IRAQI
Academic Scientific Journals

University of Kirkuk Journal For Administrative and Economic Science (2024) 14 (1): 223-235

UKJAES

University of Kirkuk Journal For Administrative and Economic Science

ISSN 2222-2995
University of Kirkuk Journal For Administrative and Economic Science

**Musa Othman Rasul, Kamal Muhammed Ali & Wstabdullah Shaho Muhammed. Using Factor Analysis to Identify the Most Influential Reasons Affecting Immigration among the Youth in Sulaimani Province.** ***University of Kirkuk Journal For Administrative and Economic Science*** **(2024) 14 (1):223-235.**

# Using Factor Analysis to Identify the Most Influential Reasons Affecting Immigration among the Youth in Sulaimani Province

Othman Rasul Musa[1], Muhammed Ali Kamal[2], Shaho Muhammed Wstabdullah[3]

[1] *Dukan Technical Institute, Sulaimani, Iraq*
[(2,3)] *University of Sulaimani, Sulaimani, Iraq*

Othman.musa@spu.edu.iq [1]
Muhammed.kamal@univsul.edu.iq [2]
Shho.wstabdullah@univsul.edu.iq [3]

**Abstract:** This research has implemented to define the most significant factors for the immigration of the youth and its influence on each other, as well as the classifying relationship among these variables were described through applying the factor analysis. The data of the research was collected in various places in Sulaymani in 2022. Likert scale of a five-point (1 to 5) is used for rating the statements. The total number of questionnaires distributed was 165, but the response rate was 93.9%. In the findings, it is concluded that there are four factors that cause hopelessness, which they comprised of a number OF variables, namely economic instability, political instability, marginalizing the youth, ideological and political conflict, unhealthy education system, and absence of job opportunities. Thus, it is recommended that government should improve their plans concerning youth's abilities and decreasing their rate of poverty in society. Also, it is recommended that political parties should take actions to prevent more instability.

**Keywords:** Factor Analysis, Immigration, Unique Variance, Common Variance, Principal Factor Method.

# استخدام التحليل العاملي لتحديد الأسباب الأكثر تأثيراً في الهجرة لدى الشباب في محافظة السليمانية

م.م. عثمان رسول موسى [١]، م.م. محمد علي كمال [٢]، م. شاهو محمد وستا عبد الله [٣]

[١] *معهد تقني دكان، السليمانية، العراق*
[(٢،٣)] *جامعة السليمانية، السليمانية، العراق*

**المستخلص:** وتم وصف العلاقة التصنيفية بين هذه المتغيرات من خلال تطبيق التحليل العاملي، وكذلك العلاقة التصنيفية بين هذه المتغيرات. تم جمع بيانات البحث في أماكن مختلفة في السليمانية عام ٢٠٢٢. ويستخدم مقياس

ليكرت ذو الخمس نقاط (١ الى ٥) لتصنيف العبارات. وبلغ إجمالي عدد الاستبيانات الموزعة (١٦٥) استبياناً، لكن نسبة الاستجابة بلغت ٩٣,٩٪. وتوصلت النتائج إلى أن هناك أربعة عوامل تسبب اليأس، والتي تتكون من عدد من المتغيرات، وهي عدم الاستقرار الاقتصادي، وعدم الاستقرار السياسي، وتهميش الشباب، والصراع الأيديولوجي والسياسي، ونظام التعليم غير الصحي، وغياب فرص العمل ... وبالتالي، يوصى بأن تعمل الحكومة على تحسين خططها المتعلقة بقدرات الشباب وتقليل نسبة الفقر في المجتمع. كما يوصى بأن تتخذ الأحزاب السياسية إجراءات لمنع المزيد من عدم الاستقرار.

**الكلمات المفتاحية:** التحليل العاملي، الهجرة، التباين الفريد، التباين المشترك، طريقة العامل الرئيسي.

**Corresponding Author: E-mail: Othman.musa@spu.edu.iq**

## 1 Introduction

Factor analysis is a statistical approach for demonstrating a huge number of variables that are related together so that a potentially small number of unnoticed variables can be detected. In another expression, using a matrix with a low-rank through an explicit probabilistic linear model, the method approximates a matrix-shaped data set. Factor analysis is a method that uncovers the underlining structure and decreases the complexity and of the data set. The age of this method is more than a hundred years. The factor model in psychology goes back at least to Spearman (1904), who from time to time is credited for inventing factor analysis. Later on, not only in psychology, but the technique was also applied to bioinformatics, signal processing, marketing and finance, economics, social science and etc. The hidden factors, which are revealed by factor analysis, lead us to understand the observed variables more. Usually, factor analysis can be grouped into two types; confirmatory factor analysis, in this type, constraints regarding the factor loadings are pre-determined. While explanatory factor analysis is the other type, where constraints are not applied such as those of type one. Lately and specifically for analysis of high-dimensional data and large matrices, factor analysis has become a tool that is mostly used in dimension reduction. Factor analysis has many similar sides to the approximation of low-rank matrix, which has applications in fields such as personalized learning, collaborative filtering, and signal processing. If we compare factor analysis with the singular value decomposition (SVD) or principal component analysis (PCA), factor analysis uses heteroscedastic noise assumption for each variable, which in contrast with constant noise variance is more rational in many applications. Unexpectedly, it is not easy to choose the number of factors which are problems we face with factor analysis. There is not existence of best-performing methods that are accepted by most, regarding the problems of traditional factor analysis that contain a small number of variables compared to the big sample they have (for example, Peres-Neto et al. (2005).

