism and economic development in Southeast Asian countries [10], analysis of the influence of population aging on public debt [11], analyzing the influence of several attributes of bad credit, bank loans, inflation, and CO2 emissions on the economic growth of several countries [12], consumption modeling by considering the cross section of each type of product category and household [13], and to analyze the increase in infant mortality rates in order to obtain several recommendations that support reducing infant mortality rates in Java East [14]. This method is also applied to poverty depth index modeling in districts/cities of Papua Province. The Fixed Effect Model (FEM) model was chosen as the best model with an R-square value of 82.5% [15]. The panel data regression model is also used to analyze the impact of several e-commerce attributes on economic growth in Indonesia. The results of this study indicate that the attributes of e-commerce that have a very significant effect are computer users [16]. Panel data is also used for modeling income inequality based on economic growth attributes. The results of this study show that the attributes of economic growth have a significant impact on income inequality [17]. A similar study was also conducted to model the percentage of poor people based on several attributes, namely: the rate of economic growth, the average length of schooling, the ratio of dependents, and life expectancy. The results of this study, the attributes that have a significant effect are the average length of schooling, the ratio of dependents, and life expectancy [18].

The advantages of panel data regression are the greater number of observations, reducing the occurrence of multicollinearity, predictive results tend to be more effective, and obtain more information. [19]. The panel data model is also very good for predicting poverty in Bengkulu Province with the Fixed Effect Model (FEM) as the best model. The significant attribute in this model is GRDP, with a Mean Absolute Percentage Error (MAPE) value of 6.59% and a Mean Square Error (MSE) of 5.48 [20].

## 2. RESEARCH METHODS

This research methodology can be seen in **Figure 1**. The initial step is data collection and cleaning or what is often referred to as extraction, transformation, and load (ETL). Next, an initial analysis of the characteristics of the attributes used as response and predictor variables is carried out. The Variance Inflation Factor (VIF) test is used to ensure that multicollinearity does not occur in the attributes as predictors before modeling.



**Figure 1. Methodology**

The next step as the main process is panel data regression modeling. There are three regression models to be tested to select the appropriate model, namely the Common Effect Model (CEM), Fixed Effect Model (FEM), and Random Effect Model (REM). In the selected appropriate model, a heteroscedasticity test was carried out to ensure that the variance between groups of observations was homogeneous. Next, an evaluation of the model is carried out based on the values of the Mean Square Error (MSE), Mean Absolute Deviation (MAD), and Mean Absolute Percentage Error (MAPE). The final stage is to predict and analyze GRDP as a statistic to measure the economic level or aggregation of gross value added based on the structure of the regional revenue and expenditure budget realization in each regency/city in East Java Province.

## 2.1 Dataset

The data sources for this research are from the Central Bureau of Statistics for East Java Province and the Directorate General of Fiscal Balance of the Ministry of Finance, namely: which were:

- a) Gross Regional Domestic Product (GRDP) data based on current prices/constant prices according to business field of East Java Province in 2014 – 2021.
- b) Data on the structure of the Regional Revenue and Expenditure Budget for 2017 – 2021 consists of: Locally Generated Revenue (LGR), Transfer to Regional and Village Fund (TRVF), Other Revenue (OR), Employ Expenditures (EE), Goods and Services Expenditures (GSE), Capital Expenditures (CE), Other Expenditures (OE), and Regional Financing Revenues (RFR).

## 2.2 Panel Data Regression

There are three panel data regression model approaches, namely the Common Factor Model (CEM), Fixed Effect Model (FEM), and Random Effect Model (REM). The common effect model is the same as Ordinary Least Square (OLS), all data are combined, both cross section data and time series data. The CEM equation is expressed in **Equation (1)**.

$$Y_{it} = \beta_0 + \sum_{k=1}^p \beta_k X_{kit} + e_{it} \quad (1)$$

where:

- $Y_{it}$  : response variable for the  $i$ -th object in the  $t$ -period
- $\beta_0$  : intercept of regression model
- $\beta_k$  : the slope/parameter of the  $k$ -th variable
- $X_{kit}$  : the  $k$ -th predictor variable for the  $i$ -th object in the  $t$ -period
- $e_{it}$  : residual/error for the  $i$ -th object in the  $t$ -period
- $i$  : cross section units (1, 2, ..., N)
- $t$  : time series units (1, 2, ..., T)
- $p$  : the number of attributes as a predictor variable

In the Fixed Effect Model, the underlying assumption is that there are different effects between objects and periods. In this model, differences in the characteristics of objects and periods are accommodated in intercepts/constants so that their values change with respect to objects and periods. Fixed Effect Model used in this study is expressed in **Equation (2)**.

$$Y_{it} = \alpha_i + \gamma_t + \sum_{k=1}^p \beta_k X_{kit} + e_{it} \quad (2)$$

where:

- $Y_{it}$  : response variable for the  $i$ -th object in the  $t$ -period
- $\alpha_i$  : constant in unit cross section  $i$
- $\gamma_t$  : constant in the  $t$ -period
- $\beta_k$  : the slope/parameter of the  $k$ -th variable
- $X_{kit}$  : the  $k$ -th predictor variable for the  $i$ -th object in the  $t$ -period
- $e_{it}$  : error for the  $i$ -th object in the  $t$ -period
- $i$  : cross section units (1, 2, ..., N)
- $t$  : time series units (1, 2, ..., T)
- $p$  : the number of attributes as a predictor variable

