ISSN 2348-3156 (Print)
International Journal of Social Science and Humanities Research ISSN 2348-3164 (online)
Vol. 4, Issue 1, pp: (714-723), Month: January - March 2016, Available at: www.researchpublish.com

Applying Principal Component Analysis for Qualification and Quantification of the Contribution of Key Foreign Exchange Earners in Rwanda (1999-2013)

[1]Muhire Didier, [2]Dr. Mung'atu K. Joseph, [3]Dr. Ndengo Marcel

[1,2]Jomo Kenyatta University of Agriculture and Technology
[3]University of Rwanda

Abstract: **The purpose of this this research project is to apply principal component analysis technique for qualification and quantification of the contribution of key foreign exchange earners in Rwanda (1999-2013). This research project provides details of principal component analysis method that helps researcher to restructure data sets specifically by reducing the number of variables. Such an approach is often called "Principal Component Analysis" or sometimes a "data reduction" or "dimensionality reduction" technique. This means that we start off with a set of variables, let say twenty; these appear to be too many variables to deal with, which in practice renders any economic forecasting alongside statistical analysis difficult if not accurate. Therefore, applying principal component analysis, this research aims at transforming a set of corrected variables into a set of uncorrected variables without any loss of information. By the end of the process we have measured the proportion of variance attributed to each eigenvalue produced associated with corresponding principal component and then interpret the PCs and finally predicted the scores. As Rwanda exports many products to international markets, it is important to qualify and quantify the contribution of each product to the reserves of foreign exchanges of Rwanda; which was the purpose of research project. On the basis of the research results, it was concluded that some products (cassiterite, coltan, wolframite and coffee) have significantly contributed on the reserves of foreign exchanges in Rwanda. Hence, it is for the Government, Private Sector or Stakeholders to allocate more funds to these to operate and produce more.**

Keywords: **Foreign Exchange Earners, Rwanda exports, Principal Component Analysis.**

1. INTRODUCTION

1.1. Back ground of the study:

The Government of Rwanda and also the financial institutions in Rwanda intervene in the foreign exchange market, among reasons, in order to defend the exchange rate and to achieve a desired amount of international reserves. The intervention in the foreign exchange market directly affects reserve money and hence has a direct impact on overall liquidity in the economy and the stance of monetary policy. In accordance with the current monetary policy framework, various mechanisms are at National Bank of Rwanda disposal to achieve the targets set in the monetary programme. Today Rwanda's foreign currency reserves innovation is a key factor in developing any sector, so the reserve of foreign exchange earners is of a big importance. Rwanda has long relied on coffee and tea as cash crops whereas major export markets include East African Community, China, Germany, and the United States. In statistics, it is often the case that researchers have a big data set to analyze and the dimension of this data set is often a huge problem when the researchers want to analyze that data set in terms of the relationship among the individual points in the data set. In order to examine the relationships among a set of p-correlated variables, it may be useful to transform the original set of variables to a new set of uncorrelated variables called "principal component." The transformation is in fact an orthogonal rotation in p-space. Assume that we have a set of data represented in terms of matrix, , where n columns are observations and m rows are the variables. We wish to transform linearly the matrix X into a Y matrix, also of dimension so that for some matrix P:

$$Y = PX \tag{1}$$

Where;

$$PX = \begin{pmatrix} Px_1 & Px_2 & . & . & Px_n \end{pmatrix} = \begin{pmatrix} p_1x_1 & p_1x_2 & \cdots & p_1x_n \\ p_2x_1 & p_2x_2 & \cdots & p_2x_n \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ p_mx_1 & p_mx_2 & \cdots & p_mx_n \end{pmatrix} = Y \tag{2}$$

With $p_i x_j \in IR^m$, $p_i x_i$ - is a standard Euclidean inner product and X - the original data being projected to the columns of P .

1.2. General objective:

The general objective of this research is to apply principal component analysis to qualify and quantify the contribution of key foreign exchange earners on Rwanda's reserves of foreign exchanges.

1.3. Specific objectives:

1. To transform a set of correlated variables contributing to foreign exchange earners into a new set of uncorrelated variables without loss of information.

2. To measure the proportion of variance attributed to each eigenvalue produced associated with corresponding principal component.

3. To interpret the PCs and their weighting to as to qualify and qualify the contribution of key foreign exchange earners in Rwanda on Rwanda's reserves of exchange.

