Analysis of Factors Affecting the Open Unemployment Rate (UOR) 2022: A Case of Banten in Indonesia

A Haya\(^1\)*, R D Lestari\(^2\), T M Crisantyme\(^3\)

\(^1\) Politeknik Statistika STIS, Jl, Otto Iskandardinata No.64C, Jakarta, Indonesia

* Corresponding author’s e-mail: aqillahaya777@gmail.com

Abstract. Unemployment is one of the many economic problems. One type of unemployment is open unemployment. The Open Unemployment Rate (UOR) in Banten Province has the largest rate in Indonesia. In this study, exploratory factor analysis was used with the aim of finding out the factors that contribute to UOR in Banten Province. The data source used is secondary data obtained from publications by Badan Pusat Statistik. Factor 1 (human development) includes the Human Development Index, Economic Growth, Population Growth Rate, Literacy Rate and Mean Years of Schooling (MYS) and Factor 2 (population) consists of only one variable, that is total population. The results of this research show that total population, HDI, and MYS have the largest contribution to UOR in Banten Province. It is hoped that the government can increase business opportunities, employment opportunities and mature human development planning to reduce UOR in Banten Province.

1. Introduction

Unemployment is one of many economic problems such as poverty, income distribution gaps or other problems faced by developing countries today. One type of unemployment is open unemployment. According to BPS Indonesia (2020), open unemployment is a person who does not have a job and is looking for a job, and some who already have a job, but have not started working yet [1].

![Figure 1. Open Unemployment Rate (UOR) in Indonesia, 2018-2022.](image)
The graph above shows that the Open Unemployment Rate (UOR) in Indonesia has fluctuated in 2018-2022. In 2018-2019, the open unemployment rate in Indonesia was still around five percent. In 2020, the open unemployment rate in Indonesia increased quite high from the previous year and reached 7.07% due to the Covid-19 pandemic which hit the world, including Indonesia. In 2021-2022 there was a significant decline, but the open unemployment rate was still relatively high. So this becomes a serious problem for Indonesia.

The unemployment problem is also experienced by all provinces in the Indonesian, including several provinces in Java, including Banten Province, DKI Jakarta, West Java, Central Java, D.I Yogyakarta and East Java. According to BPS (2022), West Java Province is currently in the first position with the highest open unemployment with a rate of 8.31%, Banten Province with an open unemployment rate of 8.09%, and DKI Jakarta Province with an open unemployment rate of 7.18%. The high value of TPT in Banten Province shows that there is no balance between the number of jobs and the number of workers in Banten. In addition, the high TPT value gives an indication of the large working-age population that is included in unemployment. The high value of TPT in Banten Province is of course a serious problem that the government needs to pay attention to. A greater role for regional government and society is needed so that regional development can be achieved or implemented properly. This is because the Open Unemployment Rate (TPT) is an indicator used to assess people's social welfare [2]. Where the higher the value of TPT, the lower the level of social welfare. Unemployment is a complex problem because it influences and is influenced by several interrelated factors. One indicator that can be used to measure unemployment is the Open Unemployment Rate (TPT). Meanwhile, population density, mean years of schooling, etc. are several indicators that can explain the magnitude of the unemployment rate. The Open Unemployment Rate is a complex problem because it affects and is influenced by many factors that interact with each other so that it is not easy to understand. This study aims to determine the determinants of the Open Unemployment Rate in Banten Province.

In this study, analysis is needed to determine the factors of the Open Unemployment Rate (UOR). The analysis that can be used is factor analysis because factor analysis aims to identify relationships between variables so that they can be classified into several sets of variables that are fewer than the initial number of variables. Factor analysis is a very sophisticated way of testing construct validity, offering efforts to simplify complexity by summarizing many elements into factors that are simpler and easier to understand [3]. This research uses several factors which include social, population, economic and educational, from which several factors will be known which factors influence the Banten Open Unemployment Rate. So it is hoped that after knowing the factors that influence the Open Unemployment Rate, it will be easier to determine policies to overcome the unemployment problem in Banten. Based on the previous description, the author is interested in conducting research entitled “Analysis of Factors Affecting the Open Unemployment Rate (UOR) in Banten Province in 2022”.