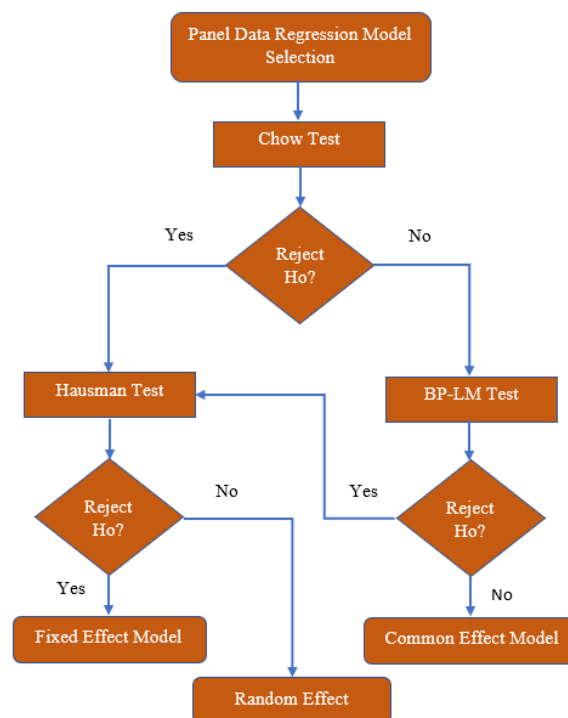
Meanwhile, the Random Effect (REM) model is stated in **Equation (3)**.

$$Y_{it} = \beta_0 + \sum_{k=1}^p \beta_k X_{kit} + e_i + u_{it} \quad (3)$$

where:

- $Y_{it}$  : response variable for the  $i$ -th object in the  $t$ -period
- $\beta_0$  : common mean value for the intercept
- $\beta_k$  : the slope/parameter of the  $k$ -th variable
- $X_{kit}$  : the  $k$ -th predictor variable for the  $i$ -th object in the  $t$ -period
- $e_i$  : error component for object  $i$
- $u_{it}$  : combined error cross section and time period
- $i$  : cross section units (1, 2, ..., N)
- $t$  : time series units (1, 2, ..., T)
- $p$  : the number of attributes as a predictor variable

Panel data regression modeling uses three models: CEM, FEM, and REM. The model selection algorithm can be seen in **Figure 2**. The Chow test is a test to choose between CEM and FEM. If the Chow  $H_0$  test is accepted, then the Breusch-Pagan - Lagrange Multiplier test (BP-LM test) is carried out.



**Figure 2. Panel Data Regression Model Selection Algorithm**

If the BP – LM test turns out that  $H_0$  is accepted, then the best model is CEM. If  $H_0$  is rejected, it is necessary to do the Hausman test, which is to choose between FEM and REM. If  $H_0$  is rejected, then the best model is FEM. Conversely, if  $H_0$  is accepted then the best model is REM.

### 3. RESULTS AND DISCUSSION

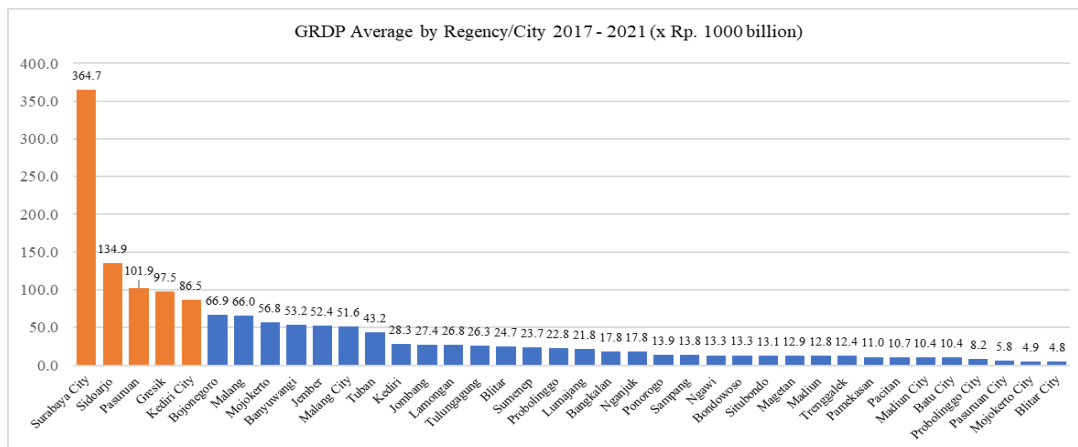
#### 3.1 Descriptive Statistics

Descriptive statistics on the attributes in this study are shown in **Table 1**. Based on the RREB structure, the Transfer to Regional and Village Fund (TRVF) has the largest average contribution compared to other attributes, namely Rp. 1588.75 billion during 2017 – 2020. The second and third largest are Employment Expenditures (EE) and Goods and Services Expenditures (GSE) with an average of Rp. 865.02 billion and Rp. 606.11 billion. While the smallest is Other Revenue of Rp. 283.63 billion. The biggest variability is Locally Generated Revenue, namely Rp. 788.50 billion, followed by the second largest variability is TRVF, namely Rp. 609.61 billion. In general, the distribution of each component of the RREB is positively skewed mainly LGR, Goods and Services Expenditures (GSE), and Capital Expenditures (CE) and have various kurtosis. While the average GRDP is Rp. 41.972.27 billion and a variance of Rp. 64.714.92 billion. GRDP variability in each regency/city in East Java Province is quite large.