4. To predict the component scores.

1.4 Hypotheses:

1. There is no correlation among many independent variables responsible for the foreign exchange earners.

2. There is a no proportion of variance attributed to each eigenvalue produced.

3. There is no contribution of key foreign exchange earners in Rwanda on Rwanda's reserves of exchange.

4. There are no component scores to predict.

Hence, this research project will be carried on coffee, tea, pyrethrum, cassiterite, coltan and wolframite exports in Rwanda since 1999 to 2013 to help in data analysis by the use of principal component analysis techniques. The next chapter introduces the literature review.

2. LITERATURE REVIEW

2.1. Review of previous studies on the subject of study:

2.1.1. Use of principal component analysis:

The history of statistics shows that the method principal components analysis is a multivariate statistical method almost used by all scientific disciplines introduced by Pearson in 1901.

(Hervé Abdi and Lynne J. Williams, 2010) described Principal Component Analysis to be a multivariate technique used for data sets analyses in which observations are described by several inter-correlated quantitative dependent variables with goal to extract the important information from the data set and also to represent it as a set of new orthogonal variables called "principal components", and also to display the pattern of similarity of the observations and of the variables.

(Frank Wood, 2009) argues that the central idea of principal component analysis (PCA) is the reduction of dimensionalities of data sets consisting of a large number of interrelated variables, while retaining as much as possible of the variation present in the data set. This result/output is achieved by transforming to a new set of variables, the principal components (PCs), uncorrelated, and which are ordered in descendent order so that the first retain most of the variation present in all of the original variables.

(Mark Richardson, 2009) argues that:" PCA is a general name of a technique that uses underlying mathematical principals to transform a number of possibly correlated variables called principal components. Its origin lies in multivariate data analysis; however, it has a wide range of other applications.

2.1.2. Key foreign exchange earners in Rwanda:

Exports are goods and services produced in one country and sent to another country for consumption. Minerals, coffee and tea are Rwanda's main sources of foreign exchange. Crops grown in the country include coffee, tea, pyrethrum, bananas, beans, sorghum and potatoes but coffee and tea are the major cash crops for export. Minerals mined include coltan, cassiterite, wolframite, sapphires, and gold but the most exported are coltan, cassiterite and wolframite.

2.2. Conceptual framework:

This research proposal focuses on the transformations of a set of correlated variables $X_1, X_2,..., X_n)$ into a new uncorrelated set of data $(Y_1, Y_2,..., Y_p)$ with reduced dimensionalities to measure the contribution of key foreign exchange earners in Rwanda since 1999 to 2013.