2. Theoretical Background

Maria G. Egeten et al. (2023) conducted research on the Open Unemployment Rate (UOR) in cities in North Sulawesi Province. This research uses panel data regression method. The results of this research, it is known that the level of education, which is proxied by the mean years of schooling, has a significant and negative effect on UOR, so that the higher the level of education, the UOR figure will decrease, assuming the other variables have the same value or are constant. Apart from that, the variable number of labor force also has a significant and positive effect on UOR so that as the number of labor force increases, the UOR figure will also increase assuming that other factors are the same or constant [4].

Another research was conducted by Dian P. and Herniwati (2019) on the Open Unemployment Rate in Central Java Province. The research was conducted using panel data regression analysis. Based on the results of this analysis, it was found that the variables of population, education, minimum wage, and GRDP had a significant influence on the level of open unemployment in Central Java Province, where
the population variable had a positive influence, while the variables of education, minimum wage, and GRDP had an influence. which is negative for the UOR figure in Central Java Province [5].

Elika Tantri and Vita Ratnasari (2016) also conducted research regarding the factors that influence the Open Unemployment Rate (UOR) with regional coverage throughout Indonesia. The research was conducted using panel regression data analysis. In this research, the results showed that the variables Population Growth Rate, Literacy Rate and High School Gross Enrollment Rate had a significant influence on the UOR rate [6].

Another research was also conducted by Sindy T. et al (2022) regarding Modelling of Factors that Influence the UOR of the Highest Province in Indonesia as an Impact of Covid-19. The research was conducted using panel data regression analysis. The results of this analysis show different results from previous research where the variables Labor Force Participation Level, Mean Years of Schooling (MYS), Economic Growth and Population Density have a significant influence on UOR figures, while Literacy Rates have no significant effect to the number of UOR partially [7].

Research conducted by Aisyah S. and Nano P. (2017) also showed different results from previous research. This research was conducted in the research area, namely Lampung Province and used the panel data regression method. Based on the analysis, it is known that population size and HDI figures have a significant influence on the Open Unemployment Rate (UOR) while the minimum wage variable does not have a significant influence on UOR figures in Lampung Province [8].

3. Methodology
This study uses exploratory factor analysis with data from 8 regencies/cities in Banten Province in 2022. The data source used is secondary data obtained from publications by the Banten Province Central Statistics Agency (BPS). Based on studies that have been carried out in previous studies, several variables were obtained that affect the Open Unemployment Rate (UOR) in a region, that are economic growth, population growth rate, literacy rate, Human Development Index (HDI), mean years of schooling (MYS), and total population.

3.1. Latent Variables & Manifest Variables
The essential purpose of factor analysis is to describe, if possible, the covariance relationships among variables in terms of a few underlying, but unobservable, random quantities called factors. There are two types of variables, namely latent variables and manifest variables. Latent variables are variables that cannot be observed directly (unobservable) or are hidden. However, it can be assumed that latent variables/lateral variables can be explained through a mathematical model from a set of variables that can be directly observed, namely manifest variables. Manifest variables are variables that can be measured directly. Manifest variables are also indicator variables of latent variables. This variable is a variable that can represent a latent variable.

Factor analysis is a multivariate method used to analyse variables that are thought to be related to each other so that these relationships can be explained and mapped or grouped into appropriate latent factors/variables. An important goal of factor analysis is to describe the covariance relationships among many variables in terms of some random quantities called factors.

3.2. Types of Factor Analysis
There are two types of factor analysis, namely Confirmatory Factor Analysis (CFA) and Exploratory Factor Analysis (EFA). Confirmatory Factor Analysis is used to test whether a factor model specified based on a theory provides adequate fit for the covariance and correlation between manifest variables. Exploratory Factor Analysis is used to investigate the relationship between manifest variables and factors without making assumptions about the relationship between the manifest variable and a particular factor, or in other words it is used to determine the number of latent factors/variables in the model. Exploratory factor analysis is a method in exploratory research, according to Solimun et al
(2017), what is meant by exploratory research is research that has the aim of gathering information. This analysis basically carries out excavation or exploration activities on several real indicators or variables, in this way factors will be formed which will then be interpreted in order to determine the latent variables that can be obtained.