**Table 1. Statistics Descriptive and RREB Structure (Rp. billion)**

Variable	Mean	Std Dev	Min	Max	Skew	Kurtosis
LGR	510.55	788.50	135.35	5381.92	4.99	26.19
TRVF	1588.75	609.61	526.84	3922.79	0.58	1.76
OR	283.63	199.03	77.19	1296.27	2.28	6.89
EE	865.02	363.15	206.06	2396.24	1.01	2.64
GSE	606.11	565.19	125.40	4239.07	4.90	26.78
CE	419.18	376.52	47.72	2754.30	4.14	20.65
OE	447.89	265.90	16.45	1364.56	0.36	0.22
RFR	314.89	314.76	0.02	2203.23	3.17	13.77
GRDP	41972.27	64714.92	4315.00	410879.30	4.21	20.07

The average GRDP for 2017 – 2020 by regency/city, ordered from largest to smallest, is shown in **Figure 3**. The city of Surabaya ranks first with an average GRDP of Rp. 364714.8 billion or almost 3 times the average GRDP of Sidoarjo as second place, which is Rp. 134878.12 billion. Regencies/cities with the third, fourth and fifth largest GRDP respectively are Pasuruan, Gresik, and Kediri City. Meanwhile, the 5 regencies/cities with the smallest average GRDP are Madiun City, Probolinggo City, Pasuruan City, Mojokerto City, and Blitar City.



**Figure 3. GRDP Average by Regency/City 2017 - 2020**

The distribution of each component of the RREB structure is shown in the boxplot in **Figure 4**. Besides having different variability, in each component there are 2 to 6 data outliers.

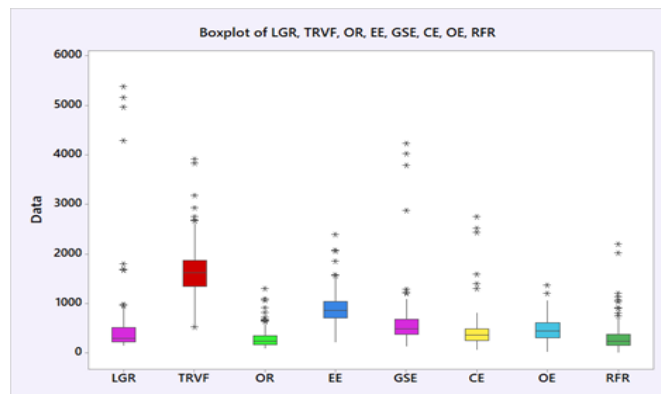


Figure 4. RREB Structure Boxplot by Component (Rp. billion)

In Figure 5, the average GRDP of East Java Province from 2017 to 2019 has increased above Rp. 1.200 billion per year or about 5%. However, due to the impact of the pandemic in 2020, the average GRDP fell by around Rp. 700 billion or 2.78%. The average economic growth in 2017 - 2019 is between 4.45% - 4.48%, and in 2020 it will decrease to -5.31%.

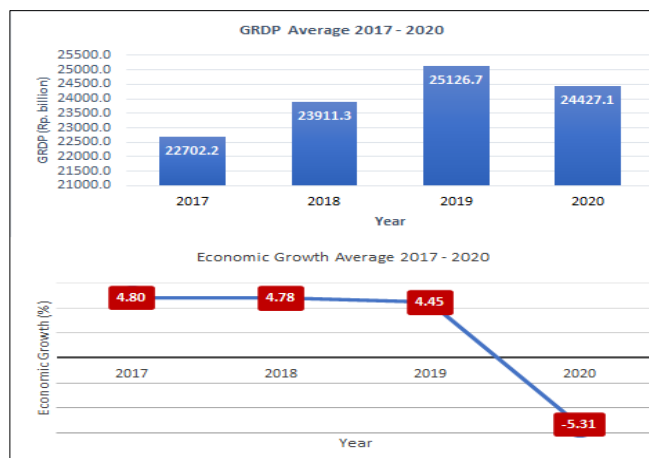


Figure 5. GRDP Average and Economic Growth Average 2017 – 2020 (Rp. billion)

The Scatterplot between GRDP and each RREB component is shown in Figure 6 and Figure 7. In general, all components have gradients that are positive with varying degrees of slope. In Figure 6, ignoring the time and area variables, partially the total variation that can be explained by LGR to GRDP is the highest, namely 92.9%. While the total variation of OR to GRDP is 56.6%.

