Journal of Mathematics and Applied Statistics
Vol. 2, No. 2, December 2024.
DOI: https://doi.org/10.35914/mathstat.v2i2.234

Principal Component Analysis: Identifiying Underlying Issues that Lead to Divorce in South Sulawesi

Eka Purnama

Institut Agama Islam Negeri Sultan Amai Gorontalo, Gorontalo, Indonesia

Email: ekapurnama@iaingorontalo.ac.id

ABSTRACT

This study aims to analyze the various factors contributing to divorce in South Sulawesi, with the goal of identifying and understanding the underlying issues that lead to marital breakdowns in the region. Utilizing Principal Component Analysis (PCA) to reduce the variables associated with divorce, thirteen factors were examined. The analysis revealed five key components: moral issues, personal will, differences of opinion, economic challenges, and criminality. Notably, these components accounted for 82% of the variance in the data, highlighting their significance in understanding the dynamics of divorce in South Sulawesi. This research provides valuable insights for policymakers and practitioners working to enhance marital stability in the region. For upcoming studies, advancements in PCA, such as kernel PCA applied to manage nonlinear data, or robust PCA to tackle outlier problems. Furthermore, the dataset could be expanded by incorporating a wider range of factors of divorce.

Keyword: Principal Component Anaysis (PCA), Variable Reducing, Divorce

Article History:
Received, 06-11-2024; Revised, 18-11-2024; Accepted, 24-11-2024.

1. Introduction

Divorce is a significant aspect of family dynamics that many couples encounter. While couples typically enter marriage with the intention of creating a harmonious and lasting relationship, they often face various challenges that can lead to the difficult decision to separate. According to Article 207 of the Indonesian Civil Code, divorce is defined as the termination of a marriage through a court ruling requested by one party, based on legally established grounds (Sofyan & Torro, 2022). This highlights that divorce is not only a personal matter but also involves legal considerations that every couple should be aware of.

In South Sulawesi, divorce continues to be a major issue that significantly impacts social order and family dynamics. It not only affects familial relationships but also deeply influences the emotional development of children caught in the separation. Children frequently feel confusion and sadness, which can hinder their growth. Moreover, the community in South Sulawesi, known for its strong family values, often stigmatizes those who have experienced divorce. This stigma can create additional emotional challenges for both parents and children as they try to adjust to their new situations.

Data from the Central Statistics Agency of South Sulawesi reveals a significant number of divorce cases in the region. In 2023, there were 14,612 recorded divorces, with 12,806 documented by The Directorate General of Religious Justice Affairs. While this represents a decrease from previous years—17,358 cases in 2022 and 15,575 cases in 2021, especially during the Covid-19 pandemic—South Sulawesi still has the highest divorce rate in Eastern Indonesia. Among the total cases, 11,483 were divorce petitions, making up approximately

78.6% (Badan Pusat Statistik Sulawesi Selatan, 2024). This not only indicates a high divorce rate but also suggests a trend where many women are proactively seeking divorce. This scenario underscores the complex social dynamics in South Sulawesi and highlights the need for a comprehensive approach to address divorce-related issues, particularly to mitigate the effects on children and families involved.

There are several reasons that contribute to divorce, including commitment, infidelity, conflict and arguing, marrying too young, financial problem, substance abuse, domestic violence (Scott et al., 2013). Other factors caused by sudden character change, lack of and poor communication, financial problems, abuse and infidelity were associated with divorce among young couples (Mohlatlole et al., 2018). Divorce become unending phenomenon in human society, several factors also lead to divorce and the dissolution of marriage. Education and employment among women or wives, different education levels among couples, employment among couples, interference from in-laws and other dependent family members, cultural issues, marital rape or sexual abuse, infertility, barrenness, and drug abuse (Emeng, 2022).

The various reasons reflect the complexity of the problems faced by couples in a marriage. To better understand the phenomenon of divorce, it is important to simplify these variables into clearer categories. For example, these causes can be grouped into categories such as moral issues, economic problems, and interpersonal conflicts. With this approach, it is hoped that we can gain a more comprehensive picture of the factors contributing to divorce, as well as formulate more effective strategies for prevention and intervention. A deeper understanding of these factors will enable relevant parties, such as social institutions and the government, to provide more appropriate support for couples experiencing difficulties in their households.

Principal Component Analysis (PCA) is a multivariate technique that analyzes a dataset where observations are described by several inter-correlated quantitative dependent variables (Abdi & Williams, 2010). The main goal of PCA is to simplify data analysis by decreasing complexity without losing essential information. PCA accomplishes this by converting the original variables into a reduced set of uncorrelated variables known as principal components, which capture the key variations in the data (Jolliffe, 2002). The principal components reflect the information from the original variables, making interpretation and further data processing easier. By using this analysis, researchers can more readily uncover patterns and relationships in the data that might otherwise be challenging to identify, enhancing the efficiency of data-driven decision-making. PCA is particularly beneficial across various disciplines, from social sciences to natural sciences, where complex data is common.