3.3. Factor Model

The common factor model states that manifest variables are a function of several latent variables. The common factor model is similar to the multiple linear regression model, with the following model:

\[ X_{ij} = \mu_i + \lambda_{ij} f_{ij} + \lambda_{i2} f_{i2} + \cdots + \lambda_{im} f_{im} + \epsilon_{ij} \]

\[ X_{i2} = \mu_2 + \lambda_{21} f_{i1} + \lambda_{22} f_{i2} + \cdots + \lambda_{2m} f_{im} + \epsilon_{i2} \]

\[ X_{ip} = \mu_p + \lambda_{p1} f_{i1} + \lambda_{p2} f_{i2} + \cdots + \lambda_{pm} f_{im} + \epsilon_{ip} \]

Or if it is made in matrix form it becomes like the following model:

\[ X_{(p \times 1)} = \mu_{(p \times 1)} + A (m \times 1) \times F_{(m \times 1)} + \epsilon_{(p \times 1)} \]

with

- \( \mu_i \) = ith variable average
- \( \epsilon_i \) = ith specific factor
- \( F_j \) = jth common factor
- \( \lambda_{ij} \) = loading of the ith variable on the jth factor
- \( A \) matriks dari faktor loading sebagai pembobot dalam fungsi

The assumptions in the model are as follows:

- \( F_i \) and \( \epsilon_i \) are independent
- \( E(F) = 0 \) dan \( E(\epsilon) = 0 \)
- \( Cov(F) = I \), the main assumption of EFA or in other words there is no correlation between factors
- \( Cov(\epsilon) = \Psi \) where \( \Psi \) is a diagonal matrix

From the model also obtained that:

\[ Cov(X) = Cov(\mu + A \times F + \epsilon) \]

\[ = Cov(\mu) + Cov(\Lambda \times F) + Cov(\epsilon) \]

\[ = \Lambda Cov(F) \Lambda' + \Psi \]

\[ = \Lambda \Lambda' + \Psi \]

From the last form it can be seen that the variation in \( X \) comes from 2 things, namely the \( F \) factor and \( \epsilon \) error. If we look at the elements that make up the covariance matrix, then

\[ Var(X_{ij}) = \sigma_{ii} = \lambda_{ij}^2 + \cdots + \lambda_{im}^2 + \psi \]

\[ \lambda_{1j}^2 + \cdots + \lambda_{mj}^2 \] referred to as communality or common variance, namely the diversity that comes from factors, and \( \psi \) is specificity or specific variance which describes the amount of variance that cannot be explained by factors.

3.4. Assumptions in Factor Analysis

- In exploratory factor analysis, multivariate normality is not required
- The variables used must be metric.
The sample must be homogeneous. Violation of this assumption increases the sample size as the number of variables increases. Reliability analysis was carried out to check homogeneity between variables.

- There are at least several pairs of variables that have a correlation of at least 0.30 and a Bartlett's Test of Sphericity (Sig.) value < α.
- There should be no outliers in the data.
- In the data set, multicollinearity can be tested using the Barlett test
- Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) value > 0.50, and the MSA value of each variable is > 0.5

3.5. Determination of Factor’s Number
- Choose m, namely the number of factors needed so that the cumulative proportion of variance reaches the specified percentage, for example 80% of the total variance.
- Choose a number of eigenvalues that is greater than the average eigenvalue. If you use an R matrix, the eigen average is 1, and if you use an S matrix, the eigen average is\[\frac{\sum_{j=1}^{p} \theta_j}{p}\]
- Use a scree plot based on the eigenvalue plot of the S or R matrix. If the graph drops sharply, then follow the sloping line, select m before the downward sloping line.
- Test the hypothesis that m is the correct number of factors

3.6. Factor Rotation
Rotation is the process of changing a solution so that the resulting solution is more interpretable without changing the underlying mathematical properties. Rotation is carried out to provide a representation of the loading value of a factor to make interpretation easier. Interpretation is easier if each variable loading has a large value on just one factor and the other factors are close to zero so that the structure becomes simpler.

