COMPARISON OF FACTOR EXTRACTION AND ROTATION METHODS IN EXPLORATORY FACTOR ANALYSIS

S.M.P.Siriwardhana

(08/10308)

Dissertation submitted in partial fulfilment of the requirements for the degree Master of Science in Operational Research

Department of Mathematics

University of Moratuwa

Sri Lanka

April 2012

Declaration of the Candidate

"I declare that this is my own work and this dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any University or other institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another except where the acknowledgment is made in the text"

Signature: *UOM Verified Signature* Date : 21.04.'12

Declaration of the Supervisor

"I have supervised and accepted this thesis/dissertation for the award of the degree"

Signature: Date:

To my Parents and Husband

ACKNOWLEDGEMENT

Writing a thesis is a long journey which one cannot travel alone. It requires patience, encouragement, and support from many people, specially one's supervisor. I would like to take this opportunity to extend my sincere respect and gratitude to my supervisor Dr.T.S.G.Peiris (Senior Lecturer in Statistics, Department of Mathematics, University of Moratuwa) for his intellectual inspiration, patience, guidance and encouragement. Without him this thesis could not have been completed.

Also I would like to thank all of my colleagues with whom I studied at the University of Moratuwa. Their encouragement and support are very much appreciated.

Last but not least, I would like to deeply thank my parents and my husband who always stand by me.

ABSTRACT

Exploratory Factor Analysis (EFA) is a technique to explore the underlying factors of a large set of observed variables which cannot be measured directly. In general there are seven types of factor extraction methods. For a meaningful interpretation of occurred factor model, the extraction method usually followed by either Orthogonal rotation or Oblique rotation method. However it has not been recommended a particular method of EFA for a given set of data. Further most of the researchers are misusing Principal Component Analysis (PCA) with Exploratory Factor Analysis. Therefore, this study was carried out to investigate a possibility of recommending a particular method for a given set of data using a data set comprising seven variables on crimes. Data were analysed using the statistical software SPSS.

To illustrate the contrast of PCA and EFA, analysis was begun with Principle Component. For the comparison of different types of extraction methods under EFA, variables were extracted using Maximum Likelihood Factoring, Principle Axis Factoring and General Least Squares followed by all the Orthogonal rotation methods separately. The steps of the analysis in EFA were quite same with all the extraction methods, however the final result and the effect of the prior assumptions make difference. It is very important to confirm that KMO statistic to be greater than 0.6, prior to carry out EFA for the adequacy of sample size in order to derive valid statistical inferences. If the variables having multivariate normal distribution it is recommended to conduct Maximum Likelihood or General Least Squares. For the non normal distributions, Principle axis factoring is recommended. However it is recommended to compare the results from each method irrespective of the distribution of data set.

Among Orthogonal rotations Varimax rotation is recommended as it provides simple factor loadings to interpret. Quartimax generally does not provide simple factor loadings as in Varimax. It is not recommended to carry out all possible combinations of factor extraction methods and rotation methods to any set of data, as same results will not be produced by each combination. The recommendation given for the particular data set was confirmed using Jackknife validation method.

Keywords: Factor analysis, Generalized Least Squares, Maximum Likelihood Extraction, Orthogonal rotations, Principal Axis Factoring

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Description
FA	Factor Analysis
PCA	Principle Component Analysis
EFA	Exploratory Factor Analysis
ML	Maximum Likelihood
PF	Principle Factoring (Using for Principal Axis Factoring)
GLS	Generalized Least Squares
V	Variance

CHAPTER 01

INTRODUCTION

1.1 Concept of Factor Analysis

The Factor Analysis was first introduced in 1904 by *Charles Spearman* who is known as the father of Factor Analysis (Fabrigar, 1999). Factor analysis basically allowed the investigator to find out the underlying factor structure of observed variables which cannot be measured directly. There are two types of factor analysis namely Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) .In EFA there is a collection of many extraction methods. Maximum Likelihood, Principle Axis Factor, Weighted Least Squares , Un weighted Least Squares, Generalized Least Squares, Imagine Analysis, Minimum Residual Analysis and Alpha Factoring to name just some.

Principal Component Analysis (PCA) is sometimes misused with Factor Analysis as in both methods there is a variable reduction procedure. However, Principal Component Analysis cannot be considered as a Factor Analysis and it differs with the purpose of use. If the purpose is to reduce the information in many variables in to a set of weighted linear combinations of those variables then PCA is the appropriate method to use. Factor Analysis is used when the purpose is to identify the latent variables which are contributing to the common variance in a set of measured variables. Thus the aim of factor analysis is to reveal any latent variables that cause to the covariance in the observed variables. Therefore during factor extraction the common variance of a variable is partitioned from its unique variance and error variance and only the common variance appears in the solution. In contrast Principal Component Analysis does not differentiate between common and unique variance, as it ignores the immeasurable effects that really exists. With this issue in some situations such as, when the factors are uncorrelated and communalities are moderate, it can produce inflated values of variance accounted for by the components (Costello, 2005).

Due to the discrepancies between the solutions when applying different combinations, the researchers have tend to apply their own preferred factor extraction method for the analysis with the popularity they have been given in researches so far. Once the well known publisher (Harmen,1976) has mentioned in his book; "Modern Factor Analysis", the preferred type of factor solutions are determined on the basis of two general principles namely (i) statistical simplicity and (ii) scientific meaningfulness.

However for a given set of data, it is rather difficult to recommend a particular factor extraction and factor rotation method. Thus the common practice is to apply few combinations of factor extractions and factor rotation methods to make sure that the final results are invariant of the type of rotation & factor extraction method.

1.2 Factor Extraction Methods

There are many books written on this subject of factor analysis and its applications. Although the main purpose of factor analysis was based on the psychology at the very beginning, with the arrival of new extraction methods, the role became vital after 1950's with meteorology and medicine, political and taxonomy, archaeology and economy and many more.

Popular writers such as H.Harmen , Charles W. Muller , Jae-on Kim and Thimothy A.Brown have mentioned in their published books about the comparisons they and their researches' have made over these number of methods. Before moving on to the comparisons of different publishes, I first concern on the popular method of factor extraction which is somewhat under argument level, the Principle Component Analysis. Although Principal Component Analysis (PCA) is the default method of extractions in many popular statistical software packages, including SPSS and SAS it is under argument whether it is a true method for factor analysis.

As described in the paper "Best Practices in Exploratory Factor Analysis", PCA is only a data reduction method. It became common decades ago when computers were slow and expensive to use; it was quicker, cheaper alternative to factor analysis. It is computed without regarding any underlying structure caused by latent variables; components are calculated using all of the variance of the manifest variables, and all of that variance appears in the solution. However many authors analyze data without an a priori idea about how the variables are related. (Costello & Osborne, 2005)

In a review of article on "Use of Exploratory Factor Analysis" published by the Department of psychology in University of East Carollina, (http://core.ecu.edu/psyc/wuenschk/stathelp/efa.htm), revealed that the Maximum Likelihood extraction allows computation of assorted indices of goodness-of-fit and the testing of the significance of loadings and correlations between factors, but requires the assumption of multivariate normality. Principal Factors (Principal Axis Factoring) method has no distributional assumptions. Therefore they suggested that, the analyzer should first examine the distributions of the measured variables for normality. Unless there are severe problems like |skew| >2 or |kurtosis| >7 for measured variables it has been recommended the ML extraction. If there are severe problems as above it has been suggested to transform variables rather than using Principal Factor methods.

Further Brown (2006) has compared advantages and disadvantages in Maximum Likelihood and the other methods of Factor Analysis with his experiments on data analysis. A key advantage of the Maximum Likelihood estimation method is that it allows for a statistical evaluation of how well the factor solution is able to reproduce the relationships among the indicators in the input data. In other words it detects how closely the correlations among the indicators predicted by the factor analysis parameters, approximate the relationship seen in the input correlation matrix. However, he also emphasized that the Maximum Likelihood estimation requires the assumption of multivariate normal distribution of the variables. If the input data depart substantially from a multivariate normal distribution, important aspects of the results of ML estimators in EFA can be distorted and not trustworthy (e.g. goodness of model fit, significance tests of model parameters). Another potential disadvantage of ML estimation is, its occasional tendency to produce "improper solutions". An improper solution exists when factor model does not converge on a final set of parameter estimates, or produces an "out of range" estimate such as an indicator with communality above 1.0. However, PF has the strong advantages of being free of distributional assumptions and of being less prone to improper solutions than ML.

Unlike ML, PF does not provide goodness-of-fit indices useful in determining the suitability of the factor model and the number of latent variables. Thus PF might be preferred in instances where marked non-normality is evident in the observed measures or perhaps when ML estimation produces an improper solution may be a sign of more serious problems, such as a poorly specified factor model or a poorly behaved input data matrix. (Brown, 2006)

It is far less the number of articles and books have written on the other factor extraction methods such as Weighted Least Squares, Un weighted Least Squares, Generalized Least Squares, Imagine Analysis, and Alpha Factoring which leads us to believe that the usage of those methods may less due to some unrevealed issues.

1.3 Rotation Methods

A rotation is a linear transformation that is performed on the factor solution or the extracted factors, for the purpose of making the solution easier to interpret. It further allows you to identify meaningful factor names or descriptions. There are two main rotation methods (Orthogonal and Oblique) following the extraction methods done. The Orthogonal is carried out when the axis are also orthogonal, and Oblique is done when the new axis are not required to be orthogonal to each other.

In review article by (http://core.ecu.edu/psyc/wuenschk/stathelp/efa.htm) University of East Carollina , they provide a strong argument in favor of oblique rotations rather than orthogonal solutions. They highlighted that dimensions of interest to psychologists are not often dimensions we would expect to be orthogonal. If the latent variables are, in fact, correlated, then an oblique rotation will produce a better estimate of the true factors and a better simple structure than will an orthogonal rotation. If the oblique rotation indicates that the factors have close to zero correlations between one another, then the analyst can go ahead and conduct an orthogonal rotation (which should then give about the same solution as the oblique rotation). They also claimed that an oblique rotation often produced a slightly better simple structure than did a Varimax rotation which comes under orthogonal rotation methods, but the pattern of loadings was almost always the same with Varimax as with oblique rotation.

In Orthogonal rotations the most widely used of these is the Varimax criterion. It seeks the rotated loadings that maximize the variance of the squared loadings for each factor; the goal is to make some of these loadings as large as possible and the rest as small as possible in absolute value. The Varimax method encourages the detection of factors each of which is related of few variables. It discourages the detection of factors influencing all variables.

The Quartimax criterion, on the other hand, seeks to maximize the variance of squared loadings for each variable and tends to produce factors with high loadings for all variables. (Tryfos, 2001)

1.4 Objectives of the Study

The various combinations in Exploratory Factor extraction methods and factor rotation methods have come across with different levels of applications which some are with the illustrations by researches and methodologists and some still with the argumentries which they have failed to give an exact proof. Under these circumstances the objectives of this study are as follows.

(i) To review some of the existing Extraction methods.

(ii) To review the use of rotation methods.

(iii) To recommend a particular method of combination for a given set of data.

1.5 Secondary Data

A secondary data set obtained from SAS manual is shown in Appendix 1.1. It is consisting of seven types of variables named as average number of murders (Murder), number of rapes (Rape), number of robberies (Robbery), number of burglaries (Burglary), number of larceny (Larceny), number of assault (Assault) and number of auto crimes (Auto) per 100,000 people in different states of America.

CHAPTER 2

BASIC THEORY IN FACTOR ANALYSIS

2.1 Overview

This chapter will outline separate introductions to several methods come under factor analysis in order to have a basic understanding on each method that will be using for the comparison in the Analysis. It also provides descriptions on the key terms and output tables which are generally included in factor analysis, dealing with the outputs of SPSS software package as per this dissertation.

2.2 What is Factor Analysis?

Factor analysis is a method of data reduction technique in multivariate environments without affecting the covariance of the initial system. Basically it consists with Principal Component Analysis and Common Factor Analysis.