### A- Literature review

When a researcher is interested in investigating which variables in a group form reasonable subgroup, that are relatively independent from each other, a statistical technique involving multivariate is applied to the group of variables called factor analysis (Peres-Neto et al (2005)). Another way to say it is, factor analysis is used to identify latent factors that form basis of the variables through clustering variables that are related in the same factor (Spearman, 1904). In this study, factor analysis is applied mostly for the purpose of reduction of a large number of inter-correlated measure to a few representing factors that are used in subsequent analysis (Tabachnick and Fidell, 2013). The goal of the current research is to test the factor analysis application by distributing questionnaires to inquire the level of satisfaction among the tourists. At first, we make an assumption that there exists a correlation to a degree between all variables. Variables should be measured at least at an ordinal level. The larger size of the sample in factor analysis, the better but the ranges that are more accepted are ten-to-one ratios (Tabachnick and Fidell, 2013; Verma and Abdel-Salam, 2019). Factor analysis can be classified into two categories, namely confirmatory factor analysis (CFA) and exploratory factor analysis (EFA). Usually in the first stages of a research, exploratory factor analysis is utilized to collect data concerning the interrelationships among a group of variables, it can be useful in checking dimensionality as well (Ho, 2006). While confirmatory factor analysis which is a set of more sophisticated and complex techniques are useful in the research process to examine particular hypotheses about the formation of an underlying group of variables (Hair et al (1998); Pituch and Stevens (2016)). A number of researches investigated and put the application of factor analysis under discussion in order to reduce a set of large data and determine extracted factors from the analysis (Hair et al (2006); Pallant (2010); Cerny and Kaiser (1977); Dziuban and Shirkey (1974)). In the tourism sector, measurement of tourist satisfaction can be done by many parameters. These variables can be gathered together into different groups of factors in factor analysis, where any of the factors deal with measuring the tourist satisfaction from at least one dimension. The factors are organized in a way that the variables in it are connected together in one way or another.

(1) To demonstrate the uppermost variability of a group in the study, extraction of the significant factors is done. One of the most prominent benefits of applying factor analysis is the provision of precious inputs so that the policy and decision makers concentrate on a small number of factors instead of huge numbers of factors. People working in the tourism sector make lots of effort to have information about the destination choice of their tourist or customer. Probably there are endless worries about taking the opinions of the tourist about different topics.

(2) Matters such as transportation, safety, cost of accommodation, communication, local food, culture, weather condition, travel video making, photography, trekking, medical treatment, nature, water supply, recreation activities, natural resources, environment, mountaineering and others are discovered when responses from the surveys among the tourists are collected in addition to the literature review (Bartlett, 1951). Using factor analysis, we can cluster several variables in unalike components. The policy maker or the researcher can optimize the components rather than focusing on many issues, that is to contribute to the growth of the tourism sector. This paper contributes in emphasizing the pros of factor analysis. Firstly, it is used to make developments in a questionnaire. While analyzing, irrational questions can be taken out from the final version of the questionnaire. In organizing the questionnaire factor analysis assists in the classification of questions various parameters. Secondly, it is used in data simplification. This study motivates researchers in the identification of factors using factor analysis which is a step-by-step process.

(3) At times, having a clear vision about the variables in the questionnaire is impossible due to adding more items or statements. Factor analysis helps in choosing few factors among a huge number of variables which are related together, this way the number of factors can be managed and dealt with better, before analyzing them in topics such as multivariate analysis or multiple regression of variance (MacCallum et al, (1999); Dhakal (2017)) Consequently, a small number of extracted factors can be investigated rather than testing all the parameters, which assists in demonstrating the group characteristic variations.

**B- Objectives**

The goal of the current research is to point out the most important factors that affect youth migration, using factor analysis. In order to help decision-makers in the government to find an appropriate explanation for migration and its reasons through obtaining statistical indicators. Consequently, coming up with conclusions and recommendations that assist them to take suitable decisions to set boundaries for this phenomenon.

## 2 Research Methodology

The questionnaire is prepared in a way to collect primary data. Various places in Sulaymani are taken to collect data in 2022. The research contains answers from the participants older than 18 years old. The questionnaire contains of statements and questions connected to the variables, which were prepared considering the literature review. The Likert scale of a five-point (1 to 5) is used for rating the statements which strongly agree is indicated by 5. The total number of distributed questionnaires was 165, but the responded rate was 93.9% which means 155 participants answered the questions and stated their reactions to the statements. IBM SPSS version 23 is used to apply all the statistical analysis.