There are two types of rotation, namely orthogonal rotation and oblique rotation. Orthogonal rotation (more for uncorrelated factors) is a rotation such that the angles of the two common factor axes form a 90° angle. Examples of this type of rotation method are Quartimax, Varimax, and Equimax. Oblique rotation (allowing correlated factors), namely rotation that produces an angle between the two common factor axes that does not have to be 90°.

4. Results and Discussion
Descriptive analysis on six variables was carried out to find out the general description of the characteristics of Banten Province. Information related to descriptive statistical analysis is presented in the following table.

<table>
<thead>
<tr>
<th>Regency/City</th>
<th>Total population</th>
<th>HDI</th>
<th>Economic growth</th>
<th>Population Growth Rate</th>
<th>Literacy Rate</th>
<th>MYS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pandeglang Regency</td>
<td>1307090</td>
<td>65.84</td>
<td>3.42</td>
<td>1.54</td>
<td>98.30</td>
<td>7.13</td>
</tr>
<tr>
<td>Lebak Regency</td>
<td>1433853</td>
<td>64.71</td>
<td>3.86</td>
<td>1.93</td>
<td>95.22</td>
<td>6.59</td>
</tr>
<tr>
<td>Tangerang Regency</td>
<td>3352472</td>
<td>72.97</td>
<td>5.47</td>
<td>1.87</td>
<td>99.01</td>
<td>8.92</td>
</tr>
<tr>
<td>Serang Regency</td>
<td>1678915</td>
<td>67.75</td>
<td>5.04</td>
<td>1.97</td>
<td>97.85</td>
<td>7.78</td>
</tr>
<tr>
<td>Tangerang City</td>
<td>1930556</td>
<td>78.9</td>
<td>5.98</td>
<td>1.05</td>
<td>98.36</td>
<td>10.84</td>
</tr>
<tr>
<td>Cilegon City</td>
<td>450271</td>
<td>73.95</td>
<td>4.50</td>
<td>2.01</td>
<td>98.41</td>
<td>10.34</td>
</tr>
</tbody>
</table>
From the table above, it can be seen that the highest population in Banten Province is South Tangerang City, the highest HDI is in South Tangerang City, the highest economic growth is in Tangerang City, the highest population growth rate is in Serang City, the highest literacy rate is in Kota Tangerang, and the highest mean years of schooling is in South Tangerang City.

4.1. Factor Analysis

4.1.1. Compilation of Correlation Matrix (Allowance for Variables)

a. MSA Test

Measure of Sampling Adequacy (MSA) is used to determine whether indicators can be used in factor analysis. The calculation results obtained MSA values for each variable item are as follows:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Measure of Sampling Adequacy (MSA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1 (total population)</td>
<td>0.5972</td>
</tr>
<tr>
<td>X2 (human development index)</td>
<td>0.6363</td>
</tr>
<tr>
<td>X3 (economic growth)</td>
<td>0.7460</td>
</tr>
<tr>
<td>X4 (population growth rate)</td>
<td>0.8720</td>
</tr>
<tr>
<td>X5 (literacy rate)</td>
<td>0.5077</td>
</tr>
<tr>
<td>X6 (mean years of schooling)</td>
<td>0.5592</td>
</tr>
</tbody>
</table>

Based on the MSA values in the table above, it can be seen that all variables have an MSA value greater than 0.5 so that all variables can be used.

b. Bartlett Test

\[ p\text{-value} = 0.0109 < 0.05 \]

Testing the feasibility for factor analysis using Bartlett's test of Sphericity. The Bartlett test aims to determine whether there is a correlation between variables. Based on the calculation, a p-value of 0.0109 is obtained, which is less than \( \alpha = 0.05 \) so that H0 is rejected, which means that there is a correlation between variables, or in other words it is feasible to do factor analysis.

c. Kaiser-Meyer Olkin (KMO) Test

KMO criterion: 0.5826

Checking sample adequacy uses the Kaiser Meyer Olkin (KMO) index with the criterion, \( KMO \geq 0.5 \): The sample size is sufficient for factoring. After carrying out calculations, the KMO value was obtained at 0.5826, which is greater than 0.5, so the sample size was sufficient to be factored using factor analysis.