The Common factor Analysis comprises with many methods which identify the internal structure of a set of variable. Some basic objectives of Factor Analysis are:

I. To determine how many common factors are needed to explain the set of variables.
II. To find the extent to which each variable is associated which each of set of common factors.
III. To provide interpretation to the common factors.
IV. To determine the amount of each factor possessed by each manifest variable.

The figure 2.1 shows the composition of factor analysis with the basic definition of each main type of analysis.

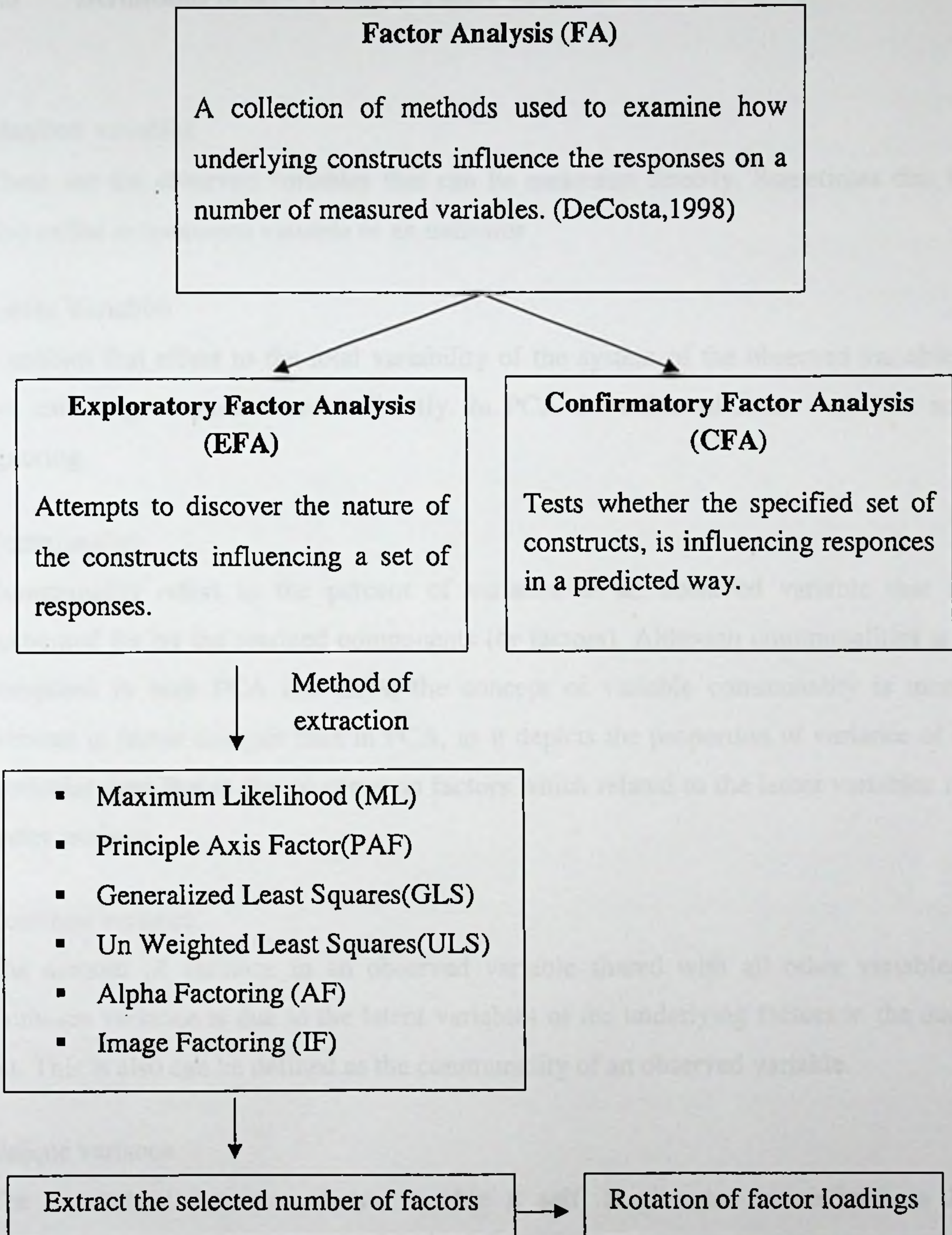


Figure 2.1: Structure of methods under Factor Analysis

The structural details of EFA and PCA and geographical representation of both analysis systems will be discussed in separate section in 2.4 and 2.5 respectively.

2.3 Definitions of Key Terms in Factor Analysis

Manifest variables

These are the observed variables that can be measured directly. Sometimes this is also called as measured variable or an indicator.

Latent Variables

Variables that effect to the total variability of the system of the observed variables but cannot get a measurement directly. In PCA the effect of latent variables are ignoring.

Communality

Communality refers to the percent of variance in an observed variable that is accounted for by the retained components (or factors). Although communalities are computed in both PCA and EFA, the concept of variable communality is more relevant in factor analysis than in PCA, as it depicts the proportion of variance of a particular item that is due to common factors which related to the latent variables in factor analysis.

Common variance

The amount of variance in an observed variable shared with all other variables. Common variance is due to the latent variables or the underlying factors in the data set. This is also can be defined as the communality of an observed variable.

Unique variance

The amount of variance of the variable it self. It also can be defined as 1-communality. If we regard this uniqueness is due to another factor, then this factor is unique to that variable only. In PCA there is no separation of Unique and Common variance, it accounts all the variance in an observed variable while in EFA it consider only the Common variance.

Correlation matrix

Correlation is a measure of the relation between two or more variables. Correlation coefficients can range from -1.00 to + 1.00 . The value of -1.00 represents a perfect negative correlation while a value of +1.00 represents a perfect positive correlation. A value of 0.00 represents a lack of correlation. With the different methods under Exploratory Factor analysis, the correlation matrix is compiled with different types of variable correlations. It can be between observed variables and between estimated or predicted variables.

Total variance of a data set

Total variance is more relevant when talk about the Principal Component Analysis. Prior to the analysis of each data set the first step is to standardize the observed variables (mean = 0, variance = 1). The total variance in the data set is simply the sum of the variances of these observed variables. On the other hand total variance can be defined as the sum of the common variance and unique variance when applying to the factor analysis.

Factor loadings

Correlation coefficients between the variables (rows) and factors (columns). On the other hand it can be defined as the contribution of certain factor amount of certain observed variable loaded on specific factor.

Eigen values

Indicate the amount of variance explained by each principle component or each factor.

Scree plot

This is a method of selecting the number of factors to inclusion. It is a plot of the Eigen values of the correlation matrix associated with each component and look for a "break" between the components with relatively large Eigen values and those with small Eigen values. The components that appear before the break are assumed to be meaningful and are retained for rotation; those appearing after the break are assumed to be unimportant and are not retained.

Factor matrix (matrix with factor ladings)

This table gives how much each manifest variable loads onto each of the latent variable before rotation. After rotation the same matrix comes as "Factor rotated matrix". From the Factor rotated matrix we can clearly interpret the newly defining factors for further analysis.

2.4 Principal Component Analysis (PCA)

The purpose of Principal Component Analysis is to derive a small number of components which are independent and that can be accounted for most of the variance in the large set of measured variables. Those derived components are defined simply as linear combination of measurements which are optimally weighted and there is no separation of variances, will contain both common and unique variances. A model of Principal Component Analysis of 5 variables reduced to 2 components can be represented as shown in figure 2.2 (Costa, 1998).

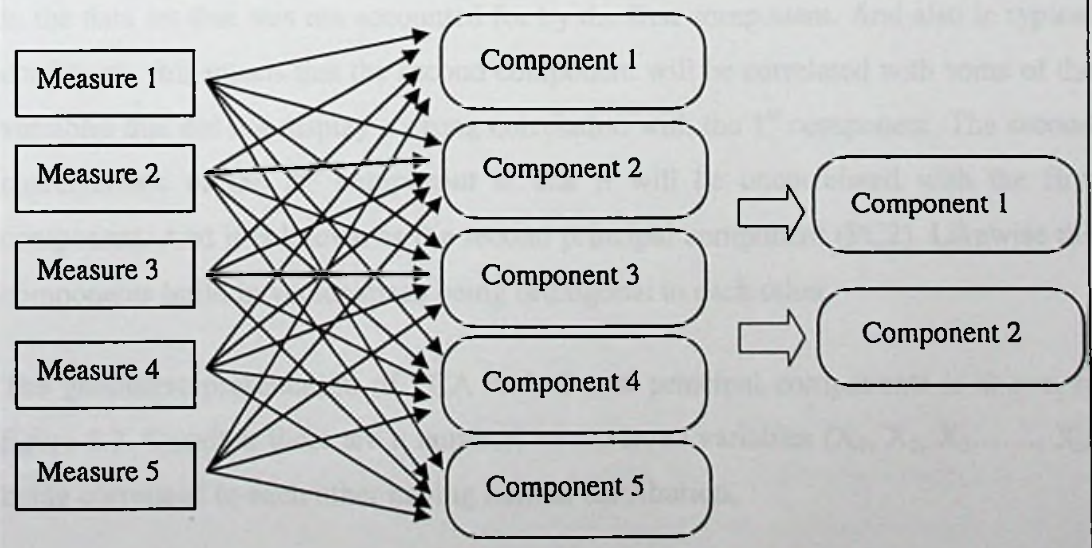


Figure 2.2: The model for Principle Component

There is no special consideration on differentiating the unique and common variance, as Principal Component Analysis ignores the immeasurable effects that really exist. It only attempts to keep a meaningful amount of variance of the initial system with the retained components. If the total variance of the system is $\sum_i^n V(X_i)$ and the variability explained by the retained components is $\sum_i^p V(PC_i)$, PCA attempts to keep the ratio of ($\sum_i^p V(PC_i)$ / $\sum_i^n V(X_i)$) greater than 75% . This is an arbitrary (reasonable) value. The computation of this Principal Components and the geometrical representation of those PCs in vector space will be discussed in next sub topic.

2.4.1 Computation of principal components from manifest variables

The first component extracted accounts for a maximal amount of total variance in the observed variables and it is known as first principal component (PC1). Under typical conditions, this means that the first component will be correlated with many of the observed variables. The second component extracted will have two important characteristics. First, this component will account for a maximum amount of variance in the data set that was not accounted for by the first component. And also in typical conditions, this means that the second component will be correlated with some of the variables that did not display a strong correlation with the 1st component. The second characteristic of the 2nd component is that it will be uncorrelated with the first component. And it is known as the second principal component (PC2). Likewise the components build in vector space being orthogonal to each other.

The geometric presentation of PCA with 2 new principal components is shown in figure 2.3. Consider there are n numbers of observed variables (X_1, X_2, X_3,……, X_n) being correlated to each other having normal distribution.

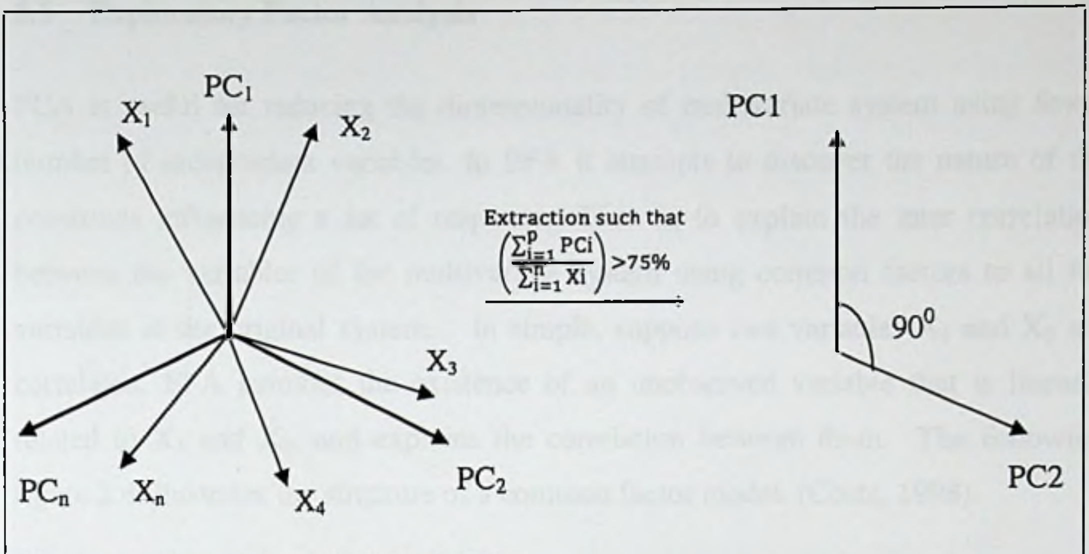


Figure 2.3: Geometric representation of PCA

In general, Principal component analysis chooses a coordinate system for the vector space consisting of all linear combination of vectors. The first principal component points in the direction of maximum variation in the data. The dimensional reduction is achieved by ignoring dimensions which do not explain much variation. According to this computation the basic model for a Principal Component comes as a linear combination of n number of observed variables such that;

$$PC_i = \sum_{j=1}^{n} a_{ij} X_j \; ; \; i = 1,2, \dots p$$

For i =1 ; $PC_1 = a_{11}X_1 + a_{12}X_2 + \dots + a_{1j}X_j + \dots + a_{1n} X_n$

where PC_i is the i^{th} Principal Component , a_{ij} is the regression coefficient or the weight for j^{th} observed variable (X_j) , used in creating i^{th} Principal Component. The regression coefficient for the same observed variable will differ according to the weight it gives creating the new Principal Component. Such that in PCA the observed data reduced to small number of new components remaining maximum amount of total variability of the initial system.