### A- Cronbach's Alpha

Cronbach's Alpha is used to evaluate the questionnaire's reliability. It is an easy way to identify if a score is reliable or not. The assumption under which Cronbach's alpha is implemented is that there are several items that are used for the measurement of the same underlying construct. An example of this is that, there are questions covering different things in the youth migration satisfaction survey, but it is to measure the overall satisfaction when clustered together.

It is used to make measurements of internal consistency. As said above, it is known to measure scale reliability. Cronbach's alpha is as follows:

$$\alpha = \frac{nr^{-}}{1 + r^{-}(n-1)}$$

Where, $\boldsymbol{r}^{-}$ is the mean correlation of mean among the items and n indicates the items' number. The range of Cronbach's alpha is between (0 and 1). If the value of Cronbach's alpha is usually bigger than (0.7), it is acceptable. A high correlation among test items can be noted when there is a high level of alpha (Snedecor and Cochran, 1989).

### B- Factor Analysis Model

The model of factor analysis is based on a linear function, illustrates (p) the number of variables to a sample size of (n), with the composition of (p) unique factors where each variable (m<p), (MacCallum, R.C., Widaman, K. F, 1999). The model is structured as follows:

$$X_{p\times 1} = \mu_{p\times 1} + A_{p\times m}F_{m\times 1} + U_{p\times 1}$$

Where;

X: is the vector of the random variables
μ: is the vector of the mean variable
A: is the common factor matrix of the variables
F: is the random common factor
U: is the vector of the unique factor

From the above formula which describes variable (j) and single value (i) denoted by the total of weighted value in factor, the model is as follows:

$$\mathrm{Z_{ij} = a_{i1}F_{1j} + a_{i2}F_{2j} + \cdots \dots \dots + a_{jm}F_{mi} + U_{ji}}$$

Where;

$\mathbf{Z_{ij}}$: Represents single or variables normative value
$\mathbf{F_{mi}}$ ...$\mathbf{F_{1j}}$: Represent single common factor normative value
$\mathbf{U_{ji}}$: Represents normative value of the unique factor of variables (j) and single (i)
$\mathbf{a_{i1}}$ ........ $\mathbf{a_{jm}}$: Represents the transfer the value of weighted in common factor

### (1) Basic Assumption of Factor Analysis

In factor analysis, covariance matrix or correlation matrix is being used. The analysis depends on whether there exists the correlation between the variables under study or not. That is because factor analysis tries to demonstrate this correlation through a number of factors. The response variable in factor analysis has a mean of (0) and a variance of (1), which is a standard variable. The units of measurement have different variables, (Dhakal, B, 2017). Therefore, the model of factor analysis becomes as follows:

X = AF+U

Common factor (F) is independent among the factors,

E (F) = mx1*cov (F) = E (FF') = φ (mxm)

The unique factor (U) is independent among the factors,

E (U) = (pxp)

Covariance matrix

$$Cov (U) = E (UU') = \Psi pxp = \begin{pmatrix} \Psi_1 & 0 & \dots & 0 \\ 0 & \Psi_2 & \dots & 0 \\ . & 0 & \dots & 0 \\ 0 & 0 & \dots & \Psi_p \end{pmatrix}$$

Each of the unique and common factors is independent,

$$E (F' U') = \begin{bmatrix} E(FF') & E(FU') \\ E(FU') & E(UU') \end{bmatrix} = \begin{bmatrix} \emptyset_{mxm} & 0 \\ 0 & \Psi_{pxp} \end{bmatrix}$$

Where;

ϕ : is the covariance matrix in common factor (F)

Ψ : is the factor unit

The covariance matrix for (X) variable is derived from the main form of factor analysis, as follows:

X =μ + AF +U

(X -μ)(X -μ)' = (AF + U) (AF +U)'= (AF +U) AF' +U'=AF (AF)' + U (AF)' +AFU' + UU'

Σ = Cov (X) = E(X -μ) = (X -μ)'= AE (FF') A' +E (UF') A' + AE (FU) + E (UU')

E (FF') A' = E (UF') = E (FU') = E (UU') =Ψ

Σ = AA +Ψ '

**(2) Factor Analysis Variance**

Con the sums of squared logs for any factor are quantities at the bottom of each column; which demonstrate the amount of the total considered variance by that factor among the observed variables.

Var (Xi) = 1 = a2i1 Var (f1) +a2i2 (f2) + ...+a2im Var (fm) + Var (ei) =a2i1 + a2i2 + ... + a2im Var (ei)

The quantity a2i1 + a2i2 + ... + a2im is known as a communality of xi, which is part of the variance, is connected to the common factors. Where Var (ei) is known as the specificity of xi that is not connected to the common factors, (Spearman, C, 1904).

The correlation between two values xi and xj is given by:

rij= ai1aj1 + ai2aj2 + ... + aimajm .......... (2.5)

As a result, if two test scores have high logs on the same factors, they can only be highly correlated. Also, -1 ≤ aij ≤ 1 for every i and j because the communality cannot be more than (1).

