UNIVERSITY OF CAPE COAST

MULTIVARIATE ANALYSIS OF FOOD PRICES IN GHANA

BY

FRANCIS SEGLAH

Thesis submitted to the Department of Mathematics and Statistics of the School of Physical Sciences, College of Agriculture and Natural Sciences, University of Cape Coast in partial fulfilment of the requirements for award of Master of Philosophy in Statistics

MAY 2014

DECLARATION

Candidate's Declaration

I hereby declare that this thesis is the result of my own original work and that no part of it has been presented for another degree in this university or elsewhere.

Candidate's Signature: …………………………… Date: …………………

Name: FRANCIS SEGLAH

Supervisors' Declaration

We hereby declare that the preparation and presentation of this thesis were supervised in accordance with the guidelines on supervision of thesis laid down by the University of Cape Coast.

Principal Supervisor's Signature: ………………… Date: …………………

Name: DR. BISMARK K. NKANSAH

Co-supervisor's Signature: ………………………. Date: …………………

Name: MR. FRANCIS EYIAH-BEDIAKO

ABSTRACT

This study is concerned with the evaluation of the price levels of food items in various markets across Ghana. The objective was to determine the main dimensions along which to identify specific markets in which items are high-priced and those that are low-priced. To this end, data on prices in 2008 of fifteen food items in 100 selected markets from all the regions of the country were obtained from the Statistical, Research and Information Directorate of the Ministry of Food and Agriculture. The data thus obtained were multivariate in structure since prices of 15 food items were recorded from each of 100 markets. Principal Component Analysis, which is a procedure for data reduction and summarization, is considered appropriate for analyzing this high dimensional data set. Initial exploration of the data, using zero-order and partial correlation analysis and eigen-analysis among others further informed the choice of the technique.

It was found that there are two important dimensions along which the levels of prices can be determined. These are the *weighted sum of all the food items* and the *weighted sum of only food items that are considered as the main constituent of a typical local diet*. A cluster analysis of the first component scores revealed that only a single market, Tepa in the Ashanti Region, was generally the lowest-priced market in the country. However, markets that are predominantly located in the Northern, Upper East and Upper West Regions are generally high-priced specifically the Bongo market in the Upper East region. There are, however, not much extreme-priced markets on the second component. The study shows that it will be economically beneficial to be mindful of the location and the size of the market from which one intends to make purchases of certain categories of food items.

ACKNOWLEDGEMENTS

I owe my Supervisors an exceptional gratitude. In fact, words alone cannot sufficiently express it.

My elation will be incomplete if I have not say thank you to all my Lecturers from the Department of Mathematics and Statistics who delivered that splendid and unparalleled professionalism during my course and research works. The administration staff is not forgotten at all.

I also acknowledge all the contributions made by the School of Graduate Studies for completion of my Master of Philosophy studies.

To my friends, I say without you, all the discussions will yield no dividends. The next to benefit from my acknowledgement is my beloved wife, Mrs. Georgina Esinam Seglah, and my children who stood behind me all this while with encouraging messages and best wishes.

DEDICATION

To my cherished wife, Esinam

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

LIST OF ABBREVIATIONS

CPI	Consumer Price Index
EUSIPCO	European Signal Processing Centre
FAO	Food and Agriculture Organization
FASDEP II	Food and Agriculture Sector Development Policy Document II
FPI	Food Price Index
IFAD	International Food and Agriculture Development
IITA	International Institute of Tropical Agriculture
KMO	Kaiser-Meyer Olkin's Measure of Sphericity
MakIS	Market Information System
MANOVA	Multivariate Analysis of Variance
MoFA	Ministry of Food and Agriculture
MSU	Marketing Support Unit
OECD	Organisation for Economic Co-operation and Development
SAPs	Structural Adjustment Programmes
SRID	Statistics, Research and Information Directorate
TLUs	Tropical Livestock Units
UN	United Nations
US	United States of America
VCR	Video Cassette Recorder
WFP	World Food Programme

CHAPTER ONE

INTRODUCTION

1.1 Background

The market price of a commodity is the short-term equilibrium price that is decided on daily basis resulting from short-term to medium-term stable conditions prevailing in the market. The determined market price, also known as the equilibrium price, is the price at which the equilibrium quantity prevails.

Seasonal variability in food supply and prices due to climatic changes and other occurrences make it difficult for Ghana to meet its food demands all year round. Poor rural road infrastructure limits the effective distribution of food and thereby increases market prices arbitrarily. Poor nature of roads to production centres and inadequate market information also lead to weak market integration between local, district and regional markets (FASDEP II Document, 2009). However, the commodity's own price, price of related goods, inflation rate, income levels of consumers, cost of production, the mixed economy are among factors that influence price determination in Ghana.

The Statistics Research and Information Directorate (SRID) under the Ministry of Food and Agriculture (MoFA) regularly analyses changes in wholesale and retail prices of individual food items.

According to IFAD (2011), food price tends to have a major impact on food security, at both household and country levels. Many of the world's poorest people spend more than half their income on food. Price hikes for cereals and other staples can force them to cut back on the quantity or quality of their food. This may result in food insecurity and malnutrition, with tragic implications in both the short and long term. Prices for cereals and other major food commodities have experienced two global spikes recently – one in 2007 to 2008, the other in 2010 to 2011. And they have generally remained higher than during the period between the 1980s and the early years of this century. Prices have also experienced hikes or remained at higher levels in many developing countries. The reasons for these different though inter-related phenomena are not identical. In developing countries, the key factors behind inadequate supply are low and stagnating productivity in agriculture, a deteriorating natural resource base, and weak rural and agricultural infrastructure and markets.