2.5 Exploratory Factor Analysis

PCA is useful for reducing the dimensionality of multivariate system using fewer number of independent variables. In EFA it attempts to discover the nature of the constructs influencing a set of responses. That is, to explain the inter correlation between the variables of the multivariate system using common factors to all the variables in the original system. In simple, suppose two variables X_1 and X_2 are correlated. EFA assumes the existence of an unobserved variable that is linearly related to X_1 and X_2, and explains the correlation between them. The following figure 2.4 illustrates the structure of a common factor model. (Costa, 1998).

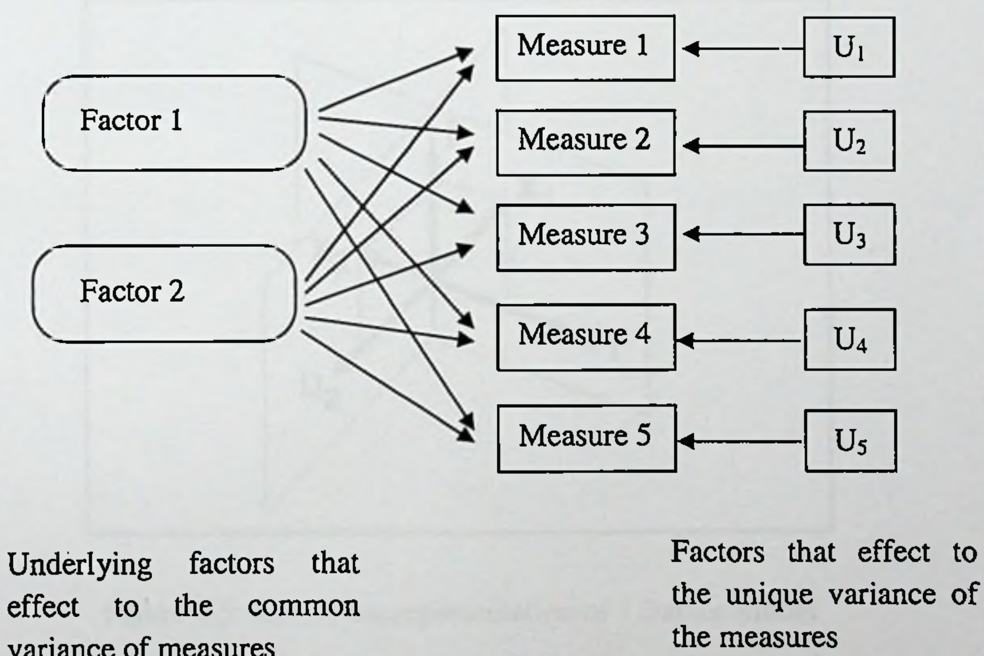


Figure 2.4: The structural model for 2-Common Factors

This model shows that each observed response (i^{th} measure) is influenced partially by unique factors (U_i) and at the same time partially influenced by underlying common factors (Factor 1 and Factor 2) that cannot be measured directly.

The ultimate goal of the factor analysis is to find out this unobserved variable from the structure of the original variables. This estimated unobserved variable is called a common factor. As the FA perform by examining the pattern of correlation between the observed measures, one can get a priori idea such that the measures that are highly correlated are likely influenced by the same factors, while those that are relatively uncorrelated are likely influenced by different factors.

2.5.1 1 – Factor model for EFA

The geometry of the relationship between original variable comprises with the two common factors and unique factors is illustrated in figure 2.5.

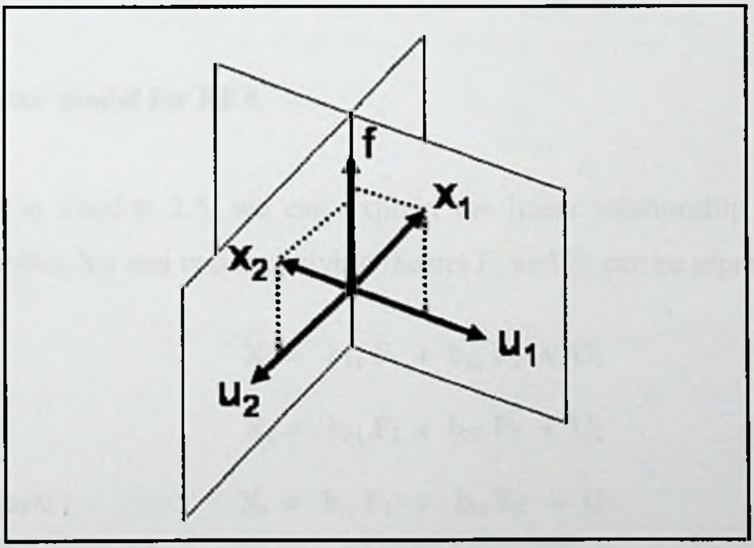


Figure 2.5: Geometric representation of 1-Factor Model

The correlated variables X1 and X2 are shown schematically in this figure. Each of these vectors are decomposed in to linear combination of a common factor F and its unique factor U_i assuming that there is only one underlying factor. This is known as 1-Factor model and it is represented as,

$$X_i = b_i F + U_i \, ; \, i = 1, 2$$

Then; $X_1 = b_1 F + U_1$

$$X_2 = b_2 F + U_2$$

The unique factors U_1 and U_2 which contribute to the unique variance of each factor are uncorrelated with the common factor F and with each other. So that all three components F, U_1 and U_2 are mutually orthogonal.

The number of dimensions is increasing with the total number of common factors and the unique factors as all those are independent or orthogonal from each other. As in contrast in PCA the number of dimensions are equal to the number of Components built.

2.5.2 2- Factor model for EFA

As discussed in Section 2.5, we can explain the linear relationship between the observed variables X_is and two underlying factors F_1 and F_2 can be represented as;

$$X_1 = b_{11} F_1 + b_{12} F_2 + U_1$$

$$X_2 = b_{21} F_1 + b_{22} F_2 + U_2$$

In general ; $X_i = b_{i1} F_1 + b_{i2} F_2 + U_i$

The terms U_1 and U_2 are known as the unique factors of observed variables or sometimes it defines as the error term when the hypothesized factors are not exactly define the whole variability of the observed variable. We will discuss the way of calculation the variability of Xi's later on this section comparing methods of PCA and EFA. The parameters b_{ij} are referred to as loadings. For an example b_{21} is called the loading of variable X_2 on factor F_1 or the correlation coefficient between the variable and the factor.

When computing the variability of variable X_i we should make some assumptions. The unique factors or the error terms are independent from each other, therefore $E(U_i) = 0$ and $V(U_i) = \sigma_i^2$. Since the factors cannot be measured we make an assumption as they have measured in standardized form. And these factors are independent from one another and with unique factors, therefore $E(F_i) = 0$ and $V(F_i) = 1$. The variance of Xi can be calculated with these two assumptions.

$$X_i = b_{i1} F_1 + b_{i2} F_2 + U_i$$

$$V(X_i) = b^2_{i1} V(F_1) + b^2_{i2} V(F_2) + V(U_i)$$

$$V(X_i) = \underbrace{b^2_{i1} + b^2_{i2}}_{\text{Communality}} + \underbrace{\sigma_i^2}_{\text{Unique variance}}$$

The variance of Xi can be described by two parts, communality which is explained by the common factors F_1 and F_2 , and the specific or the unique variance that is unique for the specific variable or the part which is not accounted by the common factors. Due to this separation of variability the percentage of variance explained by the factors after the extraction are getting less in EFA methods in contrast with the Principle Component Analysis. At the same time the amount of the contribution of each factor (let us say Factor 1) in explaining sum of the squared observed variance or the extraction sums of squared loadings ($\sum b^2_{i1}$; $i = 1, 2\ldots7$) also having a value below the initial Eigen values. This inequality is due to the extraction of factors considering only the common variance when conducting EFA methods.

2.5.3 Steps to follow in Exploratory factor analysis

1) Collect Measurements

There should be a sufficient number of measurements, as factor analysis is a technique that requires a large sample size with variables that has measured in same experimental units. Many experiments have done in order to find out the adequacy of sample size. In general the sample size with 50 cases is very poor, 100 is poor, 200 is fair, 300 is good, 500 is very good, and 1000 or more is excellent. Over 300 cases is probably adequate, but the communalities after extraction should probably be above 0.5 (Field, 2005). In SPSS there is a statistical test called Kaiser-Meyer- Olkin (KMO) to test the adequacy of sample size.

2) Standardizing collected data

Depending on the matrix which going to be used for analysis, we need to do standardization for the data if those have measured in different scales. If the data have measured in same scale, then it is better to use covariance matrix. If the data have measured in different scales there are two ways, either we can use correlation matrix or first standardize the variables and take covariance matrix which is same as correlation matrix.

$$\text{Standardization of variables: } Z_{ij} = \frac{(X_{ij} - \mu_i)}{\sigma_i}$$

Where X_{ij} is the j^{th} observation of i^{th} manifest variable, μ is the mean of i^{th} manifest variable and σ_i is the standard deviation of i^{th} manifest variable. Then Z_{ij} represents how much amount of an observation, deviates above or below the mean of the respective manifest variable.

3) Obtain the Correlation Matrix

Obtain the correlation matrix between all the variables. This matrix can be used to check the pattern of the relationships before moving to the real analysis. The correlation matrix obtained should be checked before proceed the other steps, with the problems of Singularity and Multicollinearity (check whether Determinant > 0.00001) which depicts that there is no highly correlated (R> 0.8) or perfectly correlated (R = 1.00) variables. At the same time the correlation matrix should check with the hypothesis of not being an identity matrix (H_0: $\sum = I$ vs H_1: $\sum \neq I$). The Bartlett's test of sphericity will test this hypothesis.

4) Select the number of Factors to inclusion

There are number of methods to determine the "optimal " number of factors by examining the data. Kaiser criterion, Scree test, Method of proportion of variance accounted, and the interpretability criteria are some of the methods use in common. One of the most popular method The Kaiser Criterion states that you should use the number of factors equal to the number of eigenvalues of the correlation matrix that are greater than one.

5) Extract the initial set of factors.

This step is more important as one should know the exact purpose for the extraction; whether it is only for data reduction or for determine the underlying factors. In this specific topic of exploratory factor analysis which aims to identify the underlying factors, there are several methods that can be applied as mentioned in Figure 2.1.

6) Rotate the extracted factors to the final solution

By rotating the factors one can attempt to find a factor solution which has a simplest interpretation. And the method of rotation can be selected according to the analyzers' ultimate goal of having correlated factors or uncorrelated factors.

3) Interpret the factor structure.

Each of the measures will be linearly related to each of the common factor and to the unique factor. The strength of the relationship is illustrated from the respective Factor loading, produced by the rotation.