**(3) Communalities**

To gather the least transferring factors, a number of common variables exist. The variables are part of the total variance, which is used in determination of the variables overlaps among the common factors and the variables.

Hi = $\sum_{k=1}^{m}$ aki .......... (2.6)

Where,

a ki: is the factor matrix (F) weighting factor (k) and variables(i)

J=1, 2, 3 ...p: number of the variables.

K=1,2, 3....p: number of the factors.

We should estimate amount of first common (hj) when using the main factor analyzing the method as one of the methods of preserving factor analysis. Communality values are between (0) and (1) because they are a part of the total variance, (Tucker, L. R., MacCallum, 2021).

0≤ hj ≤ 1

**(4) The Rotated Factor Matrix**

There are several ways to obtain the rounded factors' matrix. We have used Varimax method which is the most common orthogonal rotation method, it was proposed in 1958 by Kaiser. This method relies on the simplification of factors composition through the loading square variances (Dziuban, C.D. and Shirkey, E.C.,١٩٧٤ ).

$$S_p^2 = \frac{1}{n}\sum_{j=1}^{n}(a_{jp}^2)^2 - \frac{1}{n^2}\sum_{j=1}^{n}(a_{jp}^2)^2$$

Where ajp is the row element and (p) in the inverse matrix and (j) of the column.

**(5) Principal Factor Method**

The format matrix is written as follows:

$$\begin{bmatrix} X_1 \\ . \\ . \\ . \\ X_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{1m} \\ . & . & . \\ . & . & . \\ . & . & . \\ a_{p1} & a_{p2} & a_{pm} \end{bmatrix} \begin{bmatrix} F_1 \\ . \\ . \\ . \\ F_m \end{bmatrix} + \begin{bmatrix} U_1 \\ . \\ . \\ . \\ U_m \end{bmatrix}$$

The following steps are followed to estimate the factor matrix:

Correlation Coefficient Matrix

The standard values of the variables have the same measurement units of using variance matrix. (Variance Covariance Matrix), Peres-Neto P. R. Jackson, D. A and Somers, K. M, 2005).

$\text{Rij}=\frac{\sum_{k=1}^{n}\text{Xijk}}{n-1}$.......... (2.7)

Where:

$$R=\begin{bmatrix} 1 & r_{12} & r_{1p} \\ . & . & . \\ r_{p1} & r_{p2} & 1 \end{bmatrix}$$

The square of the correlation coefficient can be calculated for each variable with the rest of the variables (R2rest). Firstly, the common factor is estimated to replace the single digits to be symbolized (Rr) and called Reduced Correlation Matrix. From the sample values (Eigen Values), the extraction of the correlation matrix is done as follows:

Rr - λ1a = 0…......... (2.8)

**(6) Kaiser-Meyer-Olkin (KMO)**

The suitability of the data for factor analysis is measured using KMO. KMO test examines the adequacy of the sample size.

For the complete model and for each variable, the KMO test measures the adequacy of the sampling process. The formula of the KMO test is as follows:

$$KMO_j = \frac{\sum_{i\neq j} R_{ij}^2}{\sum_{i\neq j} R_{ij}^2 + \sum_{i\neq j} U_{ij}^2}$$

Where, Uij is the matrix of partial covariance and Rij is the matrix of correlation. The value of KMO changes between (0) to (1). The values of KMO between 0.6 to 0.69 are mediocre, while the values between 0.7 to 0.79 are middling and values between 0.8 to 1.0 mean that the sampling process is adequate. On the other hand, values less than 0.6 mean the sampling is not adequate. The results of the factor analysis without any doubts are not quite suitable for the analysis of the data, if the KMO value is smaller than 0.5. When we have a sample, its size is less than 300; we need to test the average communality from the retained items. For a sample of size less than 100 and average value of 0.6 is acceptable, while for the samples of sizes more than 100 and less than 200, an average between 0.5 and 0.6 is accepted (Lavrakas, 2008; Guttman, 1954).

**(7) Bartlett's Test of Sphericity**

The null hypothesis (H0), which states that the variables are orthogonal, is tested using Bartlett's Test of Sphericity. The original matrix of correlation indicates that the variables are unsuitable for structure detection as they are not related together. While the alternative hypothesis (H1) states that the variables are not orthogonal. That means the variables are correlated enough to a point the matrix of correlation diverges from the identity matrix in a significant way. A factor analysis could possibly be worthwhile for the data set when the significant value is less than 0.05. The determinant of the correlation matrix |R| is calculated to measure the relation between the variables in total. Under (H0), |R| equals to (1) but if the variables are correlated strongly, then |R| ≈ 0.

The Bartlett's test of Sphericity is given by:

$$x^2 = -\left(n - 1 - \frac{2p+5}{6}\right) * ln\,|R|$$

In the above formula n is the total size of the sample, p indicates the number of variables and R is the matrix of the correlation (Lavrakas, 2008; Guttman, 1954).