The study uses principal component analysis to reduce data, specifically in health, to distinguish between children with Congenital Adrenal Hyperplasia (Ljubicic et al., 2021). PCA can also be applied to meteorological forecasting. The pressure gradient, temperature gradient, and high-altitude guiding winds are the three main factors that influence strong winds, forming three groups of strong wind prediction factors. Principal Component Analysis (PCA) transforms multiple indices into a smaller set of linearly independent comprehensive indices (Ling et al., 2021). In ecology, it is used to explore environmental data in Saihanba Forest, China, to develop environmental assessment models (Liu et al., 2022). In the fields of mechanics and engineering, PCA (Principal Component Analysis) visualization is employed to identify the key chemical elements that can improve the monitoring of machine productivity and efficiency. The analysis was conducted on five machines to determine the primary factors contributing to wear. The results revealed that Calcium (Ca) and Magnesium (Mg) are the two dominant elements responsible for the wear in the machines (Parhusip et al., 2022). In social studies, PCA has been used to analyze social and behavior rhythms in the elderly. The aim was to analyze the factorial structure of the Italian version of Brief Social Rhythm Scale (BSRS). Concerning the Analysis of Principal Components, the study shows all the items of the instrument correlated to a single component and confirms that the BSRS is consistent and can be useful for the study of the regularity of SBR in old adults. (Cossu et al., 2022).

2. Method

This type of research is quantitative and employs Principal Component Analysis (PCA) to conduct an in-depth analysis of the data. Through this method, the researcher aims to identify patterns and relationships among the involved variables, while reducing data complexity without losing important information.

Principal Component Analysis (PCA) is a multivariate method that examines a data table where observations are characterized by multiple interrelated quantitative dependent variables. Its purpose is to extract significant information from the table and express it as a collection of new orthogonal variables known as principal components (Palazzolo et al., 2023). The concept behind PCA is straightforward, it aims to reduce the dimensionality of a dataset while retaining as much variability (i.e., statistical information) as possible (Jollife & Cadima, 2016). By using PCA, information from several observed variables on the same subjects can be combined into a smaller number of variables, making data analysis and interpretation easier (Greenacre et al., 2022).

In the initial phase of this research, secondary data collection was carried out by gathering information related to the number of divorce cases. This data may include divorce statistics from various sources, such as reports from the Central Statistics Agency of South Sulawesi. By collecting this data, the researcher can begin to analyze existing patterns and identify factors contributing to divorce. Once the data is gathered, the next step is to apply Principal Component Analysis to extract the most relevant components and understand the dynamics behind the phenomenon of divorce occurring in society.

This research was conducted by collecting data on the number of divorce cases from 24 regency/municipality in South Sulawesi. The data was obtained from the Central Statistics Agency of South Sulawesi. In this study, thirteen variables contributing to divorce were analyzed, namely: adultery (X_1), drunk (X_2), drugs addict (X_3), gambling (X_4), split up (X_5), jail (X_6), polygamy (X_7), domestic violence (X_8), disability (X_9), constant disputes and quarrel (X_{10}), forced marriage (X_{11}), apostate (X_{12}), and economic (X_{13}).

Once the data is gathered, Principal Component Analysis (PCA) is performed to pinpoint significant components and minimize data dimensions while retaining essential information. RStudio software supports this analysis, enabling the researcher to efficiently manage the data and achieve precise outcomes. The analysis aims to reveal underlying patterns in divorce cases in South Sulawesi and highlight the key factors contributing to this issue.

Let $X = (X_1, X_2, \dots X_m)$ e a vector consisting of m original variables transformed into a vector of n new variables $C = (C_1, C_2, \dots C_n)$ where $n < m$. This relationship can be expressed as follows:

$$C_i = b_{1i}X_1 + b_{2i}X_2 + \cdots + b_{mi}X_m \tag{1}$$

or

$$\begin{aligned} C_1 &= b_{11}X_1 + b_{21}X_2 + \cdots + b_{m1}X_m \\ C_2 &= b_{12}X_1 + b_{22}X_2 + \cdots + b_{m2}X_m \\ &\vdots \\ C_n &= b_{1n}X_1 + b_{2n}X_2 + \cdots + b_{mn}X_m \end{aligned} \tag{2}$$

where:
X_m represent the m initial variables
C_i represent the i-th component, where $i = 1,2, \dots, n$
b_{mn} represent the weights of the variables from m to n
n is the number of new components/variables.

The Principal Component Analysis (PCA) process consists of consists of several stages consisting of data preparation, data standardization, forming a variance-covariance matrix, calculating eigenvalues and eigenvectors to determine the number of new components, selecting the principal components and interpreting the principal components.

1. Data Collection

In this initial stage, relevant data is gathered and prepared in a suitable format for analysis. It is crucial to ensure that the collected data encompasses all necessary variables and is representative of the phenomenon being studied. Additionally, data cleaning is performed to address any missing values or outliers that could impact the analysis results.

2. Data Standardization

Once the data is collected, the next step is standardization. This is essential when the existing variables have different scales. Standardization helps transform the data into a uniform scale, allowing each variable to contribute proportionally to the analysis. This process typically involves subtracting the mean from each variable and dividing by its standard deviation. As a result, PCA can identify patterns and structures within the data without bias introduced by scale differences.

PCA relies on a strict yet effective assumption: linearity. This assumption simplifies the problem significantly by narrowing down the possible bases. As a result, PCA is confined to representing the data as a linear combination of its basis vectors (Shlens, 2014). Data standardization ensures the linearity of variables by putting them on the same scale.

3. Variance-Covariance Matrix

After the data is standardized, the next step in PCA analysis is to form the variance-covariance matrix. This matrix serves as the foundation for analyzing the relationships between the existing variables. By using the variance-covariance matrix, we can measure the strength of the relationship between two variables, providing deep insights into their interactions (Zhang et al., 2018). Each element in this matrix reflects the covariance between pairs of variables, indicating whether these variables tend to increase or decrease together. This matrix is an important step before performing eigen decomposition, which will help us identify the most significant principal components.

The variance of a variable can be calculated using the formula (Johnson & Wichern, 2007):

$$S_{jj} = \frac{1}{m-1}\sum_{j=1}^{m}\left(X_{ij} - \bar{X}_J\right)^2 \tag{3}$$

Where:

S_{jj} is the variance.

m is the number of observations.