2.6 Methods of Exploratory Factor Analysis

The methods are varying with the composition of correlation matrix that computes for the extraction of determined number of factors. The diagonal of the correlation matrix is replacing with estimated communalities. The way of estimating communalities varies with the method of extraction. Then the new eigen values are created for the extracted factors. And also with some of the methods there are specific assumptions that have to be made before the analysis begun. This section describes each of the extraction method available in software SPSS, giving a basic idea on how it differs with other extraction methods. The details for the definitions are gathered from the stat notes of Factor Analysis sited by the North Carolina State University under public administration program. These extraction methods are, Maximum Likelihood Factoring, Principal Axis Factoring, Un weighted Least Square, Generalized Least Squares, Alpha Factoring, Image Factoring and Principal Component Analysis.

2.6.1 Maximum Likelihood Factoring

Maximum Likelihood method assumes multivariate normality. Correlations are weighted by each variable's uniqueness (uniqueness is 1 minus communality of the variable). Maximum Likelihood factoring generates a Chi-square goodness-of-fit test. From that the researcher can increase the number of factors one at a time until a satisfactory goodness-of-fit is obtained.

2.6.2 Principal Axis Factoring

A method of factor analysis in which the factors are based on a reduced correlation matrix using a priori communality estimates. These initial communalities are determined by the squared multiple correlation of the variable with the other variables. The communalities are inserted in the diagonal of the correlation matrix and the extracted factors are based only on the common variance, with specific and error variance excluded.

2.6.3 Unweighted Least Square

Based on minimizing the sum of squared of differences between observed and estimated correlation matrix, not counting the diagonal.

2.6.4 Generalized Least Square

Based on adjusting Unweighted Least Squares by weighting the correlations inversely according to their uniqueness. (more unique variables are weighted less) Uniqueness is defined as 1- h^2 , where h^2 is the communality of the observed variable. Like in Maximum Likelihood Factoring, in Generalized Least Squares also generates a chi- square goodness-of-fit test.

2.6.5 Alpha Factoring

Based on maximizing the reliability of factors, assuming variables are randomly sampled from a universe of variables. All other methods assume cases to be sampled and variables fixed.

2.6.6 Image Factoring

Image factoring based in the correlation matrix of predicted variables rather than actual variables, where each variable is predicted from the others using the multiple regression.

2.7 Methods of Rotations

With the purpose of interpreting the ultimate factors in a meaningful way there are two major types of rotations. Those are Orthogonal Rotation methods and Oblique Rotation methods. Orthogonal rotations impose the restriction that the factors can not be correlated. And there are three main sub criterions called Varimax, Quartimax and Equamax. Oblique rotation allows the factors to be correlated with one another. Methods of Oblimin and Promax come under the Oblique rotations.

Crawford and Ferguson have introduced a general equation for the rotations, thus Varimax, Equamax and Quartimax rotations are special cases with different substitutions. (http://www.stata.com/bookstore/stata12/pdf/mv_glossary.pdf)

k	Rotation
0	Quartimax
1/p	Varimax
f/2p	Equamax

$$C(\Lambda) = \frac{1-k}{4} \langle \Lambda^2, \Lambda^2 (11' - I) \rangle + \frac{k}{4} \langle \Lambda^2, (11' - I) \Lambda^2 \rangle$$

Where A is the matrix to be rotated, T is the rotation and $\Lambda = AT$. Further p is the number of rows of A and f is the number of columns of A.

2.7.1 Varimax criterion

This is the most widely used rotation method. It allows the analyzer to interpret the results much easier than the other methods. Varimax criterion creates rotated loadings in such a way that maximize the variance of the squared loadings for each factor and the goal is to make some of these loadings as large as possible and the other loadings as small as possible in absolute value. Therefore one can identify the best factors, each related to few variables without the factors that influence all variables.

2.7.2 Quartimax criterion

When compared to Varimax Criterion, Quartimax Criterion also tries to maximize the variance of squared loadings, but not for each factor but for each variable. So it tends to produce factors with high loadings for all variables. Therefore it's difficult to interpret the results as we do in Varimax criterion.

2.7.3 Equamax criterion

This criterion is less in use. As it does not maximize or minimize the variances of squared loadings it is hard to interpret the results.

CHAPTER 03

RESULTS AND DISCUSSION ON FACTOR EXTRACTION AND ROTATION METHODS

3.1 Preliminary Analysis of Data

3.1.1 Descriptive statistics

Some of the descriptive statistics of observed variables: the number of murders (Murder), the number of rapes (Rape), the number of robberies (Robbery), the number of assaults (Assaults), the number of burglaries (Burglary), the number of larcenies (Larceny) and the number of autos (Auto) are shown in table 3.1.1.

Table 3.1.1: Descriptive statistics of observed variables

Variable	Mean	Standard Deviation	Sample Variance	Skewness	Kurtosis
Murder	7.4	3.9	14.9	0.29	-0.7
Rape	25.7	10.8	115.8	0.596	-0.26
Robbery	124.1	88.3	7805.5	1.67	4.06
Assault	211.3	100.3	10050.7	0.61	0.107
Burglary	1291.9	432.5	187017.9	0.56	0.44
Larceny	2671.3	725.9	526943.5	0.37	-0.31
Auto	377.5	193.4	37401.4	1.69	3.99

The results in Table 3.1.1 indicate that there is a low mean (7.4) and standard deviation (3.9) to the number of murders when compared to other variables. The number of larcenies (Larceny) is having the highest mean (2671.3) and standard deviation (725.9).

The values of kurtosis and skewness, are useful to get an idea about the normality of each variable. For univariate data Y_1, Y_2...,Y_n the formulas for kurtosis and skewness can be written as follows; where $\bar{Y}$ is the mean, s is the standard deviation and n is the number of data points.

$$\text{Skewness} = \frac{\sum_{i=1}^{N}(Y_i-\bar{Y})^3}{(N-1)s^3}$$

$$\text{Kurtosis} = \frac{\sum_{i=1}^{N}(Y_i-\bar{Y})^4}{(N-1)s^4} - 3$$

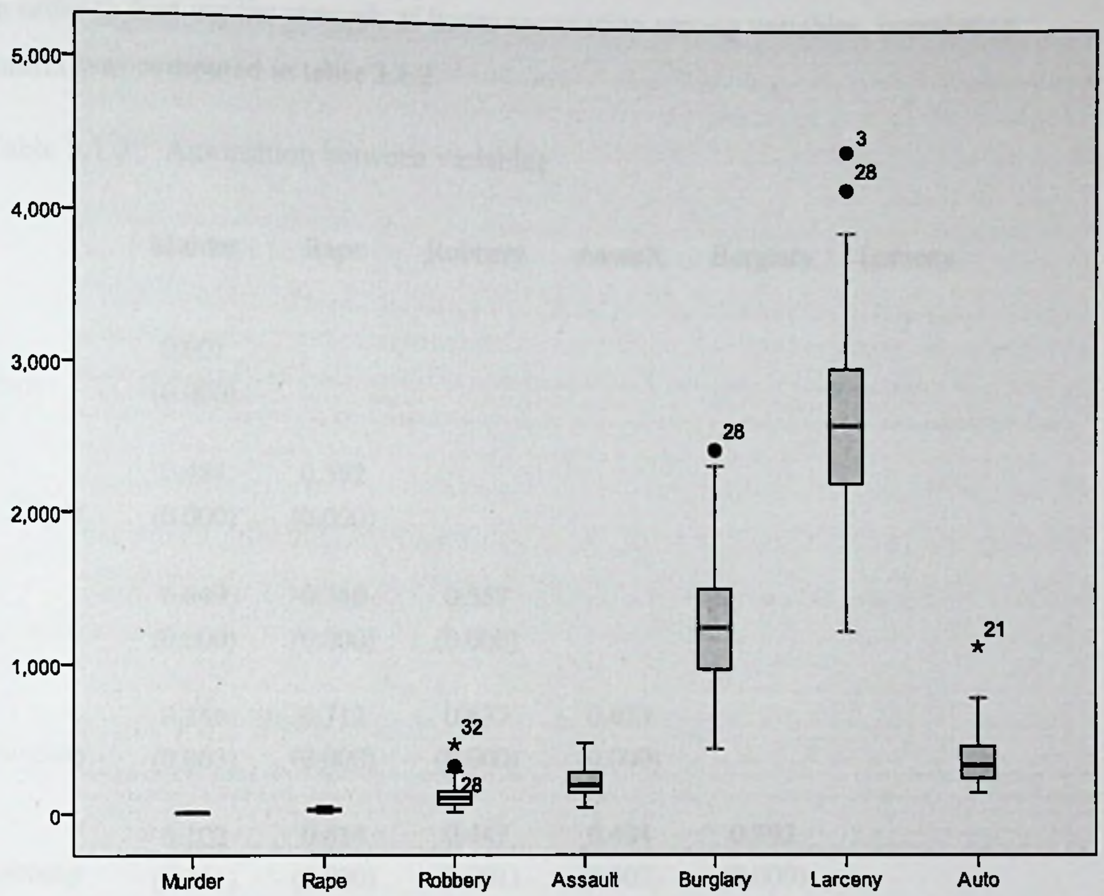
Skewness tells about the symmetry of the distribution. In a normal distribution the skewness is zero which indicates that the variable has a perfectly symmetric distribution. Kurtosis tells about the peakedness of the distribution. Kurtosis in a normal distribution is three and it looks like a bell- shaped which not too peaked or flat. Therefore for a perfect normal distribution the kurtosis and skewness should be three and zero respectively.

When considering the skewness of these seven variables it can be assumed that all variables except for Robbery and Auto are having skewness close to zero indicating that it can be assumed all variables are normally distributed except Robbery and Auto. At the same time the values of kurtosis in Robbery (4.06) and Auto (3.99) are deviate from three while the other variables are having Kurtosis close to three. Therefore we can consider that the number of murders, number of rapes, number of assault, number of burglary and number of larceny are having fairly normal distributions while number of robberies and number of autos are having non-normal distributions.

3.1.2 Box plot of observed variables

In order to see the variability and skewness of each variable in a pictorial form, box plot was drawn as shown in figure 3.1 for each variable.

Figure 3.1: Box plot of manifest variables



When consider the inter quartile range of each variable, the dispersion of the number of burglaries and number of larcenies are higher than that of the other variables indicating that those two variables have higher variability. On the other hand the number of murders, number of rapes and number of robberies are having a small variation when compared to other four variables.

In this method the observations above the upper limit $[Q_3 + 1.5(Q_3\text{-}Q_1)]$ and below the lower limit $[Q_1 - 1.5(Q_3\text{-}Q_1)]$ are taken as outliers. (Q3 = 3rd quartile and Q1 = 1st quartile). The whiskers of number of robberies, number of burglaries and number of auto crimes indicate of having skewed distributions confirming the results obtained in descriptive analysis. According to the above subjective criteria it can be seen that there are few outliers in Robbery, Burglary, Larceny and Auto (Fig: 3.1).

3.1.3 Association between variables

In order to find out the strength of linear association among variables, correlation matrix was computed in table 3.1.2.

Table 3.1.2 : Association between variables

	Murder	Rape	Robbery	Assault	Burglary	Larceny
Rape	0.601 (0.000)					
Robbery	0.484 (0.000)	0.592 (0.000)				
Assault	0.649 (0.000)	0.740 (0.000)	0.557 (0.000)			
Burglary	0.386 (0.003)	0.712 (0.000)	0.637 (0.000)	0.623 (0.000)		
Larceny	0.102 (0.241)	0.614 (0.000)	0.447 (0.001)	0.404 (0.002)	0.792 (0.000)	
Auto	0.069 (0.317)	0.349 (0.007)	0.591 (0.000)	0.276 (0.026)	0.558 (0.000)	0.444 (0.001)

Parenthesis indicates the probability value for the significance of correlation at 5% .

According to the correlation matrix it reveals that there are significant correlations ($p < 0.05$) between any two observed variables except between number of Larcenies & number of murders and between number of autos & number of murders. The highly dependence structure among variables justifies the use of Factor analysis. This was further confirmed using KMO statistics shown in Table 3.1.3.