## 3 Conclusions and Recommendations

From the research we came up with the following conclusions:

1. There are four factors which the Total Initial Eigenvalues are more than one, these four factors explain (53.046%) of the total variance.
2. the most significant among all the four factors is the first factor as it explains (30.803%). This factor covers the variables ($X1$: economic instability, $X2$: political instability, $X3$: marginalizing the youth, $X5$: ideological and political conflict, $X9$: unhealthy educational system and $X12$: absence of job opportunities) with the values (0.731, 0.673, 0.761, 0.513, 0.641 and 0.536) respectively.
3. The second most significant among all the four factors is the second factor as it explains (10.002%). This factor covers the variables ($X4$: existence of an attractive factor abroad, $X6$: absence of absolute freedom, $X7$: absence of a healthy domestic education, $X8$: continuous external threats, $X11$: limitations on freedom of expression, $X15$: culture and traditions and $X17$: broken families) with the values (0.496, 0.608, 0.609, 0.728, 0.448, 0.695 and 0.537) respectively.
4. The third most significant among all the four factors is the third factor as it explains (6.231%). This factor covers the variables ($X16$: unemployment, $X18$: religious factors, $X19$: absence of financial supports for marriage and $X20$: desperate emotional and social relationships) with the values (0.530, 0.698, 0.616 and 0.513) respectively.
5. The least most significant among all the four factors is the fourth factor as it explains (5.010%). This factor covers the variables ($X10$: employment opportunities abroad, $X13$: effect of media and social media and $X14$: unequal social status) with the values (0.557, 0.589 and 0.770) respectively.

### A- Tables and figures

Bartlett's Test of Sphericity

The null hypothesis (H0), which states that the variables are orthogonal, is tested using Bartlett's Test of Sphericity. The original matrix of correlation indicates that the variables are unsuitable for structure detection as they are not related together. While the alternative hypothesis (H1) states that the variables are not orthogonal.

factor analysis could possibly be worthwhile for the data set when the significant value is less than 0.05. The determinant of the correlation matrix |R| is calculated to measure the relation between the variables in total. Under (H0), |R| equals to (1) but if the variables are correlated strongly, then |R| ≈ 0.

The Bartlett's test of Sphericity is given by:

$$x^2 = -\left(n - 1 - \frac{2p+5}{6}\right) * ln\,|R|$$

In the above formula n is the total size of the sample, p indicates the number of variables and R is the matrix of the correlation (Lavrakas, 2008; Guttman, 1954).

## B- Descriptive statistics

### Gender

**Table (1):** Gender

| G+Ender##$+ | Frequency | Percent |
|---|---|---|
| Male | 79 | 51.0 |
| Female | 76 | 49.0 |
| Total | 155 | 100.0% |

51.0 % of the participation of this research is male and 49.0% are female.

### Age

**Table (2):** Age Distribution

| Age distribution | Frequency | Percent |
|---|---|---|
| Less than 20 | 11 | 7.1 |
| 20 – 30 | 135 | 87.1 |
| 30 – 40 | 6 | 3.9 |
| 40 and more | 3 | 1.9 |
| Total | 155 | 100.0% |

Table (2) shows that 87.1% of participants are (20 to 30) years old, 7.1% of the participants are below (20) years old, 3.9 % of the participants are (30 to 40) years old and 1.9% of the 155 respondents are 40 years and above.

### (1) Education Level

**Table (3):** Education level

| Residents | Frequency | Percent |
|---|---|---|
| Primary | 4 | 2.6 |
| Secondary | 16 | 10.3 |
| Diploma | 31 | 20.0 |
| Bachelor | 101 | 65.2 |
| High school | 3 | 1.9 |
| Total | 155 | 100.0% |

Table (3) demonstrates that out of the total participants: 2.6% has a Primary degree; 10.3% has a Secondary degree, 20.0% has a Diploma degree, 65.2% has Bachelor and 1.9% has a High school degree.

**Table (4):** Economic status

| Residents | Frequency | Percent |
|---|---|---|
| Good | 52 | 33.5 |
| Not bad | 90 | 58.1 |
| Bad | 13 | 8.4 |
| Total | 155 | 100.0% |

Table (4) demonstrates that out of the total participants: 33.5% has good economic status, 58.1% has not bad economic status and 8.4% has bad economic status.

### (2) The Result of factors discussion

### (A) Kaiser-Meyer-Olkin and Bartlett's Test

The statistic of KMO changes between 0 and 1; which 0 means that the sum of partial correlations is big in relation to the sum of correlations, showing diffusion of the correlations' pattern. On the other hand, 1 means that correlation patterns are relatively dense and so factor analysis should yield dependable and clear factors. The recommendation of Kaiser (1974) considers accepting values bigger than (0.5) as acceptance. Moreover, the values of KMO between (0.5 and 0.7) are mediocre, (0.7 and 0.8) are good, (0.8 and 0.9) are great, and values of KMO above (0.9) are superb.

In our research the values of KMO equals (0.852), which falls into the fourth range that is the KMO value is great. While, as shown below in table (5) Bartlett's test is highly significant ($p < 0.05$). As a result, the factor analysis is appropriate to be applied throughout this research.