X_{ij} represents each individual observation.

$\bar{X}_J$ is the mean of the observations.

The covariance between two variables can be calculated using the formula:

$$S_{ij} = \frac{1}{m-1}\sum_{j=1}^{m}\left(X_{ij} - \bar{X}_I\right)\left(X_{ij} - \bar{X}_j\right) \tag{4}$$

Where:

S_{ij} is the covariance between variables i-th dan j-th

m is the number of observations.

X_{ij} represents each individual observation.

$\bar{X}_I$ is the mean of the observations of I-th

$\bar{X}_J$ is the mean of the observations of J-th

The variance-covariance matrix S can be expressed with the following equation:

$$S_{m\times m} = \begin{bmatrix} S_{11} & S_{12} & \dots & S_{1m} \\ S_{21} & S_{22} & \dots & S_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ S_{m1} & S_{m2} & \dots & S_{mm} \end{bmatrix} \tag{5}$$

The matrix S is symmetric, with variances along the diagonal and covariances in the off-diagonal positions.

The correlation matrix is used to measure the linear relationship between two variables. The correlation coefficient r can be calculated using the following formula (Liu et al., 2022):

$$r = \frac{\sum_{j=1}^{m}\left(X_{ij} - \bar{X}_I\right)\left(X_{ij} - \bar{X}_j\right)}{\sqrt{\sum_{j=1}^{m}\left(X_{ij} - \bar{X}_I\right)^2 \sum_{j=1}^{m}\left(X_{ij} - \bar{X}_J\right)^2}} \tag{6}$$

Where:

r is is the correlation coefficient between variables

X_{ij} represents each individual observation.

$\bar{X}_I$ is the mean of the observations of I-th

$\bar{X}_J$ is the mean of the observations of J-th

The correlation matrix can be expressed as follows:

$$R_{m\times m} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mm} \end{bmatrix} \tag{7}$$

4. Calculation of Eigenvalues and Eigenvectors

The next step in PCA involves computing the eigenvalues and eigenvectors of that matrix. These eigenvalues and eigenvectors help identify the most important principal components within the data, allowing researchers to reduce the dimensionality while retaining crucial information. This analysis can yield valuable insights into the factors that lead to divorce in South Sulawesi. Eigenvalues reflect the amount of variance accounted for by each principal component, whereas eigenvectors show how much each original variable contributes to those components.

An eigenvalue of a square matrix A is a scalar λ such that there exists a non-zero vector $\mathbf{v}$ (the eigenvector) for which the following equation holds:

$$A\mathbf{v} = \lambda\mathbf{v} \tag{8}$$

This means that when the matrix A acts on the vector $\mathbf{v}$, it merely scales it by the factor λ.

An eigenvector of a matrix A corresponding to the eigenvalue λ is a non-zero vector $\mathbf{v}$ that satisfies the equation above (Jiang, 2022).

5. Selection of Principal Components

Researchers can decide how many principal components to keep by analyzing the eigenvalues and cumulative variance proportions, often aiming for a cumulative variance exceeding 80%. Alternatively, a scree plot can be employed to assess the eigenvalues, where components with eigenvalues greater than 1 are generally regarded as significant and chosen as principal components (Varmuza & Filzmoser, 2009).

6. Interpretation of Components

After the principal components are identified, the next step is to interpret and assign names to each component based on the variables that have significant weights. These components are linear combinations of the original variables, and for each principal component, the factor loadings will be calculated to indicate the contribution of each original variable to that component. This approach will provide clearer insights into the factors influencing divorce.

3. Result and Discussion

3.1 Descriptive Analysis

The collected data was analyzed descriptively to gain insights into its characteristics. This study focuses on 13 variables that represent different factors related to divorce. Descriptive statistics for these variables are presented in Table 1 below:

Table 1. Descriptive Statistics

Variable	Variable Labels	Min	Max	Mean	St Dev
X_1	Adultery	0	11	0.9167	2.244155
X_2	Drunk	0	45	8.083	9.183713
X_3	Drug Addict	0	9	1.167	2.03591
X_4	Gambling	0	7	2.50	2.303117
X_5	Split up	0	301	73.79	70.93106

X_6	Jail	0	7	0.875	1.596532
X_7	Polygamy	0	20	2.542	4.373628
X_8	Domestic Violence	0	78	13.88	16.63237
X_9	Disability	0	3	0.2917	0.6902531
X_{10}	Constant Disputes and Quarrel	0	1924	410.7	393.0821
X_{11}	Forced Marriage	0	4	0.875	1.423789
X_{12}	Apostate	0	5	1.375	1.555146
X_{13}	Economy	0	44	16.62	15.69772

Table 1 provides information on 13 variables that affect the number of divorces across 24 districts and cities in South Sulawesi. Among the variables examined, several were identified as significant contributors to divorce. Notably, variable X_{10}, which pertains to ongoing disputes and conflicts between partners, shows a substantial impact, with a total of 1,924 divorce cases linked to it. Additionally, variableX_5, associated with one partner leaving, also plays a critical role, accounting for 301 cases. These results indicate that unhealthy relationship dynamics and emotional instability are key factors driving divorce in the region.