Table 3.1.3 : Kaiser - Meyer - Olkin Measure of Sampling Adequacy

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.784
Bartlett's Test of Sphericity	Approx. Chi-Square (X^2_{21}) (p = 0.000)	219.506

KMO measure of 0.784 (> 0.6) indicates high sampling adequacy for the factor analysis. On the other hand it indicates that the partial correlations are adequate to carry out the factor analysis. Barlett's test of sphericity is to test whether the correlation matrix is an identity matrix.

H_0 : Correlation matrix = I_{6x6} Vs H_1 : Correlation matrix is not an identity matrix

In other words, Bartlett's test addresses the question of whether the correlation matrix should be factored.

Test statistic,

$$BS = -\left[(n-1) - \frac{(2p-5)}{6}\right]\log|R|$$

Where R = $\prod_{i=1}^{p} \lambda_i$ and λ_i - i^th^ eigen value, n = number of observations and p = number of variables, under the hypothesis of;

$$H_0 : BS \sim \chi^2_{\frac{p^2-p}{2}}$$

According to the above result shown in table 3.1.4, the Bartlett's test statistics is highly significant (X^2_{21} test statistics = 219.51, p = 0.000) confirming that the null hypothesis can be rejected. Thus it can be further justified the association between the variables are significant, therefore it revealed in pursuing some form of dimension reduction.

As Bartlett's test confirmed that factor analysis can be carried out the data were standardized prior to the analysis. As the number of each variable is highly vary with respect to sizes, data were standardized to bring t a common scale.

3.1.4 Find out the number of components to retain

Table 3.1.4 : Eigen value analysis

Component	Eigen values	% of variance explained by each variable	Cumulative % of the variability explained
PC1	4.115	58.79	58.79
PC2	1.239	17.70	76.48
PC3	0.726	10.37	86.85
PC4	0.316	4.52	91.37
PC5	0.258	3.69	95.06
PC6	0.222	3.17	98.23
PC7	0.124	1.77	100.00

From the table 3.1.4, it is clear that the first two principal components explain relatively large variance whereas the other subsequent components explain only small amount of variance. According to the Kaiser criterion (Eigen value >1) only PC1 (4.445 >1) and PC2 (1.239 >1) can be selected as new variables. The 1st component explains 58.78% of the variability of the initial system while the 2nd component explains about 17.7% of the variability of the initial system. Therefore 76% of the variability of the initial system is explained by the first two components which is a good representation of the original set of data. Thus it can be concluded that the initial 7-D space multivariate data can be reduced to 2-D space, so that two resulting new variables (components) are linear combination of the original variables.

3.1.5 Confirmation of the number of selected components

In the above section, it was confirmed that two components are sufficient according to Keiser rule, as Kaiser (1959) recommended to retain only principal components with eigen values exceeding unity, for standardized data. The graphical approach proposed by Caltell (1966), to plot the variance accounted for each component (eigen value, in the case of correlation matrix) in order from largest to smallest. This is known as Scree plot (Figure 3.2)

Figure 3.2: Scree plot

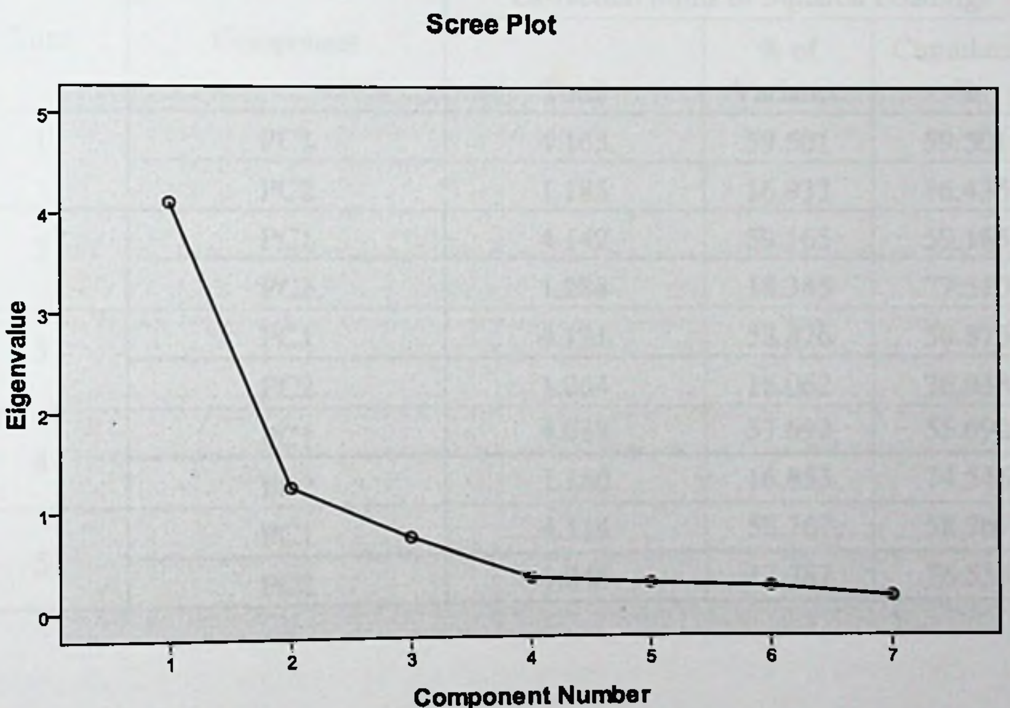


Figure 3.1.2 shows an "elbow" shape at the second eigen value, after that remaining e- values decline in approximately linear function. Therefore scree plot also shows that the first two components account for most of the total variability of the data. The remaining components account for a very small proportion of the variability (close to zero) and are probably unimportant. As the two criteria's giving the same result of selecting two components, two factor model is continuing with further extracting methods.

3.1.6 Jackknife validation

In order to assess the validity of the above solution (selection of two components) Jackknife validation was carried out. (See appendix 1.2). This method suggests carrying out PCA by holding one observation at a time and so PCA was carried out for several times. Following table 3.1.5 summarize the results of selected PCs by the kaiser criterion in each turn.

Table 3.1.5 : Summary of jackknife validation process

Turn	Component	Extraction Sums of Squared Loadings		
		Total	% of Variance	Cumulative %
1	PC1	4.165	59.501	59.501
	PC2	1.185	16.933	76.435
2	PC1	4.142	59.165	59.165
	PC2	1.284	18.345	77.510
3	PC1	4.121	58.876	58.876
	PC2	1.264	18.062	76.938
4	PC1	4.038	57.692	55.692
	PC2	1.180	16.853	74.545
5	PC1	4.114	58.767	58.767
	PC2	1.245	17.787	76.554

It was found in all turns, eigen values greater than one (Kaiser criterion) were only in the first two components. Thus we can easily recommend that 7-D original data series can be reduced to 2-D model.

3.2 Factor Extraction using Principal Component Analysis (PCA)

3.2.1 Preliminary analysis of PCA

Results of preliminary analysis of principal components is shown in table 3.2.1

Table 3.2.1: Results of the Eigen values of the correlation matrix using PCA

Component	Initial Eigenvalues			Extraction Sums of Squared		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
PC1	4.115	58.785	58.785	4.115	58.785	58.785
PC2	1.239	17.696	76.481	1.239	17.696	76.481
PC3	0.726	10.369	86.850			
PC4	0.316	4.520	91.370			
PC5	0.258	3.685	95.056			
PC6	0.222	3.172	98.228			
PC7	0.124	1.772	100.000			

The left panel of the table already explained with table 3.1.4. The number of rows in the right panel corresponds to the number of components (two) retained. It is important to consider the right panel consisting with the variability after the extraction of components. One of the major different achievements in Principal Component Analysis and Exploratory Factor Analysis illustrates the above table. The total variability of the extracted first two components remains in constant (eigen value for PC1 = 4.115 and eigen value for PC2 = 1.239. calculation in table 3.2.3) emphasizing that the components have extracted from the whole variability of the observed variables.

In contrast when conducting EFA, the values of extraction sums of squared or the percentage variance of factors after extraction will be less than that of in PCA.

Communality refers to the proportion of each variable's variance that can be explained by the retained factors. An observed variable will display a large communality if it loads heavily on at least one of the retained components. Although communalities are computed in both PCA and EFA methods, the concept of variables communality is more relevant in EFA than in PCA. In Principle component analysis we compute new few components comprising with linear combination of observed variables, rather than building an underlying factor combination for an observed variable as in EFA.

In Principal Component Analysis as it is assumed that all variances are common or it takes in to account all the variability in observed variables, the initial communalities are all equal to 1. So the diagonal of the correlation matrix contains with ones (1s') when extracting components with PCA. The initial communalities and final communalities of PCA are shown in table 3.2.2.

Table 3.2.2 : Communalities of the components after extraction

Variable	Initial Communality	Final Communality
Murder	1.000	.861
Rape	1.000	.803
Robbery	1.000	.650
Assault	1.000	.794
Burglary	1.000	.848
Larceny	1.000	.726
Auto	1.000	.671

With the retained new components number of murders and number of burglaries, are well represented in the factor space having 86.1% and 84.4% respectively. In the variables number of Robberies (65%) and number of Autos (67.1%) are common with less percentage associated variances. All the other variables are having a good representation of the variability (> 70%). The values of the communalities for two components model can be calculated as shown in below.

Communality for i^{th} variable = (squared loading for PC_{i1} + squared loading for PC_{i2})

$$= \sum PC_{ij}^2 \ , j = 1,2$$

For example communality for the variable number of murders (Murder) is equal to the sum of the squared loadings of PC_1 and PC_2 obtained in Table 3.2.1, $((0.609)^2 + (-0.700)^2) = 0.86$.

3.2.2 Results of new components

Table 3.2.3: Factor loadings for 2- factor model using PCA without rotation

Variable (X_i)	PC1	PC2	Final Communality
Murder	.609	-.700	.861
Rape	.876	-.189	.803
Robbery	.805	.047	.650
Assault	.805	-.382	.794
Burglary	.893	.226	.848
Larceny	.725	.448	.726
Auto	.599	.559	.671
Variance (Eigen value) $= \sum a_i^2$	4.115	1.239	5.354

The results in Table 3.2.3 indicates that component loadings of the first component (PC1) for number of rapes, number of robberies, number of assaults, number of burglaries and number of larcenies are higher (> 0.7) than that for the variables of number of murders and number of autos (0.6) . However, comparing these loadings it is not easy to isolate some variables which load more on the first component than the other variables. Therefore it is not possible to classify a meaningful component consisting with the number of murders, number of rapes, number of robberies, number of assaults, number of burglaries, number of larcenies and number of autos. So it is necessary to rotate factors to obtain meaningful factors.

Therefore the 2 components are generally transformed using various transformations to make the component loading matrix , in particularly to make loadings in component 1 more simple and more meaningful. The proportion of variance in each variable accounted for by two components (communalities) is shown in the fourth column of the table 3.2.3. Factor loadings obtained for 2- factor model using Orthogonal rotations of Varimax, Equamax and Quartimax are shown in next three tables.

3.2.3 Use of Orthogonal rotations

Table 3.2.4: Results of Orthogonal rotations on 2 components

Variable	Rotation method					
	Varimax		Equamax		Quartimax	
	PC1	PC2	PC1	PC2	PC1	PC2
Murder	-.045	.927	-.045	.927	.113	.921
Rape	.501	.742	.501	.742	.620	.647
Robbery	.614	.523	.614	.523	.693	.412
Assault	.316	.833	.316	.833	.453	.767
Burglary	.801	.455	.801	.455	.867	.312
Larceny	.833	.179	.833	.179	.851	.035
Auto	.819	.011	.819	.011	.809	-.128
Variability	2.736	2.616	2.736	2.616	3.213	2.14
Total	5.352		5.352		5.353	

The loadings of each variable on PC1 and PC2 are similar in both Varimax rotation and Equamax rotation. When considering the loadings greater than 0.6 in Varimax and Equamax rotations, variable types of number of murders, number of rapes and number of assaults are highly loaded in Component 2 with 0.927, 0.742 and 0.833

amount of loadings respectively, while number of robberies, number of burglaries, number of larcenies and number of autos are highly loaded in Component 1 with 0.614, 0.801, 0.833 and 0.819 amount of loadings respectively.