Between 2006 and 2008, international food prices doubled. The effects of the price surge reverberated globally, though the worst hit were low-income, food-deficit countries with meager stocks. In total, about 100 million poor rural and urban people were pushed into the ranks of the world hungry. While international food prices have declined since mid-2008, they are still substantially higher than prior to the price surge, and they are likely to remain at the same levels in 2010 or higher for the next decade (www.ruralpovertyportal.org, IFAD 2011)

Growing population and income in emerging and developing countries will add significantly to the demand for food in the coming decades. By 2050 the world's population is expected to have reached about 9 billion people and the demand for food to have increased by between 70% and 100%. This alone is sufficient to exert pressure on commodity prices. According to the latest OECD/FAO medium term outlook projections, prices of crops and most livestock products will be higher in both real and nominal terms during the decade to 2019 than they were in the decade before the 2007/08 price spikes. If the rate of growth of agricultural production does not keep pace with demand, upward pressure on prices will result. A demand or supply shock in a situation where the supply-demand balance is already tight, can, for the reasons explained in the previous paragraph, result in increased volatility around the upward trend." and "Food price inflation can also be a serious issue in middle income countries, where many consumers expend as much as half of their budget on basic foods. Even in the developed countries significantly higher food prices can create hardship for the least well-off, who tend also to devote a larger share of household spending to food. Nevertheless, consumers in developed countries face wider choices in terms of their ability to adjust spending on different types of foods and most developed countries have safety net mechanisms that are well suited to delivering targeted assistance to the most affected.

The sharp increase in food prices both in world markets and in local markets since 2006 has raised serious concerns about the food and nutrition situation of poor families in many countries. Particularly in urban areas, where people cannot grow their own food, household budgets have been squeezed. The rapid price increases are especially bad news for young children, as any disruption to their nutrition tends to have serious long-term implications, both in terms of stunting, and lower educational outcomes, affecting their earning potential in later life (www.younglives.org.uk, 2008).

1.2 Statement of the Problem

Policy documentation and implementation is effective if it is evidence-driven. The problem of food security in Ghana is directly linked with variability in market food prices and other market information for policy and evidence-based decision-making.

The MoFA defines “Food Security” as good quality nutritious food, hygienically packed and attractively presented, available in sufficient quantities all year round and located at the appropriate places at affordable prices (FASDEP II Document, 2009).

The United Nations (FAO/WFP), defines “Food Security” as a situation where all people, at all times, have physical, social and economic access to sufficient, safe and nutritious food that meets their dietary needs and food preferences for an active and healthy life. (UN, FAO Report, 2008).

The market information is woefully inadequate to players for effective planning. This inadequacy of market information calls for a study on the evaluation of general price levels of food items in a number of representative markets from all the regions of the country. This study is a step in this direction.

1.3 Objectives of the Study

A solution to the above problem stated and justified problems is to consider ranking the towns on the basis of an index that takes into account certain identifiable food items that contribute to the cost of feeding. Thus, the main objective of the study is determine the major dimensions, if any, along which to evaluate the price levels of some food items in order to identify the specific markets in which items are high-priced and those that are low-priced.

In pursuit of the stated objective, the study will target the following specific objectives:

1. Explore the suitability of the selected food items for measuring the general variation in prices of food items across the country;
2. Determine a few major dimensions which adequately represent the variations in prices of food items across the country;
3. Rank various markets in Ghana according to price levels of foodstuff;
4. Identify clusters of individual markets that reflect price levels in the respective markets.
5. Identify specific markets that may be regarded as highly-priced and those that may be regarded as low-priced.

1.4 Significance of the Study

Among the objectives for the Food and Agriculture Sector in Ghana are: (1) to ensure increased competitiveness and enhanced integration into domestic and international markets; and (2) to improve growth in income, and food security and emergency preparedness.

Based on the role of agriculture in the national development framework, it is justified that the following considerations be given utmost attention.

i. Price regimes information and its analysis is indispensable tool needed to reduce drastically the percentage of post-harvest losses that is attributed to low prices.
ii. Promotion of farmers' capacity to respond to scarce markets needs accurate market information.
iii. Promotion of specific commodities for markets requires information on clusters of markets based on price regimes.

iv. Market information, market intelligence, and continued innovation are necessary for success in international markets.

1.5 Literature Review

1.5.1 Dimension Reduction

Filmer and Pritchett (2001) obtained principal components as indices based on a range of assets owned by households. The choice of assets depended on the country under study but included agriculture-specific assets (such as tractors, threshers, harvesters) for agricultural wealth and household durables (e.g. Television, VCR, stove, refrigerator) for non-agricultural wealth. By construction, the mean of these indices is zero. A measure of household livestock assets expressed in Tropical Livestock Units (TLUs) is also included, as well as two dummy variables indicating whether the household used any fertilizer or pesticides during the year preceding the survey."

Gruber and Theis (EUSIPCO, 2006) affirm that an important tool in high-dimensional, explorative data mining is given by clustering methods. They aim at identifying samples or regions of similar characteristics, and often code them by a single codebook vector or centroid. One of the most commonly used partitional clustering techniques is the k-means algorithm, which in its batch form partitions the data set into *k* disjoint clusters by simply iterating between cluster assignments and cluster updates.

Elliott (1985) made use of clustering to form an empirical taxonomy of helpful and non-helpful events in counseling, as perceived by clients. In a counselling simulation, students acting as counselling volunteers identified and described 86 helpful and 70 non-helpful counsellor responses. Judges then rated the similarity of these responses to yield similarity matrices for the helpful and the non-helpful events. Average linkage cluster analysis was applied to these similarity matrices. The helpful events were grouped hierarchically in two superclusters related to task and interpersonal aspects of counselling. The largest cluster within task orientation was *new perspective*, and the largest interpersonal cluster was *understanding*.

Berven (1985) studied the use of case management simulations with counsellors. As one part of his analysis, he clustered the counsellors into three performance groups by applying Ward's method to six performance measures. Subgroups were clearly differentiated by performance styles, and were also significantly differentiated by experience level.

1.5.2 Food Prices

According to Claro and Monteiro (2010), fruit and vegetable participation in total food purchases increased as the price of these foods decreased, or as income increased. A 1% decrease in the price of fruit and vegetables would increase their participation by 0.79%, whereas a 1% increase in family income would increase participation by 0.27%. The effect of income tended to be smaller among higher income strata.