3.2 Data Standardization

Standardization aims transform the data into a uniform scale. Using RStudio, the results of data transformation are obtained in Table 2 as follow:

Table 2. Data Standardization

Regency	X_1	X_2	X_3	...	X_{13}
Selayar	-0.4085	-0.8801	-0.5730	...	-1.0591
Bulukumba	-0.4085	0.4264	-0.0819	...	0.0239
Bantaeng	4.4932	4.0198	-0.5730	...	-0.2309
⋮	⋮	⋮	⋮	⋱	⋮
Palopo	-0.4085	0.3180	-0.0819	...	1.3617

PCA assumes linearity in the relationships between variables. Therefore, if the relationships between variables in the data are nonlinear, PCA will not be able to capture important patterns. The standardized data will meet the assumption of linearity and ensure that the data has the same scale.

3.3 Variance Covariance Matrix

Calculating the variance-covariance matrix measures the relationships between pairs of variables. Prior to this calculation, it is essential to standardize the data. Standardization is important when variables are measured on different scales, as it ensures that each variable contributes equally to the analysis. Using RStudio, both the variance-covariance matrix and the correlation matrix were generated, as illustrated in Table 3 below:

Table 3. Correlation Matrix

Variable	X_1	X_2	X_3	...	X_{13}
X_1	1.0000	0.8315	0.0127	...	0.0127
X_2	0.8315	1.0000	0.1244	...	0.2909
X_3	0.0127	0.1244	1.0000	...	0.4482
⋮	⋮	⋮	⋮	⋱	⋮
X_{13}	0.0127	0.2909	0.4482	...	1.0000

Table 3 shows the existence of variables that are strongly correlated, such as X_1 and X_2, which have a correlation value of 0.8315. This high value signifies a meaningful relationship between the two, implying that X_1 and X_2 tend to change in similar ways. As a result, these variables can be combined into a single component for additional analysis, enhancing the understanding of their interactions and influences, and simplifying the interpretation of the research findings.

3.4 Eigen Values and Eigen Vectors

The next step involves calculating the eigenvalue decomposition and the cumulative proportions of variance, which are presented in Table 5 below:

Table 5. Eigen Value, Variance Proportion and Cumulative Proportion

Component	Eigen Value	Variance Proportion	Cumulative Proportion
PC_1	4.0571	0.3121	0.3121
PC_2	2.6224	0.2017	0.5138
PC_3	1.5873	0.1221	0.6359
PC_4	1.3363	0.1028	0.7387
PC_5	1.0682	0.08217	0.82088
PC_6	0.6200	0.0477	0.8686
PC_7	0.5636	0.04335	0.91193
PC_8	0.4139	0.03184	0.94377
PC_9	0.3302	0.0254	0.9692
PC_{10}	0.2122	0.01633	0.98550
PC_{11}	0.1033	0.00795	0.99345
PC_{12}	0.0451	0.00347	0.99692
PC_{13}	0.0400	0.00308	1.00000

Table 5 presents information regarding the eigenvalues, variance proportions, and cumulative proportions for each component formed. Although there is no standard guideline for determining the number of components based on cumulative variance proportions, a common range considered is between 70% and 90%. In this study, it has been established that the number of components selected is based on cumulative variance proportions exceeding 80%. Thus, from the analysis conducted, five components were identified that meet this criterion. This decision is crucial to ensure that the selected components can explain the majority of the variance in the data, making the results more representative and relevant for further interpretation. Using this criterion helps simplify the model while retaining significant information from the existing variables.

3.5 Selection of Principal Components

The determination of the number of principal components can also be conducted through a scree plot analysis, which illustrates the eigenvalues of each component. In this method, components with eigenvalues greater than one are considered to have a significant contribution to the total variance. Figure 1 shows the scree plot of eigenvalues generated from the analysis. The graph clearly indicates an inflection point that helps in identifying the optimal number of components. The point at which the eigenvalues begin to decline significantly signifies that sub sequent components contribute less to the variance, thus serving as a reference for determining the relevant number of components for further analysis.

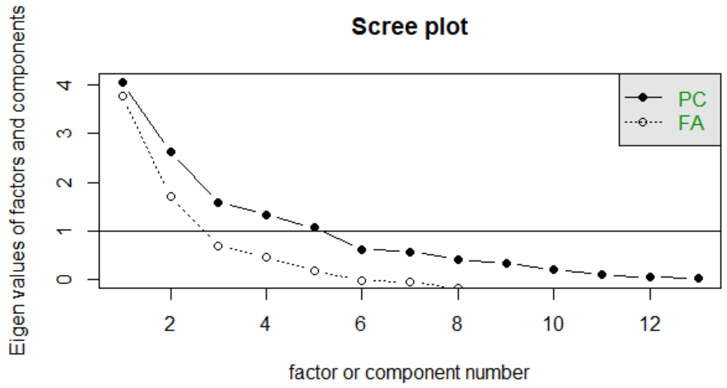

Figure 1. Scree plot of eigen value

Figure 1 clearly shows that there are five components with eigenvalues exceeding 1, which indicates their significance and the ability to create five distinct factors. Once the number of components is determined, the next step is to conduct a rotation. This rotation aims to optimally group the variables into new components, clarifying the factor structure and making the analysis results easier to interpret. By rotating, each variable can be aligned with the most relevant component, improving the clarity and significance of the relationships among the variables in this study. This step is essential to ensure that the resulting components accurately represent the patterns found in the data.

3.6 Interpretation of Components

From the analysis of the 13 variables conducted, five main components were successfully formed along with their decomposed variables. The following Table 6 shows the distribution of these variables into each component. Each component reflects a group of interrelated variables that share similar characteristics. By mapping the variables into these components, the study can provide deeper insights into the data structure and the relationships between the variables.