When considering the Quartimax rotation, with greater than 0.6 loadings, component one is the highly loaded variables although the loaded values are different, same types of variables have loaded for each component as in Varimax and Equamax rotations. Therefore we can select similar factors from the 3 methods indicating results obtained using Principal Component Analysis are invariant of the rotation method.

3.2.4 Summary of Principal Component analysis

The two components extracted using PCA, were rotated since the interpretation is difficult with the component loadings in component matrix. Using three methods namely Varimax, Equamax and Quartimax by keeping the variance accounted for by the two components remain the almost the same (with 5.35, 76.5%). Further the variance encountered by each variable separately were almost same for all three rotation methods except few loadings in Quartimax rotation varies with the corresponding values in other methods. Thus when selecting the variables for the components with > 0.6, same set of variables loads in 1st component while the rest of the variables loads in 2nd component regardless the method of rotation, indicating the resulted analysis of Principal Component is invariant with the rotation method. Variables number of murders, number of rapes and number of assault load highly on one Principal Component while number of robberies, number of burglaries, number of larcenies and number of autos load highly on the other Principal Component. Thus the ultimate models for the selected two components can be written as linear combinations of each selected standardized variables separately.

PC1 = 0.614 Robbery + 0.801 Burglary + 0.833 Larceny + 0.819 Auto

PC2 = 0.927 Murder + 0.742 Rape + 0.833 Assault

3.3 Factor Extraction using Maximum Likelihood Factoring (ML)

3.3.1 Preliminary analysis of ML

Results of preliminary analysis of ML is shown in table 3.3.1

Table 3.3.1 : Results of eigen values of the correlation matrix using ML

Factor	Initial Eigen values			Extraction Sums of Squared		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.115	58.79	58.79	3.804	54.34	54.34
2	1.239	17.70	76.48	0.962	13.75	68.09

The table 3.3.1 indicates the difference between the percentage variances of initial and after extraction. The amount of percentage variance when extract by Principal Component method was remaining the same value with the initial Eigen values in the left panel of the Table 3.2.1 (pg 32). But in Maximum Likelihood amount of extracted variance getting less. This effect is because in Common Factor Analysis methods, only the common variance is concerned rather than considering the total variance consisting with the unique variance.

Summary results of factor extraction using Maximum Likelihood Factoring is shown in table 3.3.2 with the factor loadings without rotation and Communalities of manifest variables.

Table 3.3.2 : Factor loadings and communality estimates for 2 factor model fitted to data by ML

Manifest Variable	Factor Loadings		Communality	
	Factor 1	Factor 2	Initial	Final
Murder	.578	.685	.609	.803
Rape	.852	.156	.730	.751
Robbery	.712	.086	.595	.514
Assault	.763	.329	.645	.691
Burglary	.908	-.218	.795	.871
Larceny	.753	-.474	.727	.793
Auto	.513	-.284	.486	.343
Eigen value	3.804	0.962		4.766

While in PCA it built PCs consisting linear combinations of observed variables, in ML as an EFA method 2- factor model for each variable can be written as follows illustrating the different approaches of two methods.

Murder $= 0.578\ F1 + 0.685\ F2 + U_1$

Rape $= 0.852\ F1 + 0.156\ F2 + U_2$

Robbery $= 0.712\ F1 + 0.086\ F2 + U_3$

Assault $= 0.763\ F1 + 0.329\ F2 + U_4$

Burglary $= 0.908\ F1 - 0.218\ F2 + U_5$

Larceny $= 0.753\ F1 - 0.474\ F2 + U_6$

Auto $= 0.513\ F1 - 0.284\ F2 + U_7$

Where U_i ($i = 1,2,\ldots,7$) is the unique variance of each observed variable.

Considering the factor loadings for each variable in table 3.3.2, new variances explained by the selected two factors that have been extracted from ML method is shown below. But note that these are not the exact factor models for the observed variables as the factor rotation yet to be done. Therefore with the purpose of calculating the new eigen values (variances of factors after extraction or the sums of squared loadings on Factors), consider the general formulas for factor model of each variable.

Contribution of F1 explaining sums of squared observed variance (eigen value) :

$= (0.578)^2 + (0.852)^2 + (0.712)^2 + (0.763)^2 + (0.908)^2 + (0.753)^2 + (0.513)^2$

$= 3.804$

Contribution of F2 explaining sums of squared observed variance (eigen value) :

$= (0.685)^2 + (0.156)^2 + (0.086)^2 + (0.329)^2 + (-0.218)^2 + (-0.474)^2 + (-0.284)^2$

$= 0.962$

The percentage variances accounted for factor1 and factor2 are 54.34% $\left(\frac{3.804}{7} * 100\right)$ and 13.75% $\left(\frac{0.962}{7} * 100\right)$ respectively. It should be noted that the models illustrate the variances which explained by the unique factors are not accounted by any of these two factors. Although the computation is same as in PCA, in ML the factor variance with the extraction is different since the values are based only on common or shared variance. As a result the total variance accounted for 2 factors under ML is less than that for 2 factors under PCA as showed in table 3.3.1.

The initial communalities were taken as the squared multiple correlations (SMC) of the variable with the other variables in contrast initial communalities are taken as one in Principal Component Analysis (table 3.2.2, pg 33). Calculating the communalities after extraction is more appropriate in EFA methods. We can interpret the proportion of each variable's variance which is explained by the retained factors according to the factor models built. For example communality for Murder is calculated as the sum of squared loadings on each factor ($0.578^2 + 0.685^2$) resulting 0.803, indicating a higher representation of that variable in factor space. Further for all the manifest variables final communalities extracted using ML is smaller than those extracted using PCA except in larceny in Burglary. The difference is highly substantial for auto (0.671 and 0.343) and followed by Robbery (0.650 and 0.514) and Assault (0.794 and 0.691). But the extracted communalities in ML itself has a higher value except in Auto (0.343) and Robbery (0.514), implying that the extraction is well defined the model.

The Chi- Square goodness- of- fit test ($X_8^2 = 18.784$, $p = 0.016$) of ML factoring indicates that there is no systematic variance in the reduced model after two factors been extracted confirming 2 – factor model is sufficient.

Results in table 3.3.2 also indicate the necessity of conducting rotations on factor loadings in order to a clear interpretation. As it is shown in table 3.3.2, the variables number of rapes (Rape), number of robberies (Robbery) , number of assaults (Assault), number of burglaries (Burglary) and number of larcenies (Larceny) are highly loaded on factor 1 having loadings greater than 0.7. Only for number of murders (Murder) it has a slightly high loading (0.685) on factor 2. Therefore it is clear that it is not possible to classify meaningful factors without applying a rotation method.

The final factor loadings after applying orthogonal rotations namely Varimax, Equamax and Quartimax for 2 factor model can be shown in section 3.3.3.

3.3.2 Use of Orthogonal rotations

Factor loadings for 2- factor model using ML with Varimax, Equamax and Quartimax rotations are shown below.

Table 3.3.4 : Factor loadings for 2 –factor model using ML with Varimax and Equamax rotations

Manifest Variable	Rotation method			
	Varimax		Equamax	
	Factor 1	Factor 2	Factor 1	Factor 2
Murder	-.003	.896	-.003	.896
Rape	.548	.671	.548	.671
Robbery	.486	.526	.486	.526
Assault	.369	.745	.369	.745
Burglary	.833	.422	.833	.422
Larceny	.881	.127	.881	.127
Auto	.574	.116	.574	.116
Variability	2.47	2.29	2.47	2.29
Total	4.76		4.76	

As it resulted in PCA (table 3.2.4, pg35), in ML also the variance accounted by each variable separately were almost same for both methods. Further those rotations have lead for a meaningful interpretation of selected two factors. The variance accounted for by the two factors remains the same (4.76; 68.09%). According to Varimax and Equamax rotations when selecting the variables with greater than 0.5 loadings , number of murders (0.896), number of rapes (0.671), number of robberies (0.526) and number of assaults (0.745) have loaded highly on Factor 2, while number of burglaries (0.833), number of larcenies (0.881) and number of autos (0.574) loaded on Factor 1.

Table 3.3.5 : Factor loadings for 2 –factor model using ML with Quartimax rotation

Manifest Variables	Quartimax Rotation	
	Factor 1	Factor 2
Murder	.320	.837
Rape	.753	.429
Robbery	.643	.316
Assault	.612	.562
Burglary	.929	.093
Larceny	.868	-.200
Auto	.578	-.099
Variability	3.42	1.36
Total	4.78	

In contrast with Varimax and Equamax rotations the variance encountered by each variable has been maximized except in number of robberies (0.712 and 0.643) and number of assaults (0.763 and 0.612). Thus it is still difficult to interpret the results after the Quartimax rotation Further the variance accounted for by the two factors has increased to 68.28%.

Therefore using the Quartimax rotation in ML factoring has become ineffective.

3.3.3 Summary of Maximum Likelihood factoring

The total variance of the initial system has been reduced to 68.1% (< 76.48%) after extracting the factors. According to the final communalities, the selected two factors are explaining a higher proportion of variance (>0.6) of each of the variables except in number of autos (0.343) and number of robbery (0.514). Number of burglaries (0.871) is the highest explaining variable by the two factors. Factors extracted using ML were rotated using the three Orthogonal rotation methods. Both varimax and equamax rotations given the same result with a clear interpretation of two factors while Quartimax rotation was not effective for rotation with ML factoring. Thus the 2 factor models for each variable can be written as follows.

$$\text{Murder} = -0.003\,F1 + 0.896\,F2 + U_1$$

$$\text{Rape} = 0.548\,F1 + 0.671\,F2 + U_2$$

$$\text{Robbery} = 0.486\,F1 + 0.526\,F2 + U_3$$

$$\text{Assault} = 0.369\,F1 + 0.745\,F2 + U_4$$

$$\text{Burglary} = 0.833\,F1 + 0.422\,F2 + U_5$$

$$\text{Larceny} = 0.881\,F1 + 0.127\,F2 + U_6$$

$$\text{Auto} = 0.574\,F1 + 0.116F2 + U_7$$

Where U_i ($i = 1,2,\ldots,7$) is the unique variance of each observed variable.

3.4 Factor Extraction using Principal Axis Factoring (PAF)

3.4.1 Preliminary analysis of PAF

Results of preliminary analysis of PAF is shown in table 3.4.1

Table 3.4.1 Results of the eigen values of the correlation matrix using PAF

Factor	Initial Eigen values			Extraction Sums of Squared		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.115	58.79	58.79	3.822	54.601	54.601
2	1.239	17.70	76.48	0.955	13.639	68.240

The variance accounted for by the remaining two factors is less than that of the PCA (76.5%) but approximately same with ML factoring (68.09%). Principal Axis factoring also accounts only the common variance as it described in ML factoring.

Summary results of factor extraction using Principal Axis factoring is shown in table 3.4.2 with the factor loadings before rotation and communalities of manifest variables.

Table 3.4.2 : Factor loading and communality estimates for 2 factor model fitted to data by PAF

Manifest Variable	Factor Loadings		Communality	
	Factor 1	Factor 2	Initial	Final
Murder	.616	.688	.609	.852
Rape	.855	.123	.730	.746
Robbery	.742	-.026	.595	.551
Assault	.774	.287	.645	.681
Burglary	.901	-.282	.795	.891
Larceny	.689	-.429	.727	.658
Auto	.526	-.347	.486	.397
Eigen value	3.82	0.955		4.77

Factor loadings for each variable are approximately same with the factor loadings in ML before rotation. The method of calculating the variances of factors after the extraction is discussed under the ML in table 3.3.2. Therefore the percentage variances accounted for factor 1 and factor 2 are 54.6% $\left(\frac{3.822}{7} * 100\right)$ and 13.64% $\left(\frac{0.955}{7} * 100\right)$ respectively. Thus the 2-factor models for each variable in PAF also can be written with this general formula of Exploratory Factor Analysis.

$$X_i = b_{i1} F_1 + b_{i2} F_2 + U_i$$

Where b_{ij} is the factor loading on i^{th} variable from j^{th} factor. However when comparing these factor loadings it is not possible to isolate some variables which load more on the first factor than the other variables.