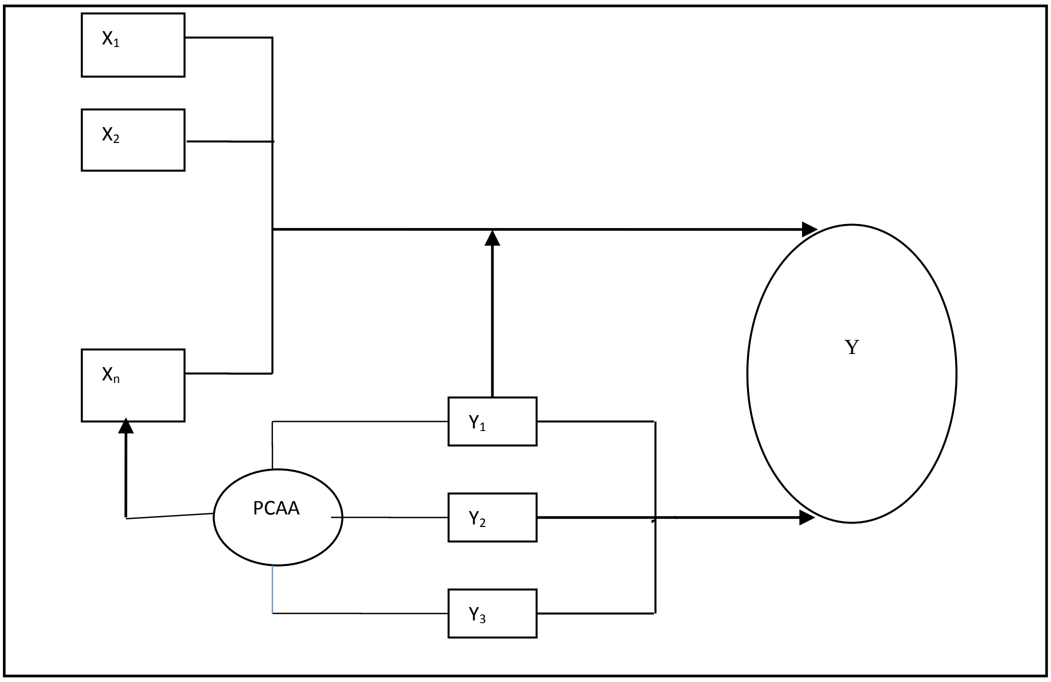

Figure 1: Conceptual frame work

This research project will deal with productions of key foreign exchange earners in Rwanda since 1999 to 2013. The sets of data to use in this work will come from two sources, namely the Central Bank of Rwanda (BNR) and the National Agriculture Export Development Board (NAEB). Secondary data covering 15 years: 1999 to 2013 will be used. Prior to this period, some data are characterized by inconsistency and missing data as well so a sample size will be taken. STATA software will be used in data analysis. The next chapter introduces the methodology.

3. METHODOLOGY

3.1. Introduction:

In this chapter, a brief but complete presentation of PCA method is done. This is followed by a brief discussion of PCA post estimation issues.

3.2. Principal component analysis:

Let $X = [X_1 \quad X_2 \quad . \quad . \quad . \quad X_p]$ a p-variate. Principal components are then defined as linear combinations of the p random variables $X_1 \quad X_2 \quad . \quad . \quad . \quad X_p$ in such a way that the covariance among then is zero. Several textbooks including references (1), (2), (5), (6) and (11) describe the PCA. Below is a brief description of the method.

Let $X = [X_1 \quad X_2 \quad . \quad . \quad . \quad X_p]$ be a random vector having the covariance matrix C with eigenvalues $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_p \geq 0$.

One considers the linear combinations

$$
\begin{aligned}
Y_1 &= a_1'^{'} X = a_{11}X_1 + a_{12}X_2 + ... + a_{1p}X_p \\
Y_2 &= a_2' \ X = a_{21}X_1 + a_{22}X_2 + ... + a_{2p}X_p \\
&. \\
&. \\
&. \\
Y_p &= a_p'^{'} X = a_{p1}X_1 + a_{p2}X_2 + ... + a_{pp}X_p
\end{aligned} \tag{3}
$$

In matrix form

$$
\begin{pmatrix} Y_1 \\ . \\ . \\ . \\ Y_p \end{pmatrix} = \begin{pmatrix} a_{11} & . & . & . & a_{1p} \\ a_{21} & . & . & . & a_{2p} \\ . & . & . & . & . \\ . & . & . & . & . \\ a_{p1} & . & . & . & a_{pp} \end{pmatrix} \begin{pmatrix} X_1 \\ . \\ . \\ . \\ X_p \end{pmatrix} \Leftrightarrow Y = PX \tag{4}
$$

By definition, principal components are those uncorrelated linear combinations $Y_1, Y_2, ..., Y_p$ whose variances in (7) are as large as possible. Let $Y_1, ..., Y_n$ be PCs where the subscripts define the first till the i^{th} PC.

Then,

$Y_1 = a_1'X$ satisfying max $Var(a_1' X)$ s.t. $a_1' \ a_1 = 1$

$Y_2 = a_2'X$ satisfying max $Var(a_2' \ X)$ s.t. $a_2' \ a_2 = 1$ and $Cov(a_1' \ X, a_2' \ X) = 0$.

The i^{th} PC is

$Y_i = a_i'X$ satisfying max $Var(a_i' \ X)$ s.t. $a_1' \ a_1 = 1$ and $Cov(a_i' \ X, a_k' \ X) = 0$, $\forall k \prec 1$.

The proportion of the total variance due to the k^{th} principal component is therefore equal to:

$$
\frac{\lambda_k}{\lambda_1 + \lambda_2 + ... + \lambda_p}; k = 1,2,..., p \tag{10}
$$

This result shows that each component of the coefficient vector $e_i' = [e_{i1}, e_{i2}, ..., e_{ik}, ..., e_{ip}]$ also merits its inspection.