In factor analysis, variables are needed that are correlated with each other. The higher the correlation value between variables, the greater the possibility that the variables are in the same factor. To determine whether a variable is correlated with other variables, the Bartlett test is carried out. From the results obtained in the Bartlett test, the p-value has a value of less than 0.05 so it can be said that the variables used are correlated and can be continued to the next process. KMO testing aims to test data adequacy requirements, both measuring the adequacy of sampling as a whole and sampling for each indicator. The KMO value is classified as sufficient if it is
greater than 0.5. Based on the output results, it can be seen that the KMO value is 0.5826 which is greater than 0.5, so the data indicates that the sample size is sufficient to proceed to the next stage of factor analysis. Based on the MSA value, there are all variables with a value of more than 0.5 so that these six variables can continue with Factor Analysis.

4.1.2. Factor Extraction. Components with Eigen Value ≥ 1

<table>
<thead>
<tr>
<th>Order of eigenvalues</th>
<th>Eigenvalues</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.6597</td>
</tr>
<tr>
<td>2</td>
<td>1.1673</td>
</tr>
<tr>
<td>3</td>
<td>0.5999</td>
</tr>
<tr>
<td>4</td>
<td>0.4995</td>
</tr>
<tr>
<td>5</td>
<td>0.0653</td>
</tr>
<tr>
<td>6</td>
<td>0.0083</td>
</tr>
</tbody>
</table>

Based on the output above, it can be seen that there are two variables that have an eigenvalue >1 so that from this criterion the number of new indicators formed is 2 indicators.

Figure 2. Scree Plot

It is known that the scree plot starts to flatten on the extraction of the initial variables into 2 factors.

4.2. Determining Factors

<table>
<thead>
<tr>
<th>Variable</th>
<th>Factor Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1 (total population)</td>
<td>0,9950</td>
</tr>
<tr>
<td>X2 (human development index)</td>
<td>0,9910</td>
</tr>
<tr>
<td>X3 (economic growth)</td>
<td>0,8440</td>
</tr>
<tr>
<td>X4 (population growth rate)</td>
<td>-0,5930</td>
</tr>
<tr>
<td>X5 (literacy rate)</td>
<td>0,6220</td>
</tr>
<tr>
<td>X6 (mean years of schooling)</td>
<td>0,9840</td>
</tr>
</tbody>
</table>
With a variable that has a loading factor, the largest variable is included in the factor group. Based on the table above, it can be seen that variable X1 (population) has the highest loading factor. This is in accordance with the results of research which states that population has a significant effect on UOR in Banten Province [10]. Population growth automatically affects the increase in the labor force, which is directly related to the problem of employment opportunities and the unemployment rate. Population growth and the work rate affect the problem of unemployment and expanding employment opportunities [11]. Variable X2 (HDI) has the highest loading factor, both of which are in accordance with the results of previous research where HDI had a significant effect on unemployment in Banten Province [12]. The mean years of schooling variable has the third largest contribution to the unemployment rate. This is supported by research where the mean years of schooling has a significant effect on the open unemployment rate in Banten Province [13]. It can also be seen that variable X4 (population growth rate) has the lowest factor loading value compared to other variables, meaning that the rate of economic growth in this study has the lowest contribution to UOR in Banten Province when compared to other variables in this study. The results of research prove that economic growth has a significant negative effect on the unemployment rate [14]. This means that any increase in economic growth in a region will have an impact on reducing the number of open unemployed. The results of research conducted by Wardiansyah et al. (2016) also prove the same thing that economic growth can reduce unemployment rates [15].

<table>
<thead>
<tr>
<th>Table 4. Factor Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor 1</strong></td>
</tr>
<tr>
<td>SS loadings</td>
</tr>
<tr>
<td>proportion var</td>
</tr>
<tr>
<td>cumulative var</td>
</tr>
</tbody>
</table>

From the output above, it explains that there are 2 components formed. The proportion of variance of Factor 1 is 0.5680, which means that the first factor is able to explain the variance of all variables by 57%. Then the second factor is able to explain from all the variables observed by 21%. It was also found that each component has its own percentage of variance where the total variance of the 2 components above is around 78%.