Roache (May 2010) in 'What Explains the Rise in Food Price Volatility?' explained that 'the macroeconomic effects of large food price swings can be broad and far-reaching, including the balance of payments of importers and exporters, budgets, inflation, and poverty. For market participants and policymakers, managing low frequency volatility (i.e., the component of volatility that persists for longer than one harvest year) may be more challenging as uncertainty regarding its persistence is likely to be higher.

It finds that low frequency volatility is positively correlated across different commodities, suggesting an important role for common factors. It also identifies a number of determinants of low frequency volatility, two of which—the variation in U.S. inflation and the U.S. dollar exchange rate—explain a relatively large part of the rise in volatility since the mid-1990s. It was pointed out that the timing of price collection was not crucial in Austria although it might play a very important role in the CPI compilation in countries with high inflation. Other issues discussed included the training of price collectors, the use of broad specifications and the greater importance of product variety over using a larger number of outlets and covering more regions. The use of electronic devices for price collection also contributed to higher quality data'. "Finally, the composition of diets can have implications for the magnitude and distribution of rising staple food prices. Households in countries where the diet is largely composed of non-tradable food staples tend to be less affected, to the extent that the prices of non-tradables do not trail the prices of tradables. For example, in our simulations Ghanaian households appear to be relatively insulated from swings in international food markets, because a large share of their diet is based on non-tradable staples such as cassava and sorghum. Should the price of these non-tradables also increase, as demand for them increases, rising food prices would have a much sharper impact. The fact that the poor are hit the hardest by rising food prices in both urban and rural areas is clearly a cause for concern. The erosion of real income in poor households not only harms their current ability to cover basic needs but has the potential to do so for some time to come thus diminishing their prospects of escaping poverty. Poor households may be forced to cope with the added stress of high food prices by depleting their asset base, reducing the number or variety of meals they consume, or reducing spending on essential non-food expenditures, such as health and education."

African Development Fund (2006) reveals that in the middle of the nineties the rural population in Burkina Faso was seriously hit by rising food prices. Whereas cotton farmers were able to cope with this shock given the simultaneous boom in the cotton sector, food crop farmers had to withdraw children from school and to let them work more intensively.

Barnett (2008) reports, the world price of wheat rose to $400 a ton, the highest level on record. That's twice the inflation-adjusted average price of wheat for the past 25 years, and twice as high as it was in May. Last year, the price of corn also hit a record of $175 a ton--more than 50 percent above the average for the previous year. Other staples, such as rice, have also hit records, ricocheting off the price of other staples as farmers switch land to high-paying commodities from other uses. Interestingly, prices have reached record highs during a time of equally record abundance: cereal crop yields last year were higher than ever, an outcome the Economist attributes to two trends: Growing demand for meat in (increasingly wealthy) China and India (livestock production requires more crops for feed); and skyrocketing demand for corn-based ethanol. As farmers have shifted croplands to feed America's growing demand for ethanol, the cost of other crops goes up, and stockpiles go down. In affluent countries like the US, the ready abundance of cheap, highly processed carbohydrates have pushed obesity and diabetes to epidemic proportions'.

Warde (1997) reports in 'Consumption, Food and Taste': Analysis of the nature of (food) consumption in present day developed economies, from a sociological perspective. Highlights are that the subject of consumption can be studied from a materialist perspective (consumer behaviour, practicalities of marketing commodities) or a cultural perspective (analysis of signs, symbols, texts, constructs).

1.5.3 Market Information

According to Robbins et al. (2000), market information services are designed to benefit farmers, traders and consumers and the services being tested in Uganda are seeking gains in farmer sales prices and improved prices for collectively sold produce. In both cases this has been achieved via the local MIS, and informal survey data with farmers working in Rakai district, claim to have received 5-15% higher returns on their sales when they are able to negotiate on known market prices, compared with farmers who simply accept prices they are offered by traders. Similarly, farmers associations in Jinja are using the local marketing agent as a link to markets and this has proven to be successful for farmers in bulking for higher value sales to larger traders. The local agribusiness centre uses the market data to forecast crop sales prices and Non-Government Organizations in northern Uganda are using the trend data to support a credit and storage scheme. However, the main target beneficiaries are the small-scale producers, as it is believed that these groups are most vulnerable to situations where market information is unavailable or asymmetric.

Robbins et al (2000) note that for a very long time, farmers in Africa have had to make decisions on what crops to plant, when to plant, where they will sell their produce and at what price. During the 1960-70s production was supported by Governments, who operated commodity marketing boards to purchase major export and staple food crops. When commodity boards were in operation the focus for market information services in African countries was to:

- Advise government on marketing policy
- Set intervention prices
- Organize marketing training for marketing authorities and cooperatives
- Document market prices as part of the Government policy analysis process.

This highly interventionist system was good for farmers, as risk was shared by the Government and farmers were able to plan production based on a known buyer price. Similar agricultural support programmes are still practiced in the major OECD nations, who currently provide approximately one billion dollars a day to support the agricultural sector. These subsidies are considered an effective use of resources, as it allows the greater part of the workforce to be employed in more remunerative industrial and service activities. Unfortunately for farmers in Africa, the commodity boards were unable to adapt to changing times and industrialization failed to occur. The commodity boards became weighed down by poor management and farmers were unable to compete with the international markets. The marketing board policies effectively led to massive internal debts coupled with poor food distribution systems, which led to food shortages in some parts of the country and excess supply in others. When the internal debts could no longer be paid, the International Monetary Fund (IMF) and World Bank were required to re-negotiate loans with most

African Governments through a process known as the Structural Adjustment Programmes (SAPs). One of the major changes required by SAPs was the reduction of the state's role in commodity marketing. This required Governments to reduce intervention support prices and at the same time to strengthen market support services such as co-operatives, credit and market information and extension services to support to the emerging private sector. Unfortunately, nearly all of these services failed to adapt in the new liberalized environment and as they become less useful, donors withdrew support and for market information several services have effectively collapsed.'

Shepherd (1997) has identified that market information comes in two main formats, (i) public dissemination of prevailing market prices and conditions and (ii) provision of price trend analysis for specific commodities to clients on demand. Shepherd (2001) recommends that the market news must be broadcast through local FM radio stations in the local language at a time when most farmers are free to listen. This information must be delivered in bulletins which aim to inform listeners about how to improve their marketing skills, and also in how to understand and use market information in negotiating for better prices.