Furthermore notice that the initial communalities of both PAF and ML are same values. Hence in PAF also the initial values of the diagonal of the correlation matrix are changing by inserting the values of squared multiple correlation (SMC) of the variable with the other variables. The final communalities of each manifest variable computed by PAF is approximately equal with the ML. When comparing the final communalities of PCA (table 3.2.2, pg 33) , it is shown that all the communalities in PAF is higher than that of the PAF except in number of burglaries (Burglary). However the remaining two factors explaining a higher proportion of variance (> 0.65) of each manifest variable except in number of autos (0.397) and number of robberies (0.551) indicating that the extraction is well defined the model.

Since the factor loadings should be improved for a clear interpretation , the 2 common factors are transformed by applying orthogonal rotations namely Varimax, Equamax and Quartimax.

3.4.2 Use of Orthogonal Rotations

Factor loadings for 2-factor model using PAF with Varimax , Equamax and Quartimax rotations are shown below.

Table 3.4.3: Factor loadings for 2-factor model using PAF with Varimax and Equamax rotations

Manifest Variable	Rotation method			
	Varimax		Equamax	
	Factor 1	Factor 2	Factor 1	Factor 2
Murder	.007	.923	.007	.923
Rape	.560	.658	.560	.658
Robbery	.574	.471	.574	.471
Assault	.391	.727	.391	.727
Burglary	.862	.385	.862	.385
Larceny	.800	.134	.800	.134
Auto	.624	.088	.624	.088
Variability	2.57	2.21	2.57	2.57
Total	4.78		4.78	

As so far in PCA and ML , in PAF also there is no any significant difference between the factor loadings in Varimax and Equamax rotations. The variance accounted for by the selected two factors remaining approximately the same (4.78, 68.28%). According to the rotated loadings when consider the loadings greater than 0.5, number of murders (0.923) , number of rapes (0.658) and number of assaults (0.727) loaded highly on factor 2. Further number of robberies (0.574), number of burglaries (0.862) , number of larcenies (0.800) and number of autos(0.624) have loaded highly on factor 1.

Table 3.4.4 : Factor loadings for 2-factor model using PAF with Quartimax rotation

Manifest Variables	Quartimax Rotation	
	Factor 1	Factor 2
Murder	.315	.868
Rape	.748	.433
Robbery	.698	.252
Assault	.611	.554
Burglary	.941	.074
Larceny	.799	-.141
Auto	.617	-.126
Variability	3.42	1.35
Total	4.77	

The factor loading of each variables after the Quartimax rotation, has maximized except in number of robberies (0.698) and number of assaults (0.611). Therefore it is not easy to interpret the 2 factor model. As also resulted in ML, Quartimax rotation is not effective for the factor rotation followed by PAF.

3.4.3 Summary of Principal Axis Factoring

Principal Axis factoring accounts the common variance of the variables for the extraction. Therefore the cumulative variance explained by the two extracted factors is lower than that of Principal Component Analysis. But it is slightly higher than that of ML, indicating that PAF has extracted high amount of common variance of observed variables than in Maximum Likelihood (68.24% and 68.09% respectively). The proportion of variance that is explaining by the remaining two factors are high in the variable types of number of burglaries (0.891) , number of murders (0.852) and number of rapes (0.746).

The factors extracted using PAF were rotated using Varimax, Equamax and Quartimax rotations. Similar with ML factoring, Varimax and Equamax rotations given meaningful factor loadings over the two factors. Further Quartimax rotation was not relevant for the rotation. However, each observed variable can be written as a linear combination of the resulted factors and with its' unique factor with the factor loadings followed by Varimax or Equamax rotation as follows.

$$\text{Murder} = 0.007\,F1 + 0.923\,F2 + U_1$$

$$\text{Rape} = 0.560\,F1 + 0.658\,F2 + U_2$$

$$\text{Robbery} = 0.574\,F1 + 0.471\,F2 + U_3$$

$$\text{Assault} = 0.391\,F1 + 0.727\,F2 + U_4$$

$$\text{Burglary} = 0.862\,F1 + 0.385\,F2 + U_5$$

$$\text{Larceny} = 0.800\,F1 + 0.134\,F2 + U_6$$

$$\text{Auto} = 0.624\,F1 + 0.088\,F2 + U_7$$

Where U_i ($i = 1,2,\ldots,7$) is the unique variance of each observed variable.

3.5 Factor Extraction using Generalized Least Squares (GLS)

3.5.1 Preliminary analysis of GLS

Table 3.5.1: Results of the Eigen values of the correlation matrix using GLS

Factor	Initial Eigen values			Extraction Sums of Squared		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.115	58.79	58.79	3.577	51.097	51.097
2	1.239	17.70	76.48	1.237	17.668	68.765

According to the Kaiser criterion two factors were extracted. The factors have extracted based on the common variance of the observed variables. Cumulative percentage of variance of extracted two factors (68.77%) is larger than that of occurred in Maximum Likelihood Factoring (68.09%) and Principle Axis Factoring (68.24%).

The results of communalities has encountered a "Heywood case" meaning one or more communalities have become greater than 1 in that specific iteration. As sited in http://www.sfu.ca/sasdoc/sashtml/stat/chap26/sect21.htm , there are several possible causes for this "Heywood case". Among those following are special.

- Bad prior communality estimates
- Too many common factors
- Too few common factors
- Not enough data to provide stable estimates
- The common factor model is not an appropriate model for the data

When the Heywood case occurred, SPSS will display the communalities based on the solution from the previous iteration. Therefore with this different situation, it is important to interpret the results carefully. The following table 3.5.2 shows the communalities and the factor loadings for 2- factors extracted.

Table 3.5.2 : Factor Loading and communality estimates for 2 factor model fitted to data by GLS

Manifest Variable	Factor Loadings		Communality	
	Factor 1	Factor 2	Initial	Final
Murder	.390	.768	.609	.772
Rape	.801	.359	.730	.796
Robbery	.657	.337	.595	.690
Assault	.650	.531	.645	.725
Burglary	.908	.051	.795	.845
Larceny	.901	-.342	.727	.937
Auto	.541	-.046	.486	.630
Eigen value	3.58	1.26		5.395

Same as in the previous Exploratory Factor extraction methods, the initial communality is the squared multiple correlation (SMC) of the variable with the other variables. The communalities after extraction is lower than that of in PCA since GLS also form in common variance of observed variables. The proportion of variance explained by the extracted two factors are high in Larceny (0.937), Burglary (0.845) and in Rape (0.796).

3.5.2 Testing the adequacy of number of factors

Like in Maximum Likelihood Factoring, Generalized Least Squares also generates a chi-square goodness-of-fit measure to test whether the extracted number of factors is adequate in representing the factor structure of the observed variables.

H_0 = There is a systematic variance of the reduced model after two factors extracted.

Vs

H_1= There is no systematic variance of the reduced model after two factors extracted.

Chi-Square Goodness-of-fit test (χ_8^2 = 13. 071, p = 0.109) indicates that at 0.05 ($\alpha = 0.05 < 0.109$) level of significance we have no evidence to reject the null hypothesis. Hence there is a systematic variance of the reduced model after two factors extracted .Therefore two factors are not enough to represent the model.

To implement the factor model the researcher can increase the number of factors one at a time until a satisfactory goodness of fit is obtained. Therefore the number of factors to be extracted was increased up to 3 and the results of goodness-of-fit test (χ_3^2 = 1.658, p = 0.646) indicated again the acceptance of null hypothesis. Further increase of factors lead to unable of computing the goodness-of-fit test as the degrees of freedom become negative. Therefore it is not possible to conduct Generalized Least Square method for this set of data.

3.5.3 Summary of Generalized Least Squares

Factors were extracted based on the common variance of the observed variables. However Heywood case occurred with the communalities in 2-factors model. Further Chi- Square Goodness-of-fit test (χ_8^2 = 13. 071, p = 0.109), confirmed that the number of factors are not enough to interpret the factor model. Although the number of factors was increased up to 3 and 4 respectively, each stage rejected the model implying that the Generalized Least Squares method is not appropriate to identify the underlying factors of this data set.

According to the descriptive statistics of the data it revealed that number of robberies and number of autos are deviate from the normal distribution. Therefore it is clear that the assumption of multivariate normality has violated with this set of data. Hence we can assume that the problems of GLS have occurred to this unsatisfactory condition. Therefore it is not appropriate to apply General Least Square method for the data which is not having multivariate normal distribution.

3.6 Summary of 2 Factor models Vs Method of Extraction

The following table shows the summary of the variances of selected two factors in each extraction method.

Table 3.6.1: Summary of 2 – Factors with the method of extraction

Method of extraction	Component/Factor	Eigen value	% Variance after extraction	Cumulative % of variance in the system
PCA	PC1	4.115	58.79	76.48
	PC2	1.239	17.70	
ML	Factor 1	3.804	54.34	68.09
	Factor 2	0.962	13.75	
PAF	Factor 1	3.822	54.60	68.24
	Factor 2	0.955	13.64	
GLS	Factor 1	3.577	2 –factor model was rejected	
	Factor 2	1.237		

CHAPTER 4
CONCLUSIONS AND RECOMMENDATIONS

Application of Principal Component Analysis (PCA) and Exploratory Factor Analysis (EFA) has been inextricably linked and thus much confusion has arisen to identify which one to be used. Once decided that EFA to be used, there are various types of factor extraction methods and factor rotation methods to make the factor loadings simpler and more meaningful. One method of component extraction is use of PCA technique. Almost all the text books on multivariate analysis are generally difficult to understand and do not clearly recommend to use a specific method on EFA for a given set of data. The study is therefore carried out to compare results of different extraction methods (including PCA) and factor rotations using set of real data.

If there is a large number of observed variables and wish to build a smaller number of components accounting the maximum variability of the initial system then it is appropriate to apply the Principal Component Analysis.

EFA consider only the common variance of the observed variables, since the common variance is due to the underlying factor structure of the observed variables. Hence contrasting with PCA, the ultimate goal of EFA is to explore those variables or factors which cannot be measured directly.

Of all factor extraction methods used in this study: Maximum Likelihood, Principal Axis Factoring and Generalized Least Square are reproduced the correlation matrix before the factor extraction. Only Maximum Likelihood and Generalized Least Square methods consider the assumption of multivariate normal distribution of the variables. For Principal Axis Factoring no distributional assumption is required. In the analyzed data set the variables types of number of robberies and number of autos were away from the normal distribution and consequently no Multivariate distribution in the observed variables. Although in Maximum Likelihood factoring there was no any computational problem of factors under this situation, Generalized Least Squares method did not compute a significant factor structure.

In that case it is recommended to use Principal Axis Factoring irrespective of normal distributions of each observed variable while Generalized Least Squares is only suitable if there is a multivariate distribution of the original system.

Application of rotation methods in PCA did not differentiate the rotated factor loadings in three orthogonal methods, Varimax, Equamax and Quartimax hence Orthogonal rotations can be applied for PCA regardless the method at the time needed a meaningful interpretation.

In contrast, Quartimax rotation is not appropriate for any of these considered Exploratory Factor extraction method while Varimax rotation gave a comparatively clear interpretable outcome. Therefore it is recommended to use Varimax rotation method followed by any of these extraction methods of Maximum Likelihood factoring , Generalized Least Square and Principal Axis Factoring.

It should be noted that the above recommendations were based on the data set used. However, the results obtained were confirmed using jackknife validation. Similar results were obtained when outliers of the data set were removed.

It should be noted the researchers have hardly used Un weighted Least Squares, Image Factoring and Alpha Factoring, may be due to lack of theoretical information.