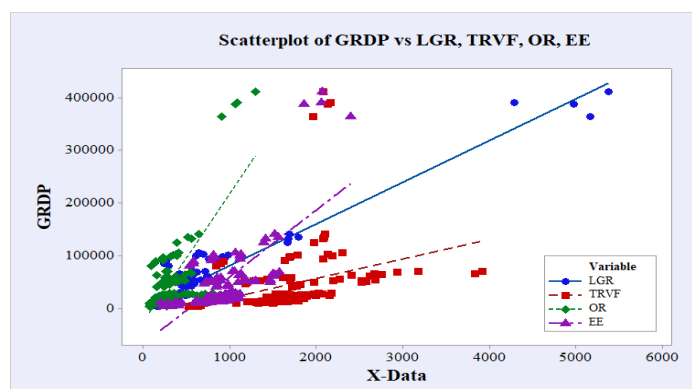


Figure 6. Scatterplot of GRDP vs LGR, TRVF, OR, and EE (Rp. billion)

Ignoring the time and area variables, partially the total variation that can be explained by GSE on GRDP is shown in Figure 7 with the highest R-squared of 89.0%. The total variation of CE to GRDP is 73.1%. While the total variation of OE and RFR is very low, respectively 1.2% and 28.4%.

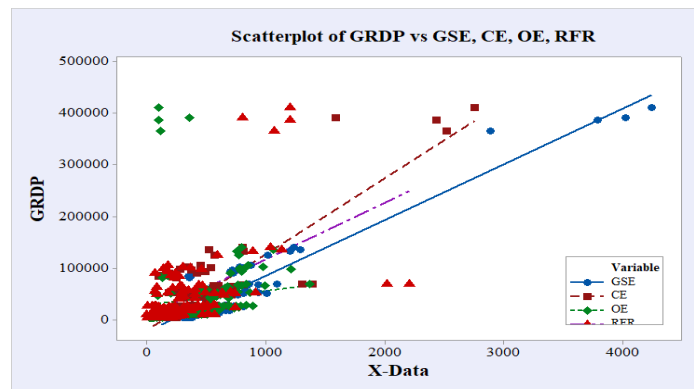


Figure 7. Scatterplot of GRDP vs GSE, CE, OE, RFR (Rp. billion)

### 3.2 Multicollinearity Test

In the original dataset, the Variance Inflation Factor (VIF) value of the ordinary least squares (OLS) model by including all the attributes of the RREB components is shown in Table 2. In the original dataset, there are 5 attributes with VIF values  $> 10$ , which means that there is multicollinearity between attributes. For this reason, a transformation is carried out into the natural logarithm. In the transformation dataset, there are 2 attributes with a VIF value  $> 10$ . This condition is better and the strategy to be applied at the modeling stage is to use the transformation dataset but not to use all the attributes in the model.

Table 2. Variance Inflation Factor (VIF)

Atributte	Variance Inflation Factor (VIF)	
	Original dataset	Transformation dataset
LGR	31.500357*	7.622151
TRVF	13.135641*	17.813869*
OR	4.581189	4.496482
EE	10.339041*	12.632361*
GSE	19.988017*	9.083803
CE	11.648862*	4.328676
OE	4.766433	6.155966
RFR	2.234953	2.005468

### 3.3 Panel Data Regression: OLS Modeling

The first panel data regression modeling to try is the CEM or OLS model. This model ignores cross-section (regency/city) and time series (time variables). Several models with  $R^2$  and  $R^2$ -adj above 80% are shown in Table 3. Although 4 models have quite high  $R^2$  and  $VIF < 10$ , only the OLS-1 model has all attribute parameters significant at  $\alpha = 5\%$ . For the OLS-2, OLS-3, and OLS-4 models, only 1 or 2 parameters are significant, respectively.

Table 3. Some OLS Models With R-Squared  $> 80\%$

OLS Model	$R^2$ -Adj	$R^2$ -Adj	Predictor Variables	VIF
OLS-1	0.8099	0.8074	LGR***	1.120691
			OE***	1.120691
OLS-2	0.8105	0.8066	LGR***	2.725564
			EE	5.077018
			OE**	2.586258
OLS-3	0.8064	0.8025	LGR***	2.286593
			TRVF***	2.044460
			OR	3.108522
OLS-4	0.8115	0.8064	LGR***	2.729575
			EE	5.641113
			OE**	2.725247
			RFR	1.357652

Significant level

\*\*\* 0.001 \*\* 0.01 \* 0.05



The total variation of the OLS model - Dummy Variables,  $R^2$  value: 99.99% and Adj.  $R^2$ : 99.99%. At the significant level  $\alpha = 5\%$ , all attributes and dummy variables are significant in the model except for the LGR attribute with  $p$ -value = 0.47294 and the Nganjuk factor (dummy variable) with  $p$ -value = 0.14669. To determine whether the CEM or FEM model is selected, the Chow test is used. Chow test results on the selected OLS model (OLS-3) obtained  $F$  value = 0.19835 and  $p$ -value = 0.9389. The conclusion is that  $H_0$  is accepted. The next step is the BP – LM test.