According to Robbins et al. (2000), public provision of market information aims to avoid any asymmetry of information in the marketplace towards one group of actors. The rational is that more equal access to market information will lead to greater uniformity in prices of a given commodity at the same time across the country and this transparency within the market will assist in reducing transaction costs, through gains in trading efficiency. At present, the market is imperfect and for example in Uganda, it is common to find a two-fold difference in the price of the same commodity in two nearby markets. Consequently, traders make profits by buying at a low price on a market in one place and simultaneously sell on a different market at a high price. This encourages arbitrage, which is the process of exchange of commodities with the objective of taking advantage of price differences that exceed transaction costs. Arbitrage is typical of a situation where, (i) markets information is not accessible, (ii) markets are highly colluded, or (iii) high transaction costs are caused by problems in the supply chain.

Robbins and Ferris (1999) report that a preliminary study on maize marketing in Uganda, revealed that agricultural markets in Uganda are characterised by long chains of transactions between farm gate and consumers, lack of competitiveness between traders, collusion at all levels of trading and poor access to appropriate market information.

Prices received by farmers for the sale of their goods are significantly less than the price they could have achieved if they had the means of transporting it themselves to assembly markets even after taking the cost of transport into account.

Ferris et al. (2000) reports that additional studies carried out by IITA revealed that small-scale farmers have identified the provision of market information as the second highest priority in their efforts to gain access to better prices and markets after roads.

1.6 Outline of Thesis

The study is organized into five main chapters. Chapter One is the introductory chapter. It discusses the background, statement of the problem at stake, justification of the study, the objectives, and the institution that collects the data and how data were collected for this study. The last part considers the relevant literature on studies on prices of food items around the world.

In Chapter Two, relevant basic theory and all the methods used to achieve the objectives are discussed. Multivariate methods, particularly, Principal Component Analysis and Cluster Analysis, are reviewed.

Chapter Three contains the methodology and preliminary analysis. This constitutes descriptive statistics, graphs plots and correlation analysis. This is done to ascertain the suitability of the data for further analysis.

Chapter Four is for further analysis of the data which constitutes the main analysis of the data. It identifies the indices which are utilized to achieve the objectives of the study.

The last chapter is Chapter Five which involves the summary of the study, discussion of findings, conclusions and recommendations.

CHAPTER TWO

REVIEW OF THEORIES AND METHODS

In this chapter, we examine and review the theory behind the techniques used in analyzing the data which is multivariate in nature. It involves reviewing formulae of various multivariate methods used in the further analysis of the study.

2.1 Multivariate Techniques

Multivariate statistics is a form of statistics encompassing the simultaneous observation and analysis of more than one statistical variable. The application of multivariate statistics is multivariate analysis. There are many different models, each with its own type of analysis. Some of these techniques are: (1) Clustering Systems — assign objects into groups (called clusters) so that objects from the same cluster are more similar to each other than objects from different clusters; (2) Hotelling's T^2 — is a generalization of Student's t statistic that is used in multivariate hypothesis testing; (3) Multivariate Analysis of Variance (MANOVA) methods — extend ANOVA methods to cover cases where there is more than one dependent variable and where the dependent variables cannot simply be combined; (4) Discriminant or Canonical Analysis — attempt to establish whether a set of variables can be used to distinguish between two or more groups; (5) Multiple Regression Analysis — attempts to determine a linear formula that can describe how some variables respond to changes in others. Regression analyses are based on forms of the general linear model; (6) Principal Components Analysis — attempts to determine a smaller set of new variables that could explain the variation in the original set; (7) Multidimensional Scaling covers various algorithms to determine a set of synthetic variables that best represent the pairwise distances between records. The original method is principal coordinate analysis; (8) Canonical Correlation Analysis tries to establish whether or not there are linear relationships among two sets of variables (covariates and response). These techniques and many others are extensively treated in standard texts such as Johnson and Wichern (2003).

Multivariate techniques are the statistical techniques used to examine and analyze data involving complexities of most phenomenon that require dimensions reduction to providing fewer but without losing sufficient information for decision making and taking.

Dimension reduction methods have in common the goal of using the correlation structure (or the variance-covariance structure) among the predictor variables to accomplish the following:

1. To reduce the number of predictor components without any substantial loss of useful information;
2. To help ensure that these components are independent;
3. To provide a framework for interpretability of results.

According to Johnson and Wichern (2003) the choice of methods and types of analyses employed are largely determined by the objectives of the study and the characteristics of the data involved. The objectives of scientific investigations, for which multivariate techniques methods most naturally lend themselves, include the following: (1) Data reduction or structural simplification; (2) Sorting and Grouping; (3) Investigation of the dependence among variables; (4) Prediction; and (5) Hypothesis construction and testing.

Organization of Multivariate Data

Multivariate data arise whenever an investigator, seeking to understand a social or physical phenomenon, selects a number of variables, $p \geq 1$, or complex characteristics to record (Johnson and Wichern 1988). The values of these variables are all recorded for each distinct *item*, *individual*, or *experimental trial*. We use the notation x_{ij} = Measurement of the i^{th} variable on the j^{th} item. Consequently, *n* measurements on p variables can be displayed as a rectangular array, called **X,** of *p* rows and *n* columns, where

$$\underset{(p\ X\ n)}{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1j} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2j} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{i1} & x_{i2} & \cdots & x_{ij} & \cdots & x_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{pj} & \cdots & x_{pn} \end{bmatrix}$$

The array **X** then contains the data consisting of all the observations on all the variables.

Types of Multivariate Techniques

The Multivariate Techniques are quite many. The commonly used ones include: (1) *Analysis of covariance structure group* — Principal Component Analysis, Factor Analysis, Canonical Correlation Analysis and Correspondence Analysis; (2) *Classification and Grouping Techniques group* — Cluster Analysis, and Discrimination and Classification.

We shall review the theory of Principal Component Analysis and Cluster Analysis, which are the main methods used in this study.