**Table (5):** KMO and Bartlett's Test

| KMO | | .852 |
|---|---|---|
| Bartlett's Test | Approx. Chi-Square | 993.458 |
| | Df. | 190 |
| | Sig. | .000 |

### (B) The Initial Factor Analysis Solution

**Table (6):** Communalities

| Questions | Initial | Extraction |
|---|---|---|
| Q1 | 1.000 | .620 |
| Q2 | 1.000 | .484 |
| Q3 | 1.000 | .640 |
| Q4 | 1.000 | .439 |
| Q5 | 1.000 | .499 |
| Q6 | 1.000 | .474 |
| Q7 | 1.000 | .435 |
| Q8 | 1.000 | .576 |
| Q9 | 1.000 | .483 |
| Q10 | 1.000 | .518 |
| Q11 | 1.000 | .492 |
| Q12 | 1.000 | .673 |
| Q13 | 1.000 | .501 |
| Q14 | 1.000 | .646 |
| Q15 | 1.000 | .590 |
| Q16 | 1.000 | .512 |
| Q17 | 1.000 | .500 |
| Q18 | 1.000 | .561 |
| Q19 | 1.000 | .514 |
| Q20 | 1.000 | .450 |

Table (6) is the table of communalities before and after extraction. For instance, 62.0% of the variance associated with question one is common.

## (C) Factor Extraction

**Table (7):** Total Variance Explained

| Component | Initial Eigenvalues | | | Extraction Sums of Squared Loadings | | | Rotation Sums of Squared Loadings | | |
|---|---|---|---|---|---|---|---|---|---|
| | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % |
| 1 | 6.161 | 30.803 | 30.803 | 6.161 | 30.803 | 30.803 | 3.418 | 17.088 | 17.088 |
| 2 | 2.000 | 10.002 | 40.805 | 2.000 | 10.002 | 40.805 | 3.011 | 15.056 | 32.144 |
| 3 | 1.246 | 6.231 | 47.036 | 1.246 | 6.231 | 47.036 | 2.095 | 10.474 | 42.618 |
| 4 | 1.202 | 6.010 | 53.046 | 1.202 | 6.010 | 53.046 | 2.086 | 10.428 | 53.046 |
| 5 | .987 | 4.933 | 57.979 | | | | | | |
| 6 | .911 | 4.553 | 62.532 | | | | | | |
| 7 | .826 | 4.129 | 66.662 | | | | | | |
| 8 | .783 | 3.913 | 70.574 | | | | | | |
| 9 | .746 | 3.730 | 74.304 | | | | | | |
| 10 | .666 | 3.329 | 77.634 | | | | | | |
| 11 | .652 | 3.262 | 80.896 | | | | | | |
| 12 | .609 | 3.047 | 83.943 | | | | | | |
| 13 | .556 | 2.782 | 86.725 | | | | | | |
| 14 | .501 | 2.507 | 89.232 | | | | | | |
| 15 | .457 | 2.284 | 91.516 | | | | | | |
| 16 | .421 | 2.106 | 93.622 | | | | | | |
| 17 | .374 | 1.871 | 95.492 | | | | | | |
| 18 | .344 | 1.720 | 97.212 | | | | | | |
| 19 | .289 | 1.446 | 98.658 | | | | | | |
| 20 | .268 | 1.342 | 100.000 | | | | | | |

Table (7) demonstrates the eigenvalues associated with each linear component (before and after) extraction and after rotation. Before extraction, 20 linear components have been identified.

The variance, which is explained by a particular linear component, can be represented by the initial eigenvalues. While eigenvalues bigger than 1 that are extracted are going to be the main factors. As a result, we came to know that in this research we have four main factors. Among which, factor 1 is the most significant factor which explains 30.803% of total variance. Next, factor 2 explains rate of variance of 10.002%. Then factor 3 comes with an explanation rate of variance of 6.231%. Lastly factor 4 comes with an explanation rate of variance of 6.010%. These four factors have an explanation rate of variance of 53.046% in total.