### 3.4 Panel Data Regression: FEM Model

The parameter estimation of the FEM model based on the selected OLS model is shown in **Table 4**, with the predictor variables of the LGR, TRVF, and OR attributes. In this model,  $R^2$  : 0.99991 and Adj.  $R^2$  : 0.99987. At  $\alpha = 5\%$  level, the attributes of LGR, TRVF, and OR are not significant in the model. That is, these attributes have little effect on the FEM model. While the area factor, all districts/cities are significant in the model. The time factor constants from 2017 to 2020 are 9.7903, 9.8385, 9.8092 and 9.8316 respectively. In the BP – LM test,  $\text{chisq} = 940$ ,  $\text{df} = 1$ ,  $p$ -value  $< 2.2e-16$ . There is a panel effect on the data.

**Table 4. Fixed Effect Model**

Atributte	Estimate	Std. Error	t-value	Pr(> t )
LGR	0.0175	0.0107	1.6364	0.1047
TRVF	-0.0071	0.0226	-0.3163	0.7524
OR	-0.0071	0.0050	-1.4428	0.1520
Factor:				
Banyuwangi	0.3054	0.0085	112.3267	< 2.2e-16 ***
Blitar	1.3092	0.0160	35.7325	< 2.2e-16 ***
Bojonegoro	-0.3164	0.0082	81.9037	< 2.2e-16 ***
Bondowoso	1.6570	0.0161	-38.5792	< 2.2e-16 ***
Gresik	1.0460	0.0135	102.6294	< 2.2e-16 ***
Jember	0.4011	0.0120	77.2815	< 2.2e-16 ***
Jombang	0.4323	0.0105	33.4149	< 2.2e-16 ***
Kediri	-0.4891	0.0211	41.1681	< 2.2e-16 ***
Batu City	-1.3681	0.0227	-23.1989	< 2.2e-16 ***
Blitar City	1.5438	0.0165	-60.3605	< 2.2e-16 ***
Kediri City	-0.5896	0.0212	93.4782	< 2.2e-16 ***
Madiun City	1.0051	0.0156	-27.7736	< 2.2e-16 ***
Malang City	-1.3447	0.0244	64.4005	< 2.2e-16 ***
Mojokerto City	-1.1600	0.0229	-55.1481	< 2.2e-16 ***
Pasuruan City	-0.8284	0.0218	-50.5918	< 2.2e-16 ***
Probolinggo City	3.0317	0.0324	-37.9468	< 2.2e-16 ***
Surabaya City	0.3827	0.0121	93.5820	< 2.2e-16 ***
Lamongan	0.1818	0.0089	31.5779	< 2.2e-16 ***
Lumajang	-0.3581	0.0086	20.4054	< 2.2e-16 ***
Madiun	-0.3489	0.0085	-41.4588	< 2.2e-16 ***
Magetan	1.2807	0.0142	-40.8742	< 2.2e-16 ***
Malang	1.1223	0.0125	90.3711	< 2.2e-16 ***
Mojokerto	-0.0319	0.0087	89.4723	< 2.2e-16 ***
Nganjuk	-0.3155	0.0079	-3.6544	0.0003996***
Ngawi	-0.5301	0.0100	-40.0059	< 2.2e-16 ***
Pacitan	-0.5008	0.0084	-53.0897	< 2.2e-16 ***
Pamekasan	1.7065	0.0129	-59.4125	< 2.2e-16 ***
Pasuruan	-0.2722	0.0080	132.4847	< 2.2e-16 ***
Ponorogo	0.2236	0.0079	-33.8760	< 2.2e-16 ***
Probolinggo	-0.2683	0.0093	28.3732	< 2.2e-16 ***
Sampang	1.9758	0.0213	-28.8875	< 2.2e-16 ***
Sidoarjo	-0.3248	0.0096	92.7144	< 2.2e-16 ***
Situbondo	0.2704	0.0084	-33.7594	< 2.2e-16 ***
Sumenep	-0.3883	0.0088	32.2057	< 2.2e-16 ***
Trenggalek	0.8602	0.0107	-44.1646	< 2.2e-16 ***
Tuban	0.3611	0.0104	80.5828	< 2.2e-16 ***
Constant of Year:				
Year	2017	2018	2019	2020
Constant	9.7903	9.8385	9.8092	9.8316

### 3.5 Panel Data Regression: REM Model

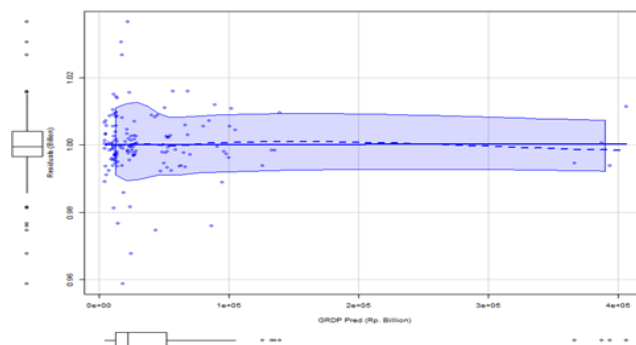
Parameter estimation of the REM model based on the selected OLS model is shown in **Table 5**, with the predictor variables being the LGR and OR attributes. The TRVF attribute is omitted by considering that the Random Effect Model dataset is only 4 years old which is only able to estimate a model with 3 parameters. On this model R-Squared: 0.78245. Adj. R-Squared: 0.77953. Based on individual testing with  $\alpha = 5\%$  level, the intercept along with the LGR and OR attributes are significant in the model. Hausman test results obtained  $\chi^2 = 549.32$ ,  $df = 2$ ,  $p\text{-value} < 2.2e-16$ . So, it is concluded that  $H_0$  is rejected so that the best model is the Fixed Effect Model.