3.3. The number of principal components:

Principal component can be generally defined as a linear combination of optimally-weighted observed variables. In reality, the number of components extracted is equal to the number of variables being analyzed. It requires the researcher to decide how many of the components are meaningful for interpretation. In general, you expect that only the first few components will account for meaningful amounts of variance, and that the later components will tend to account for only trivial variance. (Frank W., 2008).

3.4. PCA and post estimation tools:

Let C be a covariance matrix and let A the correlation matrix. For C and A, it is possible to compute

$$C = \{diag(R)\}^{-1/2} R \{diag(R)\}^{-1/2} \quad (11)$$

$$A = \{diag(R)\}^{-1} R \{diag(R)\}^{-1} \quad (12)$$

Where

R is the inverse of the correlation matrix.

One can therefore; we define “Kaiser-Meyer-Olkin” (KMO) as:

$$KMO = \frac{\sum_s r_{ij}^2}{\sum_s (a_{ij}^2 + r_{ij}^2)} \quad (13)$$

Where

$s = (i, j; i \neq j)$; r_{ij} is the correlation of variables i and j and a_{ij} is the anti-image correlation. The variable-wise measure KMO_i is defined analogously as

$$KMO_i = \frac{\sum_s r_{ij}^2}{\sum_s (a_{ij}^2 + r_{ij}^2)} \quad (14)$$

Where $s = (j; i \neq j)$

4. RESULTS AND DISCUSSIONS

4.1. Introduction:

In this chapter, secondary data from the Central Bank of Rwanda (BNR) and the National Agriculture Export Development Board (NAEB) are examined, processed and results are analyzed using STATA. These data covers the period 1999 to 2013. The current period was chosen because of availability of data, consistency.

4.2. Initial variables:

Rwanda exports different product like cassiterite, coltan, wolframite, beryllium, gold, animal skins, flowers, baskets, tea, coffee, pyrethrum, iron ore, concentrated milk, beer, raw sugar etc. This research will use a couple of productions of different initial variables. These initial variables are: X_1: Coffee productions; X_2: Tea productions; X_3: Coltan productions; X_4: Cassiterite productions; X_5: Wolframite productions; and; X_6: Pyrethrum. These are initial values from which to reduce the dimensionalities to principal components for a new set of uncorrelated variables.

4.3. Results and discussion:

This research focuses on applying principal component analysis for qualification and quantification of the contribution of foreign key exchange earners in Rwanda (1999-2013).

There are so many exports being done by the country from different products. Below is a table of our data to be used.

Table 1: Data for analysis

Year	Cassiterite	Wolframite	Coltan	Coffee	Tea	Pyrethrum
1999	1.17	0.11	4.63	26.47	19.25	0.31
2000	0.95	0.27	11.35	22.52	22.65	0.00
2001	1.13	0.40	41.10	19.36	25.74	1.77
2002	1.41	0.52	13.95	14.65	22.18	1.05
2003	4.49	0.23	6.36	15.01	22.60	1.28
2004	15.88	0.40	12.99	32.23	22.64	0.65
2005	17.86	2.58	16.87	38.27	25.29	0.00
2006	15.87	9.53	11.17	54.04	32.61	1.92
2007	31.97	19.42	19.23	35.67	34.78	3.00
2008	41.92	13.18	36.24	46.91	44.95	0.38
2009	28.58	5.76	20.24	37.29	48.72	0.64
2010	42.21	7.10	18.48	56.08	55.71	1.41
2011	96.82	16.03	38.58	74.60	61.91	4.51
2012	52.90	26.26	56.91	60.89	65.72	9.71
2013	61.07	30.05	134.57	54.90	57.19	3.98

Table 2: Summary statistics

Variable	Observations	Mean	Std. Dev.	Min	Max
Cassiterite	15	27.61533	27.63221	.95	96.82
Wolframite	15	8.789333	10.08325	.11	30.05
Coltan	15	29.51133	32.57891	4.63	134.57
Coffee	15	39.25933	18.16978	14.65	74.6
Tea	15	37.46267	16.56739	19.25	65.7
Pyrethrum	15	2.040667	2.530114	0	9.71

This table above gives the summary statistics of the data. It gives the mean, standard deviation, minimum and maximum of the data used in analysis. From the mean, it is clear that for coffee, tea, cassiterite and coltan there is no big difference. This shows that the four variables contributed too much than wolframite and pyrethrum.