Table 6. Formed Principal Components

Component	Component Labels	Variable	Variable Labels	Weight	Variance Proportion	Explained Proportion
$\boldsymbol{PC_1}$	Moral Factor	X_1	Adultery	0.95	0.29	0.36
		X_2	Drunk	0.92		
		X_4	Gambling	0.57		
		X_7	Polygamy	0.88		
		X_8	Domestic Violence	0.95		
$\boldsymbol{PC_2}$	Personal Will Factor	X_5	Split up	0.89	0.17	0.20
		X_9	Disability	0.91		
		X_{11}	Forced Marriage	0.59		
$\boldsymbol{PC_3}$	Differences of Opinion	X_{10}	Constant Disputed and Quarrel	0.88	0.13	0.16
		X_{12}	Apostate	0.81		
$\boldsymbol{PC_4}$	Economic Factor	X_3	Drug Addict	0.91	0.13	0.15
		X_{13}	Economy	0.73		
$\boldsymbol{PC_5}$	Criminality	X_6	Jail	0.86	0.10	0.12

Table 6 shows that there are five new components formed, each representing the original variables. The new components formed are as follows:

1. **First Component (PC_1)**: This component consists of five variables: adultery (X_1), drunk (X_2), gambling (X_4), polygamy (X_7), and domestic violence (X_8). This component is named the **Moral Factor**, which encompasses behavioral aspects related to norms and ethics in relationships.
2. **Second Component (PC_2)**: This component encompasses of the variables of split up (X_5), disability (X_9), and forced marriage (X_{11}). This component is named the **Personal Will Factor**, which reflects physical and emotional conditions, as well as dissatisfaction or incompatibility that can influence divorce decisions.
3. **Third Component (PC_3)**: This component consists of the variables of constant disputes and quarrels (X_{10}) and apostasy (X_{12}). This component is called the **Disagreement Factor**, which describes conflicts arising from differences in views and beliefs between partners.
4. **Fourth Component (PC_4)**: This component consists of two variables: drug addict (X_3) and economic (X_{13}). This component is named the **Economic Factor**, which reflects financial challenges and substance use that can impact the relationship.
5. **Fifth Component (PC_5)**: This component consists of one variable: imprisonment (X_6), and is referred to as the **Criminality Factor**. It reflects the impact of criminal actions that can affect the stability of relationships and families.

By establishing these components, the research can offer more focused insights into the factors that lead to divorce in South Sulawesi. This structured method enables us to pinpoint the critical elements that contribute to a higher divorce risk. The five identified components can be arranged into the following equation:

$$
\begin{aligned}
PC_1 &= 0.95X_1 + 0{,}92X_2 + 0.57X_4 + 0.88X_7 + 0.95X_8 \\
PC_2 &= 0.98X_5 + 0.91X_9 + 0.59X_{11} \\
PC_3 &= 0.88X_{10} + 0.81X_{12} \\
PC_4 &= 0.91X_3 + 0.73X_{13} \\
PC_5 &= 0.86X_6
\end{aligned} \tag{9}
$$

The equation above is based on the highest weights of each variable that show a positive correlation. These five factors explain 82% of the variance. The findings suggest that about 18% of the factors contributing to divorce are influenced by elements outside the variables examined in this study.

To understand the direction and magnitude of each component's contribution to the data variability, it can be clearly seen in the PCA biplot visualization shown in Figure 2 as follow:

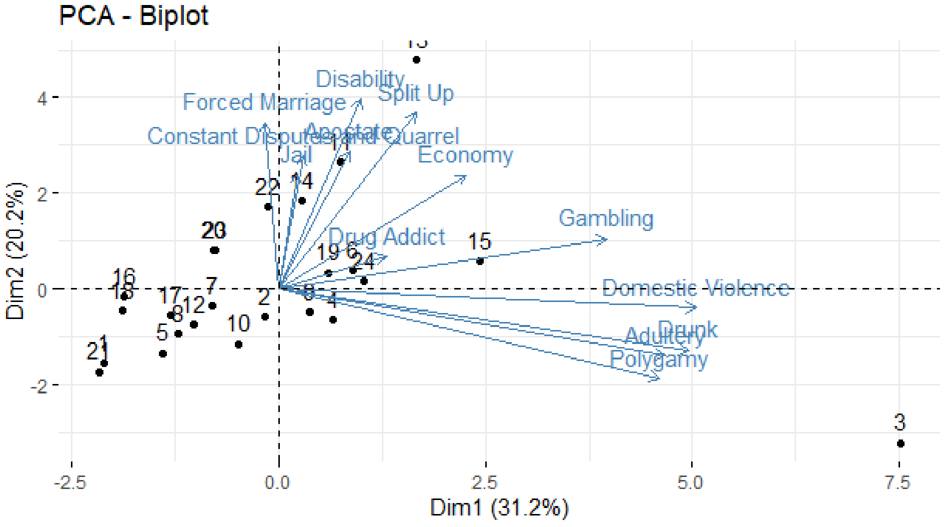

Figure 2. PCA Biplot

The biplot in Figure 2 represents the first two principal components (PC_1 and PC_2) that explain most of the data variation. The axes representing the principal components (PC_1) show how each variable contributes to the formation of the principal components, while the vectors representing these variables provide information about the direction and relationships between the variables in the principal component space.