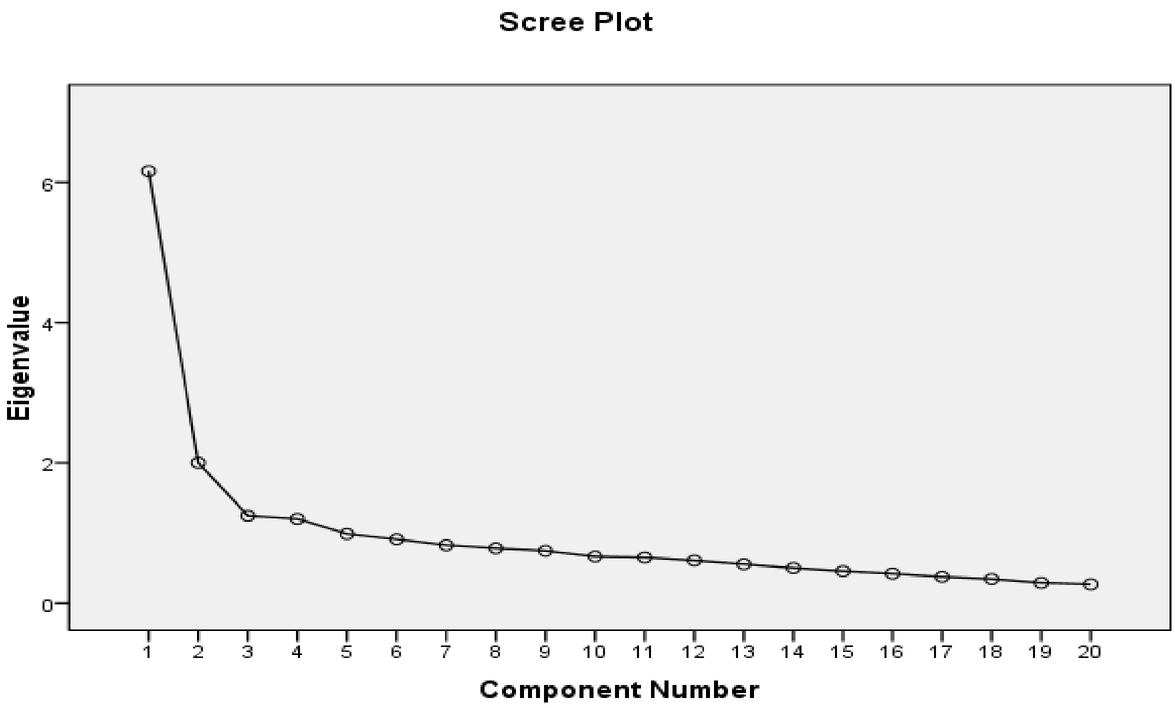


**Figure (1):** Factor Extraction of Initial Eigenvalues

## (D) Rotated Component Solution

**Table (8):** Rotated Component Matrix

| | Component | | | |
|---|---|---|---|---|
| Questions | 1 | 2 | 3 | 4 |
| Q1 | .731 | .029 | .044 | .287 |
| Q2 | .673 | .155 | .052 | .067 |
| Q3 | .761 | .071 | .220 | .084 |
| Q4 | .316 | .496 | .096 | .291 |
| Q5 | .513 | .401 | .227 | -.152 |
| Q6 | .317 | .608 | .053 | .046 |
| Q7 | .166 | .609 | .047 | .183 |
| Q8 | .143 | .728 | .158 | .009 |
| Q9 | .641 | .165 | .169 | .130 |
| Q10 | .380 | .178 | -.178 | .557 |
| Q11 | .419 | .448 | .146 | .307 |
| Q12 | .536 | -.074 | .354 | .505 |
| Q13 | .164 | .196 | .299 | .589 |
| Q14 | .027 | .183 | .138 | .770 |
| Q15 | -.163 | .695 | .106 | .264 |
| Q16 | .376 | .063 | .530 | .293 |
| Q17 | .012 | .537 | .460 | -.009 |
| Q18 | .235 | .130 | .698 | -.033 |
| Q19 | .237 | .112 | .616 | .257 |
| Q20 | -.117 | .403 | .513 | .103 |

Table (8) demonstrates the rotated component matrix, which is a factor loading the matrix for each variable on the four main factors extracted. Values of the factor loading less than 0.4 are suppressed. Therefore, factor loadings only greater than 0.4 have been displayed.

**Factor 1 includes:**

Q3, Q1, Q2, Q9, Q12 and Q5, factor 2 includes: Q8, Q15, Q7, Q6, Q17, Q4 and Q11, factor 3 includes: Q18, Q19, Q16 and Q20 while factor 4 includes: Q14, Q13 and Q10 regarding the significance of the questions in each factor from biggest to smallest.

**Table (9):** Component Transformation Matrix

| Component | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | .615 | .526 | .420 | .411 |
| 2 | -.648 | .739 | .141 | -.121 |
| 3 | -.375 | -.153 | -.142 | .903 |
| 4 | -.249 | -.392 | .885 | -.031 |

Table (9) demonstrates the coefficients of the factor scores' matrix; the observed variables are in standardized form, which means they have a mean equaling to (0), and a standard deviation equaling to (1).

## (3) Equations

1st Factor contains the variables (X1, X2, X3, X5, X9, X12) by the rate of 30.803% of total variance.

$$F1 = 0.731X1+0.673X2 + 0.761X3 +0.513X5+ 0.641X9 + 0.536X12$$

2nd Factor contains the variables (X4, X6, X7, X8, X11, X15, X17) by the rate of 10.002% of total variance.

$$F2 = 0.496X4 +0.608X6+0.609X7 + 0.728X8 + 0.448 X11 + 0.695 X15+ 0.537X17$$

3rd Factor contains the variables (X16, X18, X19, and X20) by the rate of 6.231% of total variance.

$$F3 =0.530X16 + 0.698X18 + 0.616X19 +0.513X20$$

4th Factor contains the variables (X10, X13, X14) by the rate of 6.010% of total variance.

$$F4 =0.557X10 + 0.589X13 + 0.770X14$$

## 4 Conclusions

As from the conclusions we have seen, marginalizing the youth is one of the most significant factors behind hopelessness that is why we recommend the government to take actions necessary so the youth's abilities should be considered.