2.2 Principal Component Analysis

The Principal Component Analysis (PCA) is a data dimension reduction method and a standard tool in modern data analysis which seeks to explain the correlation structure of a set of predictor variables using smaller set of linear combinations of these variables referred to as components.

Objectives of PCA

The objective of PCA is to identify a new set of orthogonal axes such that

1. The coordinates of observations with respect to each other of axes give the values for the new variables.
2. Each new variable is a linear combination of the original variables
3. The first new variable accounts for the maximum variance in the data
4. The second new variable accounts for the maximum variance that has not been accounted for by the first variable
5. The third new variable accounts for the maximum variance that has not been accounted for by the first two variables.
6. The p^{th} new variable accounts for the variance that has not been accounted for by the p-1 variables
7. The p new variables are uncorrelated

2.2.1 Derivation of Principal Components

As mentioned above, principal component is one of methods used for dimensionality-reduction and summarization. In order to determine the major dimensions underlying a high p-dimensional data set the correlation or the variance-covariance matrix of the indicator variables is subjected to a principal component analysis. In this technique, each of the p possible principal components (y_i) is expressed as a linear combination of the original variables x_j as

$$y_i = \sum_{j=1}^{p} a_{ij} x_j \ , \qquad (1)$$

where the set of coefficients, a_{ij} $(j = 1, 2, \Lambda\ , p)$, is the eigenvector of y_i. Equation (1) could be written such that the principal component scores are standardized to have a unit variance. Denoting the eigenvalue of the component (y_i) by α_i, then y_i accounts for an amount of α_i of the variation in the data. Then $\text{Var}\left(\frac{y_i}{\sqrt{\alpha_i}}\right) = 1$ and $\frac{y_i}{\sqrt{\alpha_i}}$ is a standardized principal component. Equation (1) can then be written as

$$y_i = \sum_{j=1}^{p} \beta_{ij} x_j , \qquad \beta_{ij} = a_{ij}\sqrt{\alpha_i} \qquad (2)$$

where $\sum_{j=1}^{p} \beta_{ij}^2$ is the eigenvalue of y_i and $\{\beta_{ij}, j = 1, 2, \Lambda\ , p\}$ is a vector of loadings of y_i on the variables. The matrix alternative of Equation (2) is of the form

$$\mathbf{y} = \mathbf{\Lambda}'\mathbf{x} \qquad (3)$$

where y is a $p \times 1$ vector of standardized components;

$\mathbf{\Lambda}$ is a $p \times p$ orthogonal matrix of component loadings;

$\mathbf{x}$ is a $p \times 1$ vector of indicator variables.

Further to the requirement that principal components must be linear combinations of the original variables, it is also necessary that the linear combinations, $y_1, y_2, \ldots\ldots, y_p$, are such that

1. they are uncorrelated among themselves;
2. the first principal component is the linear combination that accounts for the highest variation in the data;
3. the second principal component is the linear combination that accounts for maximum remaining variation while being uncorrelated with the first principal component; and
4. the third principal component accounts for the maximum remaining variation while being uncorrelated with the two previously determined principal components; and
5. the ***p***th principal component accounts for the maximum remaining variation, while being uncorrelated with the first $(p-1)$ principal components.

2.2.2 Extracting Principal Components

The variance in the correlation matrix is "repackaged" into *p* eigenvalues. Each eigenvalue represents the amount of variance that has been captured by one component. Each component is a linear combination of the *p* variables. The first component accounts for the largest possible amount of variance. The second component, formed from the variance remaining after that associated with the first component has been extracted, accounts for the second largest amount of variance, etc. The principal components are extracted with the restriction that they are orthogonal. Geometrically they may be viewed as dimensions in *p*-dimensional space where each dimension is perpendicular to each other dimension. Each of the *p* variable's variance is standardized to one. Each factor's eigenvalue may be compared to 1 to see how much more (or less) variance it represents than does a single variable.

With p variables there is $p\,X\,1 = p$ variance to distribute. The principal components extraction will produce p components which in the aggregate account for all of the variance in the p variables. That is, the sum of the p eigenvalues will be equal to p, the number of variables. The proportion of variance accounted for by one component equals its eigenvalue divided by p.

The following are some of the procedures that are used to determine the number of components to retain in a principal component analysis:

1. *Scree Plot*: A scree plot is a plot of the Eigenvalues against the number of factors in order of extraction. Experimental evidence indicates that the point at which the scree begins denotes the true number of factors. Generally, the number of factors determined by a scree plot will be one or a few more than that determined by the Eigenvalue criterion.
2. *Percentage of Variance accounted for by the component*. In this approach the number of factors extracted is determined so that the cumulative percentage of variance extracted by the factors reaches a satisfactory level. It is recommended that the components extracted should account for at least 60% of the variance. The proportion of variation accounted for by the i^{th} component is given by

$$\text{Proportion of variance} = \frac{\lambda_i}{\sum_{i=1}^{p} \lambda_i}$$

3. *Determination based on eigenvalues*. In this approach, only component with eigenvalues, $\lambda > 1.0$ are retained. An eigenvalue represents the amount of total variance explained by the component. Components with variance less than 1.0 are no better than a single variable, since, due to standardization, each variable has a variance of 1.0.

4. *The relevance of the component within the objective of the subject matter*. Within the context of the area under consideration, it will be appropriate to extract only those components that meet the objective of the study. These are the components that are relevant to the study.

In the determination of the number of components to retain, Larose (2006) has observed that Social Scientists may be content for their components to explain only 60% or so of the variability, since human response factors are so unpredictable, whereas Natural Scientists might expect their components to explain 90 to 95% of the variability, since their measurements are intrinsically less variable. Other factors also affect how large a proportion is needed. For example, if the principal components are being used for descriptive purposes only, such as customer profiling, the proportion of variability explained may be a shade lower than otherwise. On the other hand, if the principal components are to be used as replacements for the original (standardized) data set and used for further inference in models downstream, the proportion of variability explained should be as much as can conveniently be achieved given the constraints of the other criteria.

2.3 Rotation Procedures

We review briefly two commonly used rotation methods in the use of principal component analysis. These are the Varimax and the Quartimax rotation procedures.