4.3. Rotation Process

**Figure 3.** factor rotation
Based on the figure above, there are two forms of rotation, the first is orthogonal rotation, namely varimax rotation, which describes a limiting method so that the factors being rotated remain uncorrelated, while the oblique rotation, namely promax, describes a method that allows correlated factors after rotation. The two rotations above show what is included in factor one, namely the variable whose loading value is close to \(-1\) or \(1\) on the \(x\) axis, these variables are the human development index (X2), economic growth (X3), population growth rate (X4), literacy rate (X5), and mean years of schooling (X6). In factor one there is an educational side, an economic side and a social side, so the name of this variable is human development. While those included in factor two can be seen that the loading value is close to \(1\) on the \(y\) axis, the variable is the population (X1), so that the naming of factor two is adjusted to the name of the variable, namely the population factor.

Factor 1 (Human Development) has the highest loading factor, both of which are in accordance with the results of previous research where HDI had a significant effect on unemployment in Banten Province. If the human development index increases, the unemployment rate in Banten Province will decrease. This is in accordance with the theory explained by (Todaro, 2000) that through increasing human capital development and development to increase human productivity. Through investment in education, it is hoped that it will be able to improve the quality of Human Resources (HR), which is demonstrated by increasing a person's knowledge and skills, which will encourage an increase in work productivity. Increasing productivity can affect employment opportunities, namely by increasing productivity there will be a decrease in production costs per unit of goods. If the price of goods falls, demand for goods rises, which will encourage entrepreneurs to increase demand for labor, so that by absorbing more labor, the high unemployment rate can be reduced [12]. Factor 2 shows that the population variable has a positive and significant influence on the open unemployment rate, namely that an increase in population will increase the unemployment rate. This is because the increasingly rapid population will produce more and more workers, but this is not balanced with existing job opportunities. With a small number of job opportunities, people compete with each other to get jobs and people who are marginalized in this competition become unemployed [5].

Communality is the amount of variation in indicators that can be explained by the factors formed. The results of calculating the communality value are as follows:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Communalities (hi (^2))</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1 (total population)</td>
<td>0.0430</td>
</tr>
<tr>
<td>X2 (human development index)</td>
<td>0.9400</td>
</tr>
<tr>
<td>X3 (economic growth)</td>
<td>0.7170</td>
</tr>
<tr>
<td>X4 (population growth rate)</td>
<td>0.3870</td>
</tr>
<tr>
<td>X5 (literacy rate)</td>
<td>0.4240</td>
</tr>
<tr>
<td>X6 (mean years of schooling)</td>
<td>0.8650</td>
</tr>
</tbody>
</table>

The table above shows the magnitude of a variable that can explain a factor. The population variable value is 0.0430, which means that the population can explain a factor of 4.30 percent. The human development index variable value is 0.9400, which means the human development index can explain 94.00 percent of the factors. The economic growth variable has a value of 0.7170, which means that economic growth can explain 71.70 percent of the factors. The population growth rate variable has a value of 0.3870, which means that the population growth rate can explain a factor of 38.70 percent. The literacy rate variable has a value of 0.4240, which means that the literacy rate can explain a factor of
42.40 percent. The variable mean years of schooling has a value of 0.8650, which means that mean years of schooling can explain a factor of 86.50 percent.

5. Conclusion
The factors that influence the Open Unemployment Rate (UOR), especially in Banten Province, of the six variables are classified into two factors. Factor analysis establishes the two factors labelled as human development and population size. Factor 1 (human development) includes the Human Development Index, Economic Growth, Population Growth Rate, Literacy Rate and Mean Years of Schooling and Factor 2 (population) consists of only one variable, namely population. The variable that contributes most to the Banten Province UOR is population, followed by HDI and mean years of schooling, meaning that the education sector is the second and third highest influence on UOR after population. Economic growth is the fourth contributor, this means that the economic sector has an influence on UOR Banten Province. Finally, the literacy rate and population growth rate have the least influence on UOR in Banten Province.

Based on the conclusions in this research, the Banten Provincial government is advised to make improvements or improvements with a large contribution to the level of open unemployment, namely by increasing business opportunities and job opportunities and maturing development planning, one of which is increasing the mean years of schooling, which are expected.

References