As one other significant factor is economic instability, it is recommended that the government should make plans for a better economy and decrease the rate of poverty in society.

Another significant factor is political instability, political parties are recommended to take actions to prevent more instability.

## Appreciation

First of all, I am grateful to the Almighty God for establishing me to complete this research. My mother, I am hugely indebted to my wonderful wife (Sazgar Mohammed Sulaiman), for the help and support, and also for having the faith on me and urging me to do better, and for the hardships she went through during the period of my study, is much appreciated. Finally, but not least, I want to thank my family members, my friends and teacher both in Sulaimani Polytechnic University and Sulaimani University.

## References

1- Bartlett, M. S., “The effect of standardization on a Chi-square approximation in factor analysis,” Biometrika, (1951), 38(3/4), 337-344.. https://academic.oup.com/biomet/article-abstract/38/3-4/337/257904
2- Cerny, C.A. and Kaiser, H.F, “A study of a measure of sampling adequacy for factor analytic correlation matrices,” Multivariate Behavioral Research, (1977), 12 (1). 43-47... https://pubmed.ncbi.nlm.nih.gov/26804143/
3- Dhakal, B., “Using factor analysis for residents’ attitudes towards economic impact of tourism in Nepal,” International Journal of Statistics and Applications, (2017), 7(5), 250-257. http://dx.doi.org/10.5923/j.statistics.20170705.02
4- Dziuban, C.D. and Shirkey, E.C., “When is a correlation matrix appropriate for factor analysis?” Psychological Bulletin, (1974), 81,358-361.https://psycnet.apa.org/doi/10.1037/h0036316
5- Guttman, L., “Some necessary conditions for common-factor analysis,” Psychometrical, 19, 149-161. 1954. https://link.springer.com/article/10.1007/BF02289162
6- Hair, J. J., Black, W.C., Babin, B. J., Anderson, R. R., Tatham, R. L., Multivariate data analysis, Upper Saddle River, NewJersey,(2013),15-734, Retrieved from: https://www.google.iq/books/edition/Multivariate_Data_Analysis/VvXZnQEACAAJ?hl=en
7- Ho, R., Handbook of univariate and multivariate data analysis and interpretation with SPSS, Chapman & Hall/CRC, Boca Raton, (2006). https://doi.org/10.1201/9781420011111
8- Lavrakas, P.J., Encyclopedia of survey research methods. SAGE Publications, Thousand Oaks, (2008), 524-527. http://dx.doi.org/10.4135/9781412963947.n337
9- MacCallum, R.C., Widaman, K. F., Zhang, S. and Hong, S. “Sample size in factor analysis,” Psychological Methods, (1999), 4(1), 84-99. https://psycnet.apa.org/doi/10.1037/1082-989X.4.1.84
10- Pallant, J., SPSS survival manual: a step-by-step guide to data analysis using SPSS, Open University Press/ Mc Graw-Hill, Maidenhead, (2010),104-378.https://doi.org/10.4324/9781003117452
11- Peres-Neto P. R. Jackson, D. A and Somers, K. M. How many principal components? Stopping rules for determining the number of non-trivial axes revisited. Computational Statistics & Data Analysis, (2005), 49(4):974–997, https://doi.org/10.1016/j.csda.2004.06.015
12- Pituch, K. A. and Stevens, J., Applied multivariate statistics for the social sciences: Analyses with SAS and IBM’s SPSS (6thed.), Taylor & Francis, New York, (201٥). 435-814 https://doi.org/10.4324/9781315814919
13- Snedecor, G. W. and Cochran, W.G., Statistical Methods (8th ed.), Iowa State University Press, Iowa, 1989. https://doi.org/10.1017/S0021859600074104
14- Spearman, C. “general intelligence,” objectively determined and measured. The American Journal of Psychology, (1904), 15(2):201–292, 1904. https://doi.org/10.2307/1412107
15- Tabachnick, B.G. and Fidell, L.S., Using multivariate statistics (6th ed.), Pearson, 2013. Retrieved from: https://books.google.iq/books/about/Multivariate_Data_Analysis.html?id=VvXZnQEACAAJ&redir_esc=y7
16- Tucker, L. R., MacCallum, R.C., Exploratory factor analysis, [E-book], available: net Library e-book. 2021, Vol. 9 No. 1, 4-11, DOI: 10.12691/ajams-9-1-2. https://www.sciepub.com/reference/349063
17- Verma,J. and Abdel-Salam, aTesting statistical assumptions in research, John Willey & Sonsinc, (2019). Retrieved from: https://www.researchgate.net/publication/331461660_Testing_Statistical_Assumptions_in_Research