2.3.1 Varimax Rotation

This is a type of orthogonal rotation. Its major objective is to have a factor structure in which each variable loads highly on one and only one factor. That is, a given variable should have a high loading on one factor and near zero loadings on other factors. Such a factor structure will result in each factor representing a distinct construct. Its goal is to minimize the complexity of the components by making the large loadings larger and the small loadings smaller within each component.

This is achieved by maximizing the variance of the squared loading across variables, subject to the communality of each variable is unchanged. That is, for any given factor

$$V_j = \frac{\sum_{j=1}^{p}(\lambda_{ij}^2 - \lambda_j^2)^2}{p} = \frac{p\sum_{j=1}^{p}\lambda_{ij}^4 - (\sum_{j=1}^{p}\lambda_{ij}^2)^2}{p^2}$$

Where V*i* is the variance of the communalities of the variables within factor *j* and λ_j^2 is the average squared loading for factor *j* and *p* is as usual.

The total variance for all the factors is given by

$$V_j = \sum_{j=1}^{m} V_j = \sum_{j=1}^{m}\left(\frac{p\sum_{i=1}^{p}\lambda_{ij}^4 - \left(\sum_{i=1}^{p}\lambda_{ij}^2\right)^2}{p^2}\right) = \frac{\sum_{j=1}^{m}\sum_{i=1}^{p}\lambda_{ij}^4}{p} - \frac{\sum_{j=1}^{m}\left(\sum_{i=1}^{p}\lambda_{ij}^2\right)^2}{p^2}$$

Since the number of variables remains the same, maximizing the preceding equation is the same as maximizing

$$pV = \sum_{j=1}^{m}\sum_{i=1}^{p}\lambda_{ij}^4 - \frac{\sum_{j=1}^{m}\left(\sum_{i=1}^{p}\lambda_{ij}^2\right)^2}{p}$$

2.3.2 Quartimax Rotation

The quartimax rotation is an orthogonal rotation. The major objective of this type of rotation technique is to obtain a pattern of loadings such that all the variables have a fairly high loading on one factor, and each variable has a high loading on one other factor and near zero loadings on the remaining factors. This technique makes large loadings larger and small loadings smaller within each variable.

This is achieved by maximizing the variance of the loadings across factors, subject to the constraints that the communality of each variable is unchanged. Thus, suppose for any given variable *i*, we define

$$Q_i = \frac{\sum_{j=1}^{m}\left(\lambda_{ij}^2 - \lambda_i^2\right)^2}{m} = \frac{m\sum_{j=1}^{m}\lambda_{ij}^4 - \left(\sum_{j=1}^{m}\lambda_{ij}^2\right)^2}{m^2}$$

where Q*i* is the variance of the communalities and *m* as usual. The total variance of all variables is given by

$$Q = \sum_{i=1}^{p} Q_i = \sum_{i=1}^{p}\left(\frac{m\sum_{j=1}^{m}\lambda_{ij}^4 - \left(\sum_{j=1}^{m}\lambda_{ij}^2\right)^2}{m^2}\right)$$

2.4 Statistics Associated With Principal Component Analysis

2.4.1 Bartlett's Test of Sphericity

Bartlett's test of sphericity is a test statistic used to examine the hypothesis that the variables are uncorrelated in the population. In other words, it examines the extent to which the correlation departs from orthogonality. Thus, if the orthogonal correlation matrix has determinant of 1, it indicates that the variables are not correlated. If there is perfect correlation between two variables, then the determinant will be very close to 0. The test statistics is given by

$$\chi_v^2 = -2\left[1 - \frac{1}{6pm}(2p^2 + p + 2)\right]\ln\Lambda$$

with $v = \frac{1}{2}(p-1)(p+2)$ degrees of freedom, where

$$\Lambda = \left[\frac{\prod_{i=1}^{p} \lambda_i}{\frac{1}{p}\left(\sum_{i=1}^{p} \lambda_i\right)^p} \right]^{\frac{1}{2}}$$

and p is the number of variables and m is the number of components.

2.4.2 Kaiser-Meyer Olkin's Measure of Sampling Adequacy

The Kaiser-Meyer-Olkin's measure of sampling adequacy, usually simply referred to as the KMO, for a variable X_i is the ratio of the sum of the squared simple correlation coefficient (r_{ij}^2) between X_i and each other variable X_j $(i \neq j)$ to that same sum plus the sum of the squared partial correlation coefficient (pr_{ij}^2) between X_i and each other X_j. Thus,

$$KMO = \frac{\sum r_{ij}^2}{\sum r_{ij}^2 + \sum pr_{ij}^2} \tag{4}$$

The expression for the simple correlation coefficient (r_{ij}) between X_i and X_j is given as

$$r_{X_i X_j} = \frac{s_{ij}}{\sqrt{s_{ii}}\sqrt{s_{jj}}} \tag{5}$$

The value s_{ij} is the (i, j) element of the matrix of sum of squares and cross-product matrix of the data. The expression for the sample partial correlation coefficient between X_i and X_j controlling for the other variables, $\mathbf{Y} = (Y_1, Y_2, \Lambda, Y_q)'$ say, is given by

$$r_{X_i X_j \cdot \mathbf{Y}} = \frac{s_{X_i X_j \cdot \mathbf{Y}}}{\sqrt{s_{X_i X_i \cdot \mathbf{Y}}}\sqrt{s_{X_j X_j \cdot \mathbf{Y}}}} \tag{6}$$

The element $S_{X_i X_j \cdot \mathbf{Y}}$ is the $(i,\ j)$ entry in the variance-covariance matrix

$$\mathbf{S}_{\mathbf{XX}} - \mathbf{S}_{\mathbf{XY}} \mathbf{S}_{\mathbf{YY}}^{-1} \mathbf{S}_{\mathbf{YX}} \tag{7}$$

In particular, if X_1, X_2 and X_3 are any three variables, the first order partial correlation coefficient between X_1 and X_2 controlling for X_3 is given by

$$r_{X_1X_2 \cdot X_3} = \frac{r_{X_1X_2} - r_{X_1X_3} \times r_{X_2X_3}}{\sqrt{\left(1 - r_{X_1X_3}^2\right)}\sqrt{\left(1 - r_{X_2X_3}^2\right)}} \tag{8}$$

where $r_{X_1X_2}$ is the zero-order correlation coefficient between X_1 and X_2. It can be seen from Equation (8) that if $r_{X_1X_3}$ and $r_{X_2X_3}$ have the same sign, then controlling for X_3 reduces $r_{X_1X_2 \cdot X_3}$ (i.e., it makes it less positive or high negative as the case may be). On the other hand, if $r_{X_1X_3}$ and $r_{X_2X_3}$ have opposite signs controlling for X_3 increases $r_{X_1X_2 \cdot X_3}$. The implication is that partial correlation coefficient between two variables X_i and X_j controlling for other variables, depends to a large extent on the correlation coefficient between each of X_i and X_j and other variables.