Table 3: Principal components/correlation

Number of obs = 15
Number of comp. = 6
Trace = 6
Rotation: (unrotated = principal)
Rho = 1.0000

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	4.51847	3.81501	0.7531	0.7531
Comp2	.703457	.243168	0.1172	0.8703
Comp3	.460289	.323613	0.0767	0.9470
Comp4	.136676	.0453006	0.0228	0.9698
Comp5	.0913758	.00164557	0.0152	0.9850
Comp6	.0897302	-	0.0150	1.0000

From the table above, we have 6 eigenvalues and eigenvalues are the variances of the principal components. Let now consider the proportion of total variance due to the k^{th} PC such that:

ISSN 2348-3156 (Print)
International Journal of Social Science and Humanities Research ISSN 2348-3164 (online)
Vol. 4, Issue 1, pp: (714-723), Month: January - March 2016, Available at: www.researchpublish.com

$$\frac{\lambda_k}{\lambda_1 + \lambda_2 + ... + \lambda_p}; k = 1,2,..., p$$

The table above shows the components and how much they explain to the variation in the data (their proportions). For example, the first component explains 75.31% of the variation in the data. The second one explains about 11.72 % of the variation in the data, etc. Cumulatively, both two first components explain 87.03% of variation in the data; the three first ones explain 94.70% and so on.

In general, the four first components explain 96.98% (about 97%) of the data and the remaining 2 explain 3.02% of the variation. This means that the four first components explain more than the two last components. These results also shows that the correlation among data and leads us to conclude that our data are correlated. Also, the columnwise sum of squares of the eigenvalues should equal to unit length i.e. $(4.51847^2 + 0.703457^2 + ... + 0.0897302^2) = 1$.

Table 4: Principal components (eigenvectors)

Variable	Comp1	Comp2	Comp3	Comp4	Comp5	Comp6
Cassiterite	0.4302	-0.3378	-0.2303	0.1200	0.2252	-0.7633
Wolframite	0.4356	0.2822	0.0441	-0.6522	-0.5295	-0.1515
Coltan	0.3454	0.6906	-0.4851	0.1984	0.3212	0.1613
Coffee	0.4122	-0.4867	-0.1224	-0.3785	0.3818	0.5378
Tea	0.4417	-0.2166	-0.0334	0.5955	-0.5687	0.2807
Pyrethrum	0.3754	0.2136	0.8329	0.1516	0.3107	-0.018

The above table shows all components. This table contains component loadings, which are the correlations between the variable and the component. Because these are correlations, possible values range from -1 to +1. Taking at the first principal component for example, it has positive loadings of roughly equal size on all variables. This can be interpreted as overall contribution on the Rwanda reserve of foreign exchanges. The second principal component has positive and negative loadings. This means that sometimes there is contribution on the Rwanda reserve of foreign exchanges and also sometimes a sector needs money for from the reserve it to operate.

Considering the linear combinations and also by considering the first principal components (as they explain about 97% of variability of the data) we have the following equations:

$$\begin{aligned} Y_1 &= (0.43 \times 4.51) + (-0.33 \times 0.70) + (-0.23 \times 0.46) + (0.12 \times 0.14) \\ Y_2 &= (0.44 \times 4.51) + (0.28 \times 0.70) + (0.04 \times 0.46) + (-0.65 \times 0.14) \\ Y_3 &= (0.34 \times 4.51) + (0.69 \times 0.70) + (-0.49 \times 0.46) + (0.20 \times 0.14) \\ Y_4 &= (0.41 \times 4.51) + (-0.49 \times 0.70) + (-0.12 \times 0.46) + (-0.38 \times 0.14) \end{aligned} \quad (18)$$

The first principal component is computed such that it accounts for the highest variance in the data set. The second principal component is calculated in the same way, with the condition that it is uncorrelated with the first principal component and that it accounts for the next highest variance. This continues until a total of p-principal components have been calculated, equal to the original number of variables.