**Table 5. Random Effect Model**

Atributte	Estimate	Std. Error	t-value	Pr(> t )
Intercept	2.5737	0.3549	7.2526	4.089e-13 ***
LGR	1.1010	0.0799	13.7800	< 2.2e-16 ***
OR	0.1937	0.0946	2.0469	0.04067 *
Effects:	var	std.dev	Share	
idiosyncratic	0.2115*	0.4599	1	

### 3.6 Heteroscedasticity Test

The next step is to ensure that there is no heteroscedasticity in the residual FEM model by testing the homogeneity of the residual variance using the Breusch-Pagan test. To check the basic assumptions of heteroscedasticity in **Figure 8** is a plot between GRDP predictions and residuals. Some of the points are outside the limits of homogeneity requirements. In the Breusch-Pagan test,  $BP = 45.815$ ,  $df = 40$ ,  $p\text{-value} = 0.2436$ . The p-value is greater than  $\alpha = 5\%$ . So, it can be concluded that there is no heteroscedasticity in the residual variance and the FEM model can be used.



**Figure 8. Scatterplot GRDP Prediction vs Residual**

### 3.7 Evaluation Model

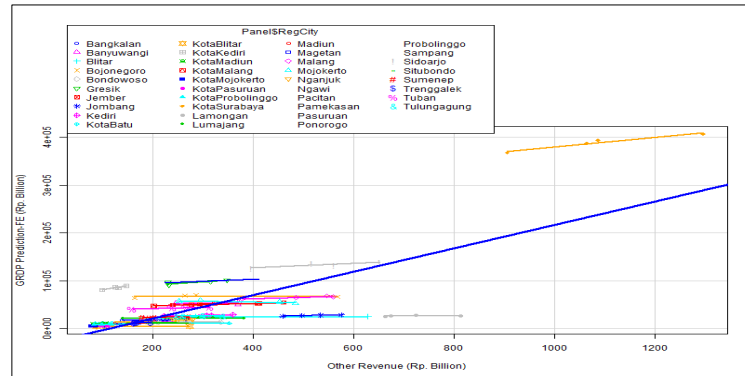
At the evaluation stage, FEM is compared with the OLS-Dummy model based on MSE, MAD, and MAPE values on the train data and test data. The results of the calculations can be seen in **Table 6**. The values of MSE, MAD, and MAPE of the Fixed Effect model on the train data are 338724.9919, 259.7182 and 0.6296, respectively. Meanwhile, the MSE value of the OLS - Dummy model in the data train is 6188603.5756 or 18 times greater than the MSE of the Fixed Effect model. The MAD and MAPE values are 1091.0522 and 2.5787 respectively or about 4 times greater than the Fixed Effect model respectively. Likewise in the test data, the MSE, MAD, and MAPE values in the Fixed Effect model are 1716324.2736, 445.7959, and 1.0692, respectively. This figure is much smaller than the OLS - Dummy.

**Table 6. Model Evaluation (FEM vs OLS - Dummy)**

Data	Model	MSE	MAD	MAPE
Train	Fixed Effect	338724.9919	259.7182	0.6296
	OLS - Dummy	6188603.5756	1091.0522	2.5787
Test	Fixed Effect	1716324.2736	445.7959	1.0692
	OLS - Dummy	20388579.3772	1831.6355	4.9556

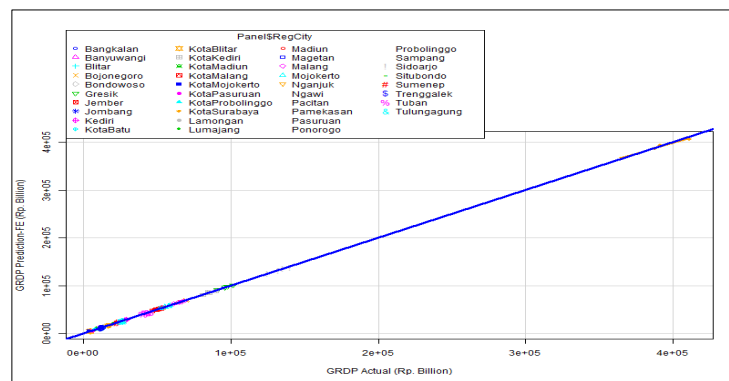
### 3.8 Model Interpretation

In the last stage is the interpretation of the model based on the prediction results using the best model, namely the Fixed Effect model. **Figure 9** is the OR scatterplot against the GRDP prediction. The blue line shows the LGR regression line on the GRDP prediction with a slope value of  $\exp(-0.0071)$  or 0.9929, which means that every increase in OR is Rp. 1 billion, there is a tendency for the average GRDP to increase by Rp. 0.9929 billion. The colorful lines in district/city data are straight-line approaches from the OR and GRDP points in each regency/city.