REFERRENCE LIST

ABDI H, Factor Rotations in Factor Analyses. University of Texas at Dallas,[Online]. Available : http://www.utdallas.edu/~herve/Abdi-rotations-pretty.pdf

Annotated SPSS Output: Factor Analysis [Online]. Available: http://www.ats.ucla.edu/stat/spss/output/factor1.htm

BROWN TA 2006, Factor Extraction, *Confirmatory Factor Analysis for Applied Research.* pp. 21-22. [Online]. Available: http://books.google.lk/books?id=KZwDkH2G2PMC&pg=PA21&lpg=PA21&dq=Methods+of+Factor+Extraction&source=bl&ots=ExPNN50STX&sig=fzpVIn2rO_w5oKVxcrhrKb51NKM&hl=en&ei=kFPrTMfMJ4G-uwP3ubiDAg&sa=X&oi=book_result&ct=result&resnum=9&ved=0CFEQ6AEwCDgK#v=onepage&q=Methods%20of%20Factor%20Extraction&f=false

COSTEELLO AB & OSBORNE JW July 2005, Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your analysis, *Practical Assessment, Research & Evolution,* vol.10, pp.1-3.

COSTA JD 1998, Overview of Factor Analysis, Department of Psychology, University of Alabama, pp. 1-4.

COOPER LG, TUCKER LR, & MEREDITH W July 1972, *Obtaining Squared Multiple Correlations from a correlation matrix which may be singular,* 37, pp.143-144

eNOTES. *Factor Analysis* [Online]. Available: http://www.enotes.com/factor-analysis-reference/factor-analysis

Factor Analysis,: Statnotes, from North Carolina State University, Public Administration program.[Online]. Available: http://www.encorewiki.org/download/attachments/25657/Factor+Analysis_+Statnotes+from+North+Carolina+State+University.pdf

FIELD Dr. A 2005, Factor Analysis using SPSS, *Discovering statistics using SPSS*, 2nd edition ed.
[This is a handout with abridged version of Chapter 15 of Field (2005)]

Glossary , Crawford Ferguson rotations, pp.662 [Online]. Available: *http://www.stata.com/bookstore/stata12/pdf/mv_glossary.pdf*

gopalprasadmalakar12 (Dec 18, 2009), *How to do Principal Component Analysis (PCA)*. [Online]. Available: http://www.youtube.com/watch?v=xr3whjZuOgY&feature=mfu_in_order&list=UL

HARMAN HH 1976, Brief History of Factor Analysis. *Modern Factor Analysis*. 3rd ed. pp. 3-6. [Online]. Available: http://books.google.lk/books?id=e-vMN68C3M4C&printsec=frontcover#v=onepage&q&f=false

Heywood Cases and Anomalies. [Online]. Available: http://www.sfu.ca/sasdoc/sashtml/stat/chap26/sect21.htm

JOE-ON KIM & MUELLER CW 1978, Introduction, *Factor Analysis Statistical Methods and Practical Issues*. pp. 7-14 . [Online]. Available: http://books.google.com/books?hl=en&lr=&id=raQzQnbET9QC&oi=fnd&pg=PA5&dq=Factor+extraction+methods+in+analysis&ots=PVrsGctpNl&sig=VlZl_yoBhWzoRC3sjqdbkvLh5tg#v=onepage&q=Factor%20extraction%20methods%20in%20analysis&f=false

LEANDRE RF, STRAHAN EJ, MCCALLUM RC & WEGENER DT 1999. *Evaluating the Use of Exploratory Factor Analysis in Phychological Research*. 4, pp 272.

LINGARD HC & ROWLINSON S. Sample size in factor analysis: why size matters.

Maximum Likelihood Factoring.[Online]. Available: http://psych.unl.edu/psycrs/statpage/mlfact.pdf

NIST/SEMATECH. Measures of Skewness and Kurtosis, *E-Handbook of Statistical Methods* [Online]. Available: http://itl.nist.gov/div898/handbook/eda/section3/eda35b.htm

Public Library Statistical Reports, *Guide to interpreting Box Plots*. [Online]. Available: http://www.plstatreports.com/PLQ26/BoxPlotGuide.htm

REILEY B, *Factor Analysis Rotation Methods*, eHOW [Online]. Available: http://www.ehow.com/list_7450638_factor-analysis-rotation-methods.html

SAS, Multivariate Analysis : Factor Analysis. *SAS(R) Stat Studio 3.11: User Guide* .[Online]. Available: http://support.sas.com/documentation/cdl/en/stsug/62259/HTML/default/viewer.htm#ugmultfactor.htm

SUHR DD, 'Exploratory or Confirmatory Factor Analysis?' : University of Northern Colorado, paper 200-31 , pp.1-4.[Online].Available: http://www2.sas.com/proceedings/sugi/200-31.pdf

SUHR DD, Principal Component Analysis Vs. Exploratory Factor Analysis: University of Northern Colorado, paper 203-30, pp.1-2. [Online]. Available: http://www2.sas.com/proceedings/sugi30/203-30.pdf

TRYFOS P 1997, Factor analysis, Chapter 14, 1-13 [Online]. Available: http://www.yorku.ca/ptryfos/f1400.pdf

UNIVERSITY EAST CAROLLINA, Review of Article on Use of Exploratory FactorAnalysis.Fabriger,L.R.,Wegener,D.T.,MacCallum,R.C.,&Strahan,E.J. (1999) Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, Vol.4, pp.272-299 [online]. Available: http://core.ecu.edu/psyc/wuenschk/stathelp/efa.htm

WIKIVERSITY, *Exploratory Factor analysis.* [Online]. Available: http://en.wikiversity.org/wiki/Exploratory_factor_analysis

WIKIPEDIA, *Factor analysis.* [Online]. Available: http://en.wikipedia.org/wiki/Factor_analysis

APPENDICES

Appendix 1.1 : Rate of different crimes per 100,000 in various states in USA

No	State	Murder	Rape	Robbery	Assault	Burglary	Larceny	Auto
1	ALABAMA	14.2	25.2	96.8	278.3	1135.5	1881.9	280.7
2	ALASKA	10.8	51.6	96.8	284	1331.7	3369.8	753.3
3	ARIZONA	9.5	34.2	138.2	312.3	2346.1	4467.4	439.5
4	ARKANSAS	8.8	27.6	83.2	203.4	972.6	1862.1	183.4
5	CALIFORN	11.5	49.4	287	358	2139.4	3499.8	663.5
6	COLORADO	6.3	42	170.7	292.9	1935.2	3903.2	477.1
7	CONNECTI	4.2	16.8	129.5	131.8	1346	2620.7	593.2
8	DELAWARE	6	24.9	157	194.2	1682.6	3678.4	467
9	FLORIDA	10.2	39.6	187.9	449.1	1859.9	3840.5	351.4
10	GEORGIA	11.7	31.1	140.5	256.5	1351.1	2170.2	297.9
11	HAWAII	7.2	25.5	128	64.1	1911.5	3920.4	489.4
12	IDAHO	5.5	19.4	39.6	172.5	1050.8	2599.6	237.6
13	ILLINOIS	9.9	21.8	211.3	209	1085	2828.5	528.6
14	INDIANA	7.4	26.5	123.2	153.5	1086.2	2498.7	377.4
15	IOWA	2.3	10.6	41.2	89.8	812.5	2685.1	219.9
16	KANSAS	6.6	22	100.7	180.5	1270.4	2739.3	244.3
17	KENTUCKY	10.1	19.1	81.1	123.3	872.2	1662.1	245.4
18	LOUISIAN	15.5	30.9	142.9	335.5	1165.5	2469.9	337.7
19	MAINE	2.4	13.5	38.7	170	1253.1	2350.7	246.9
20	MARYLAND	8	34.8	292.1	358.9	1400	3177.7	428.5
21	MASSACHU	3.1	20.8	169.1	231.6	1532.2	2311.3	1140.1
22	MICHIGAN	9.3	38.9	261.9	274.6	1522.7	3159	545.5
23	MINNESOT	2.7	19.5	85.9	85.8	1134.7	2559.3	343.1
24	MISSISSI	14.3	19.6	65.7	189.1	915.6	1239.9	144.4
25	MISSOURI	9.6	28.3	189	233.5	1318.3	2424.2	378.4
26	MONTANA	5.4	16.7	39.2	156.8	804.9	2773.2	309.2
27	NEBRASKA	3.9	18.1	64.7	112.7	760	2316.1	249.1
28	NEVADA	15.8	49.1	323.1	355	2453.1	4212.6	559.2
29	NEWHAMPS	3.2	10.7	23.2	76	1041.7	2343.9	293.4
30	NEWJERSE	5.6	21	180.4	185.1	1435.8	2774.5	511.5

No	State	Murder	Rape	Robbery	Assault	Burglary	Larceny	Auto
31	NEWMEXIC	8.8	39.1	109.6	343.4	1418.7	3008.6	259.5
32	NEWYORK	10.7	29.4	472.6	319.1	1728	2782	745.8
33	NORTHCAR	10.6	17	61.3	318.3	1154.1	2037.8	192.1
34	NORTHDAK	0.9	9	13.3	43.8	446.1	1843	144.7
35	OHIO	7.8	27.3	190.5	181.1	1216	2696.8	400.4
36	OKLAHOMA	8.6	29.2	73.8	205	1288.2	2228.1	326.8
37	OREGON	4.9	39.9	124.1	286.9	1636.4	3506.1	388.9
38	PENNSYLV	5.6	19	130.3	128	877.5	1624.1	333.2
39	RHODEISL	3.6	10.5	86.5	201	1489.5	2844.1	791.4
40	SOUTHCAR	11.9	33	105.9	485.3	1613.6	2342.4	245.1
41	SOUTHDAK	2	13.5	17.9	155.7	570.5	1704.4	147.5
42	TENNESSE	10.1	29.7	145.8	203.9	1259.7	1776.5	314
43	TEXAS	13.3	33.8	152.4	208.2	1603.1	2988.7	397.6
44	UTAH	3.5	20.3	68.8	147.3	1171.6	3004.6	334.5
45	VERMONT	1.4	15.9	30.8	101.2	1348.2	2201	265.2
46	VIRGINIA	9	23.3	92.1	165.7	986.2	2521.2	226.7
47	WASHINGT	4.3	39.6	106.2	224.8	1605.6	3386.9	360.3
48	WESTVIRG	6	13.2	42.2	90.9	597.4	1341.7	163.3
49	WISCONSI	2.8	12.9	52.2	63.7	846.9	2614.2	220.7
50	WYOMING	5.4	21.9	39.7	173.9	811.6	2772.2	282

Appendix 1.2 : Computed tables in Jackknife validation

Turn 1: With 49 observations holding 1st observation

Component	Initial Eigen values			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.076	58.228	58.228	4.165	59.501	59.501
2	1.262	18.022	76.250	1.185	16.933	76.435
3	0.730	10.424	86.674			
4	0.321	4.586	91.261			
5	0.267	3.817	95.077			
6	0.228	3.252	98.329			
7	0.117	1.671	100.000			

Turn 2 : With 49 Observations holding 2nd observation

Component	Initial Eigen values			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.002	57.171	57.171	4.142	59.165	59.165
2	1.283	18.333	75.505	1.284	18.345	77.510
3	0.753	10.762	86.266			
4	0.330	4.712	90.979			
5	0.273	3.896	94.875			
6	0.235	3.364	98.239			
7	0.123	1.761	100.000			

Turn 3 : With 49 Observations holding 3rd observation

Component	Initial Eigen values			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.039	57.700	57.700	4.121	58.876	58.876
2	1.287	18.381	76.080	1.264	18.062	76.938
3	0.722	10.308	86.388			
4	0.330	4.718	91.106			
5	0.274	3.921	95.027			
6	0.227	3.248	98.275			
7	0.121	1.725	100.000			

Turn 4: With 49 Observations holding 4th observation

Component	Initial Eigen values			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.039	57.700	57.700	4.038	57.692	57.692
2	1.287	18.381	76.080	1.180	16.853	74.545
3	0.722	10.308	86.388			
4	0.330	4.718	91.106			
5	0.274	3.921	95.027			
6	0.227	3.248	98.275			
7	0.121	1.725	100.000			

Turn 5 : With 49 observations holding 50^{th} observation

Component	Initial Eigen values			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.114	58.767	58.767	4.114	58.767	58.767
2	1.245	17.787	76.554	1.245	17.787	76.554
3	0.725	10.355	86.909			
4	0.318	4.545	91.454			
5	0.260	3.717	95.171			
6	0.221	3.160	98.331			
7	0.117	1.669	100.000			