Table 5: Explained variance by components

Components |	Eigenvalue	Proportion	SE_Prop	Cumulative	SE_Cum	Bias
Cassiterite	4.518471	0.7531	0.0785	0.7531	0.0785	0.111428
Wolframite	.7034571	0.1172	0.0498	0.8703	0.0447	0.058393
Coltan	.460289	0.0767	0.0335	0.9470	0.0183	-0.094946
Coffee	.1366764	0.0228	0.0103	0.9698	0.0113	0.00213
Tea	.0913758	0.0152	0.0069	0.9850	0.0068	0.292973
Pyrethrum	.0897302	0.0150	0.0068	1.0000	0.0000	-0.369977

In this case, an estimation command was run. The output is organized in such way that there are different equations. The first equation contains the eigenvalues while the second equation named “comp 1” contains is the first PC; for instance, standard errors of the eigenvalues. Therefore, the estimation command has also the standard errors of the principal components and it has estimated the covariances.

Scree plot:

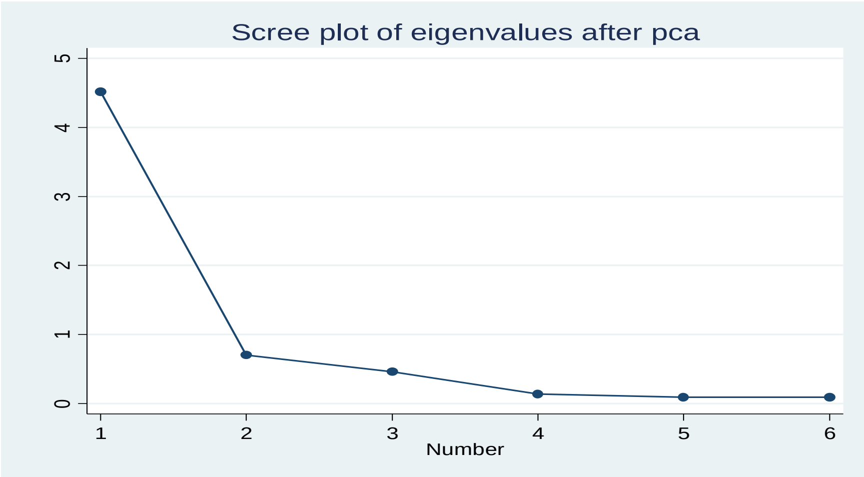

Figure 1: Scree plot

The scree plot graphs the eigenvalue against the component number. From the fourth component on, you can see that the line is almost flat, meaning that each successive component is accounting for smaller and smaller amounts of the total variance. In general, we are interested in keeping only those principal components whose eigenvalues are explaining at about 95 % of the total variation of data. Hence, the four first principal components are retained as they explain 96.98% (or about 97%) of variation of data while the remaining two components explain only about 3.02% (or 3%) of variation in data. Form this point of view; it is clear that the point of principal components analysis is to redistribute the variance in the correlation matrix to redistribute the variance to first components extracted.

Estimating the KMO (Kaiser-Meyer-Olkin) measure of sampling adequacy we get the following results.

Table 16: KMO

Variable |	kmo
Cassiterite	0.8514
Wolframite	0.7946
Coltan	0.7377
Coffee	0.8222
Tea	0.8796
Pyrethrum	0.8177
Overall	0.8214

The above results show that the KMO for each individual variable is greater than 0.5 or 50% even the overall KMO is 82.14%. This means that our data are of good choice i.e. our sample is adequate and no more or small correlations among the components. Therefore, this indicates that the component analysis is useful for the variables.

5. SUMMARY, CONCLUSION AND RECOMMENDATIONS

5.1. Summary:

This research project provides details of principal component analysis method that helps researchers to restructure data sets specifically by reducing the number of variables.

Therefore, by applying principal component analysis techniques, by the end of the process we came up to have a smaller number but which still reflects a large proportion of the information contained in the original dataset. It is in this line that by the use of it, that this research project helped to apply the technique for qualification and quantification of the contribution of key foreign exchange earners in Rwanda (1999-2013).

Applying the technique, we came up to reduce the dimensionality of our data to a new set of uncorrelated variables. Therefore, the first objective was achieved and also the first hypothesis was verified and accepted.

After the transformation, to each eigenvalue was associated a variance. This variance varied from to the first component to the last component, monotonically decreasing. Therefore, the second objective of this research project was achieved and also the second hypothesis was rejected.