Now a high value of the KMO indicates that the partial correlation coefficient between X_i and X_j is small. That is, after removing the effect of all other variables, the association between X_i and X_j is low. Thus, for any pair of variables, their association is strongly influenced by the other variables. This is an indication of generally high pairwise correlations among the variables. Consequently, we will conclude that the variables or sets of the variables belong together. Therefore, a high KMO value is an indication that there are distinct groupings among the variables, and hence, a justification for using a dimensionality reduction technique to analyze the data. On the other hand, if the KMO value is low, it indicates that the partial correlation coefficient between X_i and X_j is high. That is, after removing the effect of all other variables, the association between X_i and X_j is (still) high. Thus, for any pair of variables, their association is not much influenced by the other variables. This is an indication of generally low pairwise correlations among the variables. Consequently, we will conclude that there are no distinct sets of variables (among the original variables) that belong together. Therefore, a low KMO value is an indication that there are no distinct groupings among the variables, and hence, there is no motivation for using a dimensionality reduction technique to analyze the data. Thus, the suitability of principal component technique, for example, requires generally low elements of the matrix of partial correlation coefficients. Kaiser and Rice (1974) have provided a guide for the interpretation of the values of the KMO measure of sampling adequacy.

The guide for interpreting various values of KMO is given in Table 1.

Table 1: Interpretation of the KMO Value

KMO Value	Description
0.90 – 1.00	Marvelous
0.80 – 0.89	Meritorious
0.70 – 0.79	Middling
0.60 –0.69	Mediocre
0.50 – 0.59	Miserable
0.00 – 0.49	Don't Factor

Source: Kaiser and Rice (1974)

The value of KMO indicates the extent of homogeneity of the indicators measuring the same dimensions. This is a diagnostic measure of the appropriateness of factor analysis for a given data.

2.4.3 Correlation Coefficient between Principal Component and Variable

The i^{th} principal component $Y_i = \sum_{k=1}^{p} a_{ik} X_k$ may be written as

$$Y_i = \mathbf{a_i}'\mathbf{X}$$

Where $\mathbf{X} = (X_1, X_2, X_3, \Lambda, X_p)'$ is the p–dimensional variable vector. If we represent the p dimensional vector $\mathbf{a}'_\mathbf{j}$ by $\mathbf{a}'_\mathbf{j} = (0,0,\Lambda,0,1,0,0,\Lambda\ 0)$, where $a_{jj} = 1$, then we can write the component variable X_j as

$$X_j = \mathbf{a}'_\mathbf{j}\mathbf{X}$$

Given the Variance-Covariance matrix of $\mathbf{X}$ as $\mathbf{\Sigma}$, we infer from Equation (2) that

$$\text{Var}(Y_i) = \lambda_i \tag{9}$$

But since $a_i'\mathbf{\Sigma} a_i = \lambda_i$, it implies that $\lambda_i a_i = \Sigma a_i$. Now given $\text{Var}(X_j) = s_j^2$ we have

$$\text{Cov}(Y_i, X_j) = \text{Cov}(a_i'X, a_j'X) = a_j'\Sigma a_i$$

$$= a_j'(\lambda_i a_i)$$

$$= \lambda_i a_j' a_i$$

$$= \lambda_i a_{ij}$$

The correlation coefficient between the i^{th} principal component, Y_i, and the j^{th} variable, X_j, is generally given by the expression

$$\rho_{Y_i X_j} = \frac{Cov\left(Y_i, X_j\right)}{\sqrt{Var\left(Y_i\right)}\sqrt{Var\left(X_j\right)}}.$$

Making appropriate substitutions, we obtain an expression for the correlation coefficient, $\rho_{Y_i X_j}$, between Y_i, and the *j*th variable, X_j, in terms of the loading, λ_i, and the weights, a_{ij}, as

$$\rho_{Y_i X_j} = \frac{a_{ij}\sqrt{\lambda_i}}{s_j} \qquad (10)$$

Equation (10) suggests that the size of the correlation coefficient between a component and a variable depends to a large extent on the variability in the observations on the variable. If the values are dispersed, the correlation coefficient will more likely be small. On the other hand, if there is little spread in the values of the variable, the correlation coefficient is more likely to be large.

In this study, the prices are likely to vary from market to market. Thus, it will be necessary to ensure that this variability does not affect the determination of the component. It may be necessary to standardize the prices in the case where wide variations exist in the prices. In this case, Equation (10) reduces to

$$\rho_{Y_i X_j} = a_{ij}\sqrt{\lambda_i}$$

Thus, the formation of the components would not be influenced by variables with high variations.