The categories of each individual to the principal components analyzed through the visualization shown in Figure 2 as follow:

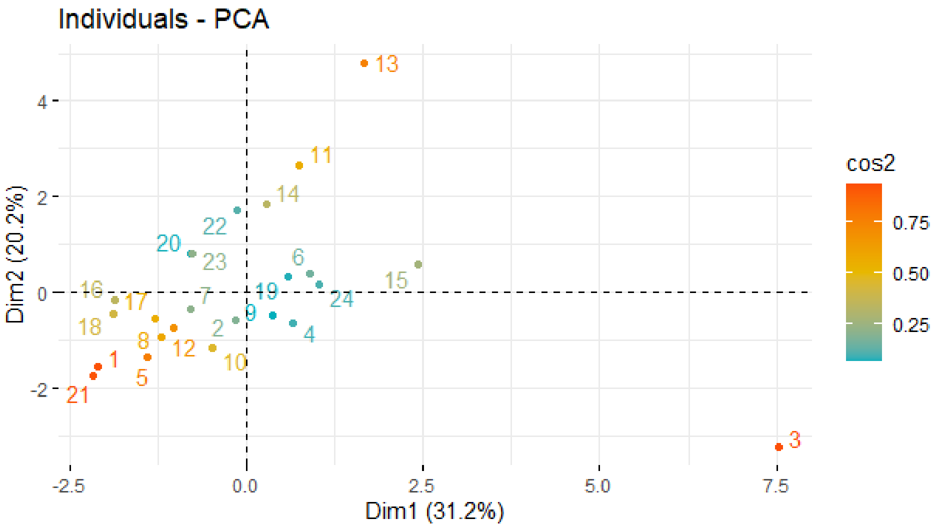

Figure 3. Individuals PCA plot

In the Figure 3, each point represents an individual in the dataset, where the position and distance between points reflect the relative contribution of each individual to the data variability captured in the principal components. The closer two points are, the more similar their contributions are to the same data pattern or structure. Conversely, individuals that are far apart from each other may indicate significant differences in their contributions to the main dimensions of the data. The different colors on the individual PCA plot indicate specific groupings or categories of individuals within the dataset, which may reflect additional characteristics or variables that differentiate each group.

The contributing of each individual to the principal components can be further analyzed through the bar chart shown in Figure 4 below:

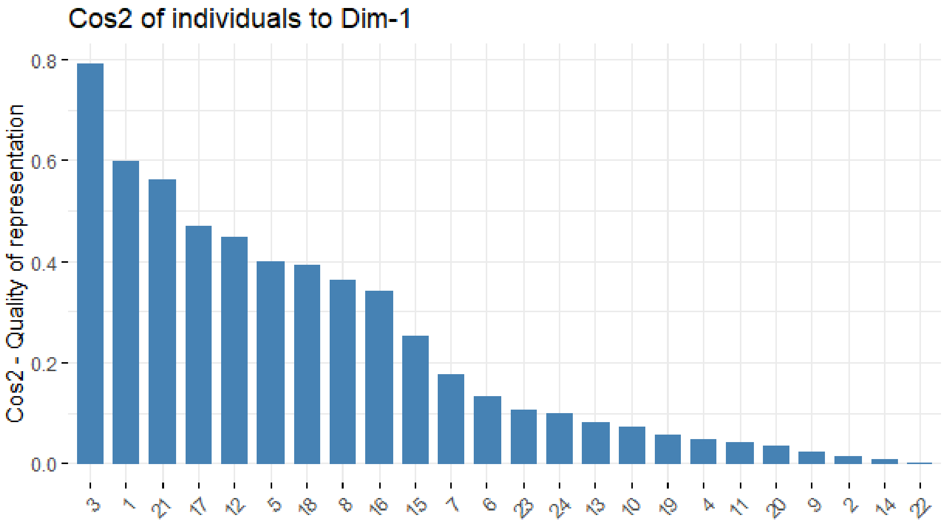

Figure 4. Contribution of individuals in PCA

Figure 4 illustrates the contribution or proportion of variation explained by the first principal component (PC_1) for each individual in the dataset. This visualization provides insights into how well PC_1 captures significant patterns or differences among individuals, which can be observed from the distribution of values reflected in the plot. The larger the proportion of variation explained by PC_1, the more dominant this principal component is in representing the data structure.

The Variables PCA plot in Figure 5 depict the contribution of each variable to the principal components generated from the Principal Component Analysis (PCA). This plot focuses on the relationship between the original variables and the principal components, illustrating the direction and strength of the correlations between the variables and the principal components.

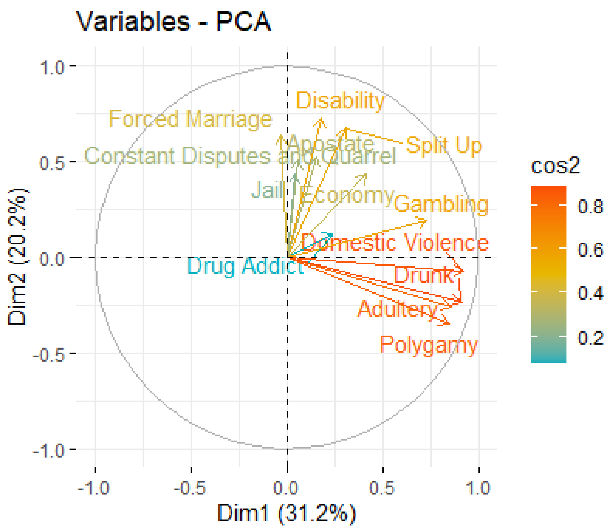

Figure 5. Variables PCA Plot

The arrows on the Variables-PCA plot represent the vectors of the variables and hold important significance. Each arrow indicates the contribution and relationship between the original variables and the principal components examined in the PCA. In Figure 5, arrows pointing to the right (or in the direction of the principal component) suggest a positive correlation with that principal component. Conversely, arrows that are perpendicular or form a 90-degree angle to the principal component indicate no correlation with the principal component being shown.