Generated principal components showed how much they explain to the variability of data so we found that the four first components explain 96.98%. This result shows that cassiterite, coltan, wolframite and coffee contribute highly on the Rwanda reserves of foreign exchanges. The fourth objective was now explained and then the fourth hypothesis was rejected.

With prediction or post estimation of the components, the overall KMO coefficient of 82.14% showed that our results are adequate. Therefore, the fourth objective was explained and also the fourth hypothesis was rejected. Component scores can be predicted.

5.2. Conclusions:

In statistics, it is hard for researchers to analyze big set of data when the dimension of this data set is often a huge. It becomes a problem when the researchers want to analyze that data set in terms of the relationship among the individual points in the data set. In order to examine the relationships among a set of p-correlated variables, it may be useful to transform the original set of variables to a new set of uncorrelated variables called principal component. These new variables are linear combinations of the original variables and are derived in decreasing order of importance. For instance, the first principal component accounts for the highest variability in the original data.

Rwanda hosts a large number of products which are considered to be key foreign exchange earners. The Government of Rwanda in collaboration with financial institutions intervenes in the foreign exchange market, among reasons, in order to defend the exchange rate and to achieve a desired amount of international reserves. The intervention in the foreign exchange market directly affects reserve money and hence has a direct impact on overall liquidity in the economy and the stance of monetary policy. The interventions in the foreign exchange market directly affect reserve money and hence have direct impacts on overall liquidity in the economy and the stance of monetary policy. In accordance with the current monetary policy framework, various mechanisms are at disposal to achieve the targets set in the monetary programme. These mechanisms include open market operations, rediscount policy, changes in reserve requirement and foreign exchange intervention. Today Rwanda's foreign currency reserves innovation is a key factor in developing any sector, so the reserve of foreign exchange earners is of a big importance. This statistical research, from results, has showed that some sectors contribute more on reserves of foreign exchanges than others. These currencies are from coltan, cassiterite, coffee and tea. It requires then for the Government of Rwanda in collaboration with other institutions and also stakeholders to invest more in those sectors as their exports contribute too much on reserves of foreign exchanges. For these, it requires then modern techniques to help to dig those kinds of minerals and also modern techniques for coffee and tea growing. Investment also in skilled engineers, modern materials, and land for these activities is need. It is for the Government of Rwanda to keep working on policies and regulations to help these sectors to grow.

REFERENCES

[1] Alexander Ilin and Tapani Raiko (2010). Practical Approach to Principal Component Analysis in the Presence of Missing Values. Finland: Aalto University School of Science and Technology.

[2] An Gie Yong & Sean Peace (2013). A Beginner's Guide to Factor Analysis (Vol. 9). Japan: University of Ottawa.

[3] Frank Wood (2008). Principal component analysis.

[4] Hervé Abdi and Lynne J. Williams (2010). Principal component analysis (Volume 2). The University of Texas.

[5] I.T. Jolliffe (2009). Principal component analysis (2nd Edition). Springer.

ISSN 2348-3156 (Print)
International Journal of Social Science and Humanities Research ISSN 2348-3164 (online)
Vol. 4, Issue 1, pp: (714-723), Month: January - March 2016, Available at: www.researchpublish.com

[6] Jonathon Shlens (2003). A Tutorial on Principal Component Analysis, Derivation, Discussion and Singular Value Decomposition (Version 1). Mountain View, CA 94043.

[7] Jonathon Shlens (2014). A Tutorial on Principal Component Analysis (Version 3.02). Mountain View, CA 94043.

[8] Lindsay Smith (2002). A Tutorial on Principal Component Analysis. USA: Cornell University.

[9] Mark Richardson (2009). Principal Component Analysis.

[10] Michael E. Tipping and Christopher M. Bishop (2011). Probabilistic Principal Component Analysis. Journal of the Royal Statistical Society, Series B, 61,Part 3, pp 611-622.

[11] Neil H. Timm (2002). Applied Multivariate Analysis. Department of Education in Pshychology, School of Education, University of Pittsburgh.

[12] Richard A., Dean W and Johnson Wichern (2007). Applied Multivariate Statistical Analysis (4th Edition). University of Wisconsin-Madison.

[13] Steven Holland (2008). Principal Component Analysis (PCA). University of Georgia, Athens, GA30602-2501.

[14] Subhash Sharma (1996). Applied Multivariate Techniques. University of South Carolina.