**Figure 9. Scatterplot Prediction of GRDP and OR**

The actual GRDP scatterplot against the predicted GRDP is shown in **Figure 10**. If the observation points for each district/city are approached by the regression line, the line will almost coincide with the panel data regression line.



**Figure 10. Scatterplot GRDP (Actual vs Prediction)**

GRDP predictions before the Covid-19 pandemic (2019) and during the Covid-19 pandemic (2020) are shown in **Figure 11**. Based on the prediction results, the average GRDP reduction in the predicted results is Rp. 1.256.3 billion with a decrease range of Rp. 118.7 – Rp. 12,907.8 billion. The pandemic in the first year (2020) had a significant impact on economic growth in the East Java regency. The biggest decline was in the City of Surabaya (Rp. 12.907.8 billion), followed by Sidoarjo (Rp. 3,642.75 billion), Gresik (2,907.35 billion), Pasuruan (Rp. 2.783.87), and the Kediri City (Rp. 2,469.94 billion). This sequence is linear with the GRDP in the five districts/cities. This also applies to other regency/cities.



In general, economic growth fell from 2018 to 2019. In 2020 economic growth fell drastically (negative) due to the Covid-19 pandemic, but in 2021 it increased significantly.

In modeling economic growth in Asian countries [7], the best model is the Fixed Effect Model (FEM) with R-squared 28.43% and Adjusted R-Squared 24.57%. Economic growth can be explained by the variables tax, government debt, household consumption and credit interest rates with a total variation of 24.57%. In the case of showing the percentage of poor people in East Kalimantan [8], the best model is the FEM model with an R-squared of 77.55%. In research [14], Adjusted R-squared was 80%. In [15], the best model is FEM with an R-square of 82.5%. The significant variables are the Human Development Index and average per capita expenditure. Meanwhile, in research [16] and [19] respectively, the Adjusted R-squared values were obtained at 40.93% and 45.49%. And in research [20], the MAPE value was 6.5%, MSE 5.5%, and RMSE 2.3%. In general, panel data regression models in previous studies obtained R-squared values below 90%. Meanwhile, in this research, FEM is also the best model with a very high R-squared value of 99.991 and Adjusted R-squared of 99.987% with a MAPE of 1.07%. What is interesting is that LGR, TRVF, and OR have less significant effects. In fact, the district/city factor has a very significant influence ( $\alpha = 0.01\%$ ).

#### 4. CONCLUSIONS

The results of this study, the Fixed Effect Model (FEM) is an appropriate model for predicting economic growth in regencies/cities in East Java. The MSE, MAD, and MAPE values of FEM are smaller than those of the OLS model with dummy variables. Cross section and period are significant to the GRDP value. What is interesting is that in the Fixed Effect model, some of the attributes that are modeled, namely Locally Generated Revenue (LGR), Transfers to Regional and Village Funds (TRVF), and Other Revenue (OR) are not significant in the model at the 95% confidence level, or it can be said to have little influence. However, all factors (cross section) have a very significant effect with a confidence level of 99.9%. This can be interpreted that local wisdom, or the characteristics of each district/city has a major contribution to economic growth. The selected model already has a very high Adjusted R-squared, namely 0.9999. That is, the total variation that can be explained by the components of the RREB structure to GRDP is very large.

The GRDP prediction results before the Covid-19 pandemic (2019) compared to the GRDP prediction results during the Covid-19 pandemic (2020) saw an average decrease of IDR. 1,256.3 billion (Rp. 118.7 – Rp. 12,907.8 billion). The pandemic in the first year (2020) had a significant impact on economic growth in East Java Regency. If the panel data regression model in this case considers several other attributes such as household consumption, investment, exports and imports, a more precise, complete, and informative model will be obtained.

#### ACKNOWLEDGMENT

The author would like to thank the Center for Research and Community Service of Politeknik Elektronika Negeri Surabaya, which has provided funding and facilities for this research.

#### REFERENCES

- [1] D. Wasani and S. I. Purwanti, "The GRDP Per Capita Gap between Provinces in Indonesia and Modeling with Spatial Regression," *J. Mat. Stat. dan Komputasi*, vol. 19, no. 1, pp. 65–78, 2022, doi: 10.20956/j.v19i1.20997.
- [2] P. Romhadhoni, D. Z. Faizah, and N. Afifah, "Pengaruh Produk Domestik Regional Bruto (PDRB) Daerah terhadap Pertumbuhan Ekonomi dan Tingkat Pengangguran Terbuka di Provinsi DKI Jakarta," *J. Mat. Integr.*, vol. 14, no. 2, p. 113, 2019, doi: 10.24198/jmi.v14.n2.19262.113-120.
- [3] E. D. A. Widyaningrum and H. Cahyono, "Pemetaan Potensi Wilayah Guna Mendorong Pembangunan Ekonomi Kabupaten Ngawi," *J. Din. Ekon. Pembang.*, vol. 3, no. 2, pp. 117–139, 2020, doi: 10.14710/jdep.3.2.117-139.
- [4] D. Junaedi and F. Salistia, "Dampak Pandemi COVID-19 Terhadap Pertumbuhan Ekonomi Negara-Negara Terdampak," *Simposium Nasional Keuangan Negara*. pp. 995–1115, 2020.
- [5] M. R. Ridha, "COVID-19 di Indonesia: aplikasi model ARIMA dan analisis kondisi perekonomian," *J. Ilmu Ekon. Terap.*, vol. 5, no. 3, pp. 416–425, 2021.
- [6] T. Tsega and T. T. Mengistie, "COVID-19 Outbreak Data Analysis and Prediction Modeling Using Data Mining Technique COVID-19 Outbreak Data Analysis and Prediction Modeling Using Data Mining Technique View project COVID-19 Outbreak Data Analysis and Prediction Modeling Using Data Mining ," *Int. J. Comput.*, vol. 38, no. 1, pp. 37–60, 2020,