The color differences in the variable vectors shown in Figure 5 represent specific categories or groups within the data, providing an additional layer of insight into the relationships among variables. The red color indicates that these variables make the largest contribution to the principal component being analyzed, meaning they play a key role in shaping the main patterns in the data. On the other hand, the blue color highlights variables that contribute less, suggesting their influence in explaining the variability of the data is minimal. Furthermore, variables with the same color are strongly correlated, indicating they move in the same direction or have a complementary relationship in explaining the principal component. This strong correlation can offer valuable insights into underlying patterns within the data that might not be immediately apparent when considering individual variables in isolation.

The contributing of each variable to the principal components can be analyzed through the bar chart shown in Figure 6 as follow:

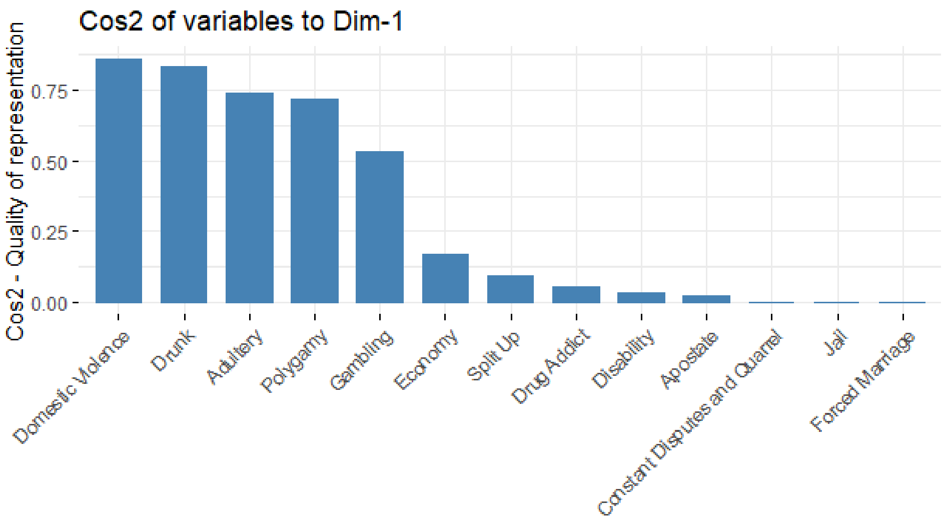

Figure 6. Contribution of Variables in PCA

The cos2 of variables in Figure 6. reflects how well the data points are represented by the first principal component. A higher cos2 value indicates that a data point is well explained by the first component, and a lower value indicates the opposite. The figure shows that the variables domestic violence, drunk, and adultery are the three most influential components on the first principal component.)

3.7 Discussion

Research on the factors contributing to divorce in South Sulawesi can be carried out by linking each factor. Here is a detailed analysis of the five factors identified in the study:

1. **Moral Factor**

The moral factor is identified as one of the main causes of divorce in this study. Behaviors like adultery, drunk, gambling, polygamy, and domestic violence highlight how breaches of social norms and values can weaken the foundation of a relationship. In South Sulawesi, where family and religious values are highly esteemed, these moral violations not only create conflict but can also result in social stigma. The repercussions of these actions often reach beyond the couple, influencing children and the extended family as well.

Adultery, drunk, gambling, and domestic violence are often the result of low morality in an individual. When a person lacks a strong moral foundation, they become more vulnerable to engaging in behaviors that harm both themselves and others. An intriguing finding in this study is that polygamy is strongly linked to moral issues. While polygamy is allowed in Islam, the predominant religion in South Sulawesi, it is still frequently viewed negatively by society. Many people consider polygamy to be in conflict with certain moral principles, despite the religious justification for it. Socially, polygamy is often seen as a sign of gender inequality, unfairness in relationships, and a potential source of family conflict.

2. Personal Will Factor

The personal will factor suggests that the decision to divorce is frequently driven by a strong desire to alter one's life circumstances. This may involve the urge to escape unhealthy relationships or arranged marriages that arise from social pressures. In this regard, it is crucial to examine how evolving societal values, particularly among younger generations, shape their views on marriage and life decisions.

3. Difference Opinion Factor

Differences of opinion are a major factor that can result in ongoing conflict. Incompatibility in perspectives, beliefs, and goals can create lasting tension in a relationship. In South Sulawesi, where many couples may come from diverse cultural and religious backgrounds, it is essential to recognize how this diversity can impact marital dynamics. Education and effective communication between partners are crucial for navigating these differences.

4. Economy Factor

Economic challenges frequently act as a major source of stress for families. Struggles to meet every day needs can result in frustration and tension that affect relationships. In South Sulawesi, where many families depend on unstable income sources, it is vital to find solutions to help couples address their financial difficulties. Providing skills training and access to economic resources can contribute to strengthening family stability.

5. Criminality Factor

The factor of criminality, especially concerning incarceration, shows that illegal activities can have lasting consequences on family structure. The absence of one partner due to legal troubles can lead to uncertainty and emotional distress for the remaining family members. Improved rehabilitation programs for individuals involved in criminal behavior, along with support for their families, can help alleviate the negative effects of this issue.

Overall, this research provides valuable insights into the factors causing divorce in South Sulawesi. By understanding these factors, stakeholders, including the government and community organizations, can formulate more effective policies and intervention programs to support families and prevent divorce. Evidence-based policies and comprehensive analysis are more likely to succeed in strengthening family relationships and reducing divorce rates in this region. Education, emotional support, and access to economic resources should be primary focuses in efforts to improve family well-being in the area.

PCA (Principal Component Analysis) is a highly useful tool in multidimensional data analysis, particularly in social and behavioral research. By helping to reduce data complexity and identify key factors, PCA allows researchers to gain deeper and more relevant insights from existing data. In the context of divorce in South Sulawesi, the application of PCA not only simplifies the analysis but also provides a clearer understanding of the interactions among various causal factors.

4. Conclusion

Based on the research findings, five new components have emerged from the dimensional reduction of the variables causing divorce in South Sulawesi. The five factors influencing divorce consist of moral factors, personal will factors, differences of opinion, economic factors, and criminality factors. By understanding these five factors, this study provides deeper insights into the causes of divorce in South Sulawesi, as well as a foundation for developing more effective policies and interventions to support family well-being and prevent divorce. For upcoming studies, advancements in PCA, such as kernel PCA applied to manage nonlinear data, or robust PCA to tackle outlier problems. Furthermore, the dataset could be expanded by incorporating a wider range of factors of divorce.

5. References

Badan Pusat Statistik Sulawesi Selatan. (2024). *Sulawesi Selatan in Figure 2024*. https://sulsel.bps.go.id/id/publication/2024/02/28/a104de42ebf8eb522608257e/provinsi-sulawesi-selatan-dalam-angka-2024.html

Abdi, H., & Williams, L. J. (2010). Principal Component Analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, *2*(4), 433–459. https://doi.org/10.1002/wics.101

Cossu, G., Agus, M., Atzori, L., Gonzales, C. I. A., Minerba, L., Ferreli, C., Puxeddu, R., Orrù, G., Scano, A., Romano, F., Pintus, E., Penna, M. P., & Carta, M. G. (2022). Principal Component Analysis of the Social and Behavioral Rhythms Scale in elderly. *Journal of Public Health Research*, *11*(1), 137–140. https://doi.org/10.4081/jphr.2021.2546

Emeng, G. I. (2022). Divorce: An Unending Phenomenon in Human Society. *HUMAN: South Asean Journal of Social Studies*, *2*(2), 2022.

Greenacre, M., Groenen, P. J. F., Hastie, T., D'Enza, A. I., Markos, A., & Tuzhilina, E. (2022). Principal component analysis. *Nature Reviews Methods Primers*, *2*(1). https://doi.org/10.1038/s43586-022-00184-w

Jiang, Y. (2022). Study on eigenvalue and eigenvector introduction. *Journal of Physics: Conference Series*, *2282*(1). https://doi.org/10.1088/1742-6596/2282/1/012004

Johnson, R. A., & Wichern, D. W. (2007). Applied Multivariate Statistical Analysis.: Pearson Prentice Hall. In *Pearson Prentice Hall*.

Jolliffe, I. T. (2002). Principal Components Analysis (2nd Edition). Springer, New York.

Jollife, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *374*(2065). https://doi.org/10.1098/rsta.2015.0202

Ling, Z., Gao, Y., & Chen, Q. (2021). Application of principal component analysis in meteorological forecast. *IOP Conference Series: Earth and Environmental Science*, *631*(1). https://doi.org/10.1088/1755-1315/631/1/012019

Liu, L., Li, L., & Mo, H. (2022). Research on Ecological Evaluation Model Based on Principal Component Analysis. *Advances in Transdisciplinary Engineering*, *23*, 429–435. https://doi.org/10.3233/ATDE220312

Ljubicic, M. L., Madsen, A., Juul, A., Almstrup, K., & Johannsen, T. H. (2021). The Application of Principal Component Analysis on Clinical and Biochemical Parameters Exemplified in Children With Congenital Adrenal Hyperplasia. *Frontiers in Endocrinology*, *12*(August), 1–8. https://doi.org/10.3389/fendo.2021.652888

Mohlatlole, N. E., Sithole, S., & Shirindi, M. L. (2018). Factors contributing to divorce among young couples in lebowakgomo. *Social Work (South Africa)*, *54*(2), 256–274. https://doi.org/10.15270/54-2-637

Palazzolo, N., Peres, D. J., Creaco, E., & Cancelliere, A. (2023). Using principal component analysis to incorporate multi-layer soil moisture information in hydrometeorological thresholds for landslide prediction: an investigation based on ERA5-Land reanalysis data. *Natural Hazards and Earth System Sciences*, *23*(1), 279–291. https://doi.org/10.5194/nhess-23-279-2023

Parhusip, H. A., Trihandaru, S., Heriadi, A. H., Santosa, P. P., & Puspasari, M. D. (2022). Data Exploration Using Tableau and Principal Component Analysis. *International Journal on Informatics Visualization*, *6*(4), 911–920. https://doi.org/10.30630/joiv.6.4.952

Varmuza, K. & Filzmoser, P. (2009). *Introduction to Multivariate Statistical Analysis in Chemometrics*. Florida: CRC Press.

Scott, S. B., Rhoades, G. K., Stanley, S. M., Allen, E. S., & Markman, H. J. (2013). Reasons for divorce and recollections of premarital intervention: Implications for improving relationship education. *Couple and Family Psychology: Research and Practice*, *2*(2), 131–145. https://doi.org/10.1037/a0032025

Shlens, J. (2014). *A Tutorial on Principal Component Analysis*. http://arxiv.org/abs/1404.1100

Sofyan, E., & Torro, S. (2022). Cerai Gugat di Kota Parepare. Predestination Journal of Society and Culture, *3*(1), 10-20. https://doi.org/10.26858/prd.v3i1.36193

Zhang, R., Du, T., & Qu, S. (2018). A principal component analysis algorithm based on dimension reduction window. *IEEE Access*, *6*, 63737–63747. https://doi.org/10.1109/ACCESS.2018.2875270