- [Online]. Available: <https://www.researchgate.net/publication/341397589>.
- [7] A. Munandar, "Analisis Regresi Data Panel pada Pertumbuhan Ekonomi di Negara-Negara Asia," *J. Ilm. Ekon. Glob. Masa Kini*, vol. 8, no. 1, pp. 59–67, 2017.
- [8] E. N. Amaliah, D. Darnah, and S. Sifriyani, "Regresi Data Panel dengan Pendekatan Common Effect Model (CEM), Fixed Effect model (FEM) dan Random Effect Model (REM) (Studi Kasus: Persentase Penduduk Miskin Menurut Kabupaten/Kota di Kalimantan Timur Tahun 2015-2018)," *ESTIMASI J. Stat. Its Appl.*, vol. 1, no. 2, p. 106, 2020, doi: 10.20956/ejsa.v1i2.10574.
- [9] A. Indrasetianingsih and T. K. Wasik, "Model Regresi Data Panel Untuk Mengetahui Faktor Yang Mempengaruhi Tingkat Kemiskinan Di Pulau Madura," *J. Gaussian*, vol. 9, no. 3, pp. 355–363, 2020, doi: 10.14710/j.gauss.v9i3.28925.
- [10] K. Yunitaningtyas, A. M. Yolanda, and Indahwati, "A panel data analysis of tourism and economic development in Southeast Asian countries," *J. Phys. Conf. Ser.*, vol. 1265, no. 1, 2019, doi: 10.1088/1742-6596/1265/1/012028.
- [11] A. Narciso, "The Impact of Population Ageing on International Capital Flows," *SSRN Electron. J.*, vol. 8, no. 4, pp. 68–77, 2012, doi: 10.2139/ssrn.1702995.
- [12] B. Ioan et al., "A Panel Data Analysis on Sustainable Economic Growth in India, Brazil, and Romania," *J. Risk Financ. Manag.*, vol. 13, no. 8, p. 170, 2020, doi: 10.3390/jrfm13080170.
- [13] J. Szwacka-Mokrzycka, "The panel data regression concept in consumption modelling," *Econ. Ann.*, vol. 185, no. 9–10, pp. 61–69, 2020, doi: 10.21003/EA.V185-06.
- [14] F. A. Wardhana and R. Indawati, "Panel Data Regression Analysis for Factors Affecting Infant Mortality Rate in East Java 2013-2017," *Indones. J. Public Heal.*, vol. 16, no. 3, p. 437, 2021, doi: 10.20473/ijph.v16i3.2021.437-448.
- [15] R. Indriani and E. O. Permatasari, "Analysis of Factors Affecting Poverty Depth Index in Papua Province Using Panel Data Regression," *Proc. Int. Conf. Math. Geom. Stat. Comput. (IC-MaGeStiC 2021)*, vol. 96, pp. 196–203, 2022, doi: 10.2991/acsr.k.220202.037.
- [16] K. D. Pratiwi, "E-Commerce and Economic Growth in Indonesia: Analysis of Panel Data Regression," *GEMA PUBLICA J. Manaj. dan Kebijak. ...*, vol. 7, pp. 171–186, 2022, [Online]. Available: <https://ejournal.undip.ac.id/index.php/gp/article/view/44116%0Ahttps://ejournal.undip.ac.id/index.php/gp/article/download/44116/21211>.
- [17] A. Alamanda, "the Effect of Economic Growth on Income Inequality: Panel Data Analysis From Fifty Countries," *Info Artha*, vol. 5, no. 1, pp. 1–10, 2021, doi: 10.31092/jia.v5i1.1176.
- [18] A. Firdaus, T. C. Dawood, and M. Abrar, "Determinants of Poverty in Indonesia: An Empirical Evidence using Panel Data Regression," *Int. J. Glob. Oper. Res.*, vol. 2, no. 4, pp. 124–132, 2021, doi: 10.47194/ijgor.v2i4.90.
- [19] "Fuzzy panel data analysis Muhammet O. Yalçın\*, Nevin Güler Dincer, Serdar Demir," vol. 48, no. 3, 2021.
- [20] A. Dewi, A. Chairunnisa, and A. Fauzan, "Implementation of Panel Data Regression in the Analysis of Factors Affecting Poverty Levels in Bengkulu Province in 2017-2020," vol. 4, no. 1, pp. 40–45, 2023, doi: 10.20885/EKSAKTA.