2.4.4 Standard Error of the Mean and Coefficient of Variation

Two basic routine statistical methods that will be of importance in this study are the standard error of the mean and the coefficient of variation. The standard error of the mean $(\overline{X}_i)$ of a variable X_i $(i = 1, 2, 3, \Lambda, k)$ is denoted by $SE(\overline{X}_i)$ and is given by

$$SE(\overline{X}_i) = \frac{S_{X_i}}{\sqrt{n_i}}. \tag{11}$$

Where n_i is the sample size on X_i. It is a measure of the spread of observed values of the variable about the mean. If we are dealing with prices of commodities in various markets, for example, then a small value of $SE(\overline{X}_i)$ shows that the price of the given commodity is stable, which is desirable. On the other hand, if the value of $SE(\overline{X}_i)$ is large, then the price of the commodity varies widely from market to market. But the implication of the value of $SE(\overline{X}_i)$, in relation to other variables, may be affected by the number of observations (n_i) on X_i. To eliminate this problem, one way is to ensure that the number of observations is the same for all the variables. In this case,

$$SE(\overline{X}_i) = \frac{S_{X_i}}{\sqrt{n}}$$

The coefficient of variation of a variable X_i $(i = 1, 2, 3, \Lambda, k)$ is denoted by $CV(X_i)$ and is given by

$$CV(X_i) = \frac{S_{X_i}}{\overline{X}_i} \tag{12}$$

usually given in percentage.

In terms of the $SE(\overline{X}_i)$, Equation (12) may be written as

$$\begin{aligned} CV(X_i) &= \frac{S_{X_i}}{\sqrt{n_i}} \times \frac{\sqrt{n_i}}{\overline{X}_i} \\ &= \frac{S_{X_i}}{\sqrt{n_i}} \times \frac{1}{\overline{X}_i} \times \sqrt{n_i} \end{aligned}$$

Assuming that $n_i = 100$, (as is the case in this study) then writing $CV(X_i)$ as a percentage, it can be simplified as

$$CV(X_i) = \frac{1}{\bar{X}_i} \times SE(\bar{X}_i) \times 100 \qquad (13)$$

Equation (13) shows that the coefficient of variation is also a measure of the spread of observed values of the variable about the mean as it is in the case for standard error. However, the use of the *CV* will be more relevant if we are comparing dispersion in two samples in which the measurements are not in the same unit. If measurements are in the same unit, then the *CV* may be used with care as it is largely influenced by the mean of the variable. The following is an illustration of the relative use of the two measures of precision.

2.5 Computation and Interpretation of Component Scores

The discussion of the computation of component scores is meaningful in relation to the data under examination. A generalization of the discussion may run into difficulties. Thus, we review this section with specific reference to the market price data.

The i^{th} principal component is given in Equation (2.1) as $Y_i = \sum_{j=1}^{p} a_{ij} X_j$. Clearly, the value of this expression will be influenced by the observed values of the variables, X_j. The variability in the values of these variables will then be reflected in the component scores. This can distort interpretation of the score. To eliminate the effect of the variation in the individual variables, we first standardize the data. Thus, the component score corresponding to Y_i is given by

$$C_i = \sum_{j=1}^{p} a_{ij} \frac{X_j - \mu_j}{s_j} \tag{14}$$

The standardization process determines the magnitude and the sign of the *j*th term in the expression in Equation (14). Assuming that all the weights, a_{ij}, are positive, we consider three typical scenarios: C_i could be a high positive value, a high negative value or close to zero. A high positive score is obtained if the values of most of the items $(X_{j,}\ j = 1, 2, \Lambda, p)$ are higher than the average values of the respective items. Thus, in relation to the market price data used for this study, a high positive score C_i indicates that the market's prices in all the commodities $(X_{j,})$ are consistently much higher than the average price for all the commodities. Such a market is a typically high-priced market.

A high negative score is obtained if the values of most of the items $(X_{j,}\ j = 1, 2, \Lambda, p)$ are lower than the average values of the respective items. Thus, in relation to the market price data, a high negative score, C_i, indicates that the market prices in most of the commodities are consistently much lower than the average price for all commodities. Such a market is a typically high-priced market.

A very small score (close to zero) is obtained if the values of most of the items are just about the same as the average values of the respective items. Thus, in relation to the market price data, a small value of C_i indicates that the market consistently has prices of all the commodities that are about the same value as the average price of all the commodities. Such a market is a typical average (moderate) -priced market.

The above discussion is basically on the effect of the standardization of the data on the sign and size of the component score assuming that the loadings are all positive. However, the magnitude of the score is also determined to a large extent by the size of the loading, a_{ij}, on the variable X_j. In order to explain this point, we consider two scenarios: the loadings are almost equal; and some of the loadings are much larger than others.

In the case where the loadings are almost equal, the principal component, Y_i, is usually referred to as a *weighted sum* of the original variables. In this case all the variables have about the same influence in the formation of the component. In relation to the market price data, a high component score means that the market has commodities with high prices, and hence, a high-priced market.

On the other hand, if a_{ij} are large for some variables, say X_k and X_t, and a_{ij} are low on all others, then the size of the component score would have been influenced by the loadings of X_k and X_t, a_{ik} and a_{it}. If the component score is high and positive, then it implies that the commodities represented by the variables X_k and X_t are high-priced in that market. If the component score is high and negative, then it implies that the commodities represented by the variables X_k and X_t are low-priced in that market. If the component score is small (close to zero) then it implies that the two commodities have average (moderate) prices in that market.

2.6 Cluster Analysis

Cluster Analysis (Sharma, 1996) is a technique used for combining observations into groups of clusters such that each group or cluster is homogeneous or compact with respect to certain characteristics; each group should be different from other groups with respect to the same characteristics. The grouping is done on the basis of similarities or distances (dissimilarities). The inputs required are similarity measures or data from which similarities can be computed. By similarity measure we mean a measure that represent the ‘‘closeness’’ or ‘‘similarity’’ between two items. There is often a great deal of subjectivity involved in the choice of a similarity measure. Important considerations include the nature of the variables (discrete, continuous, binary) or scales of measurement (nominal, ordinal, interval, ratio) and the knowledge of the subject matter. When items (units or cases) are clustered, proximity is usually indicated by some sort of distance. On the other hand, variables are usually grouped on the basis of correlation coefficients or like measures of association. Below are brief reviews of the various similarity measures.

2.6.1 Similarity Measures

Similarity measures consider Euclidean (straight-line) distance, $d(\mathbf{x},\mathbf{y})$ between two p-dimensional observations $X=\left[x_1,x_2,...,x_p\right]'$ and $Y=\left[y_1,y_2,...,y_p\right]'$. The distance $d(\mathbf{x},\mathbf{y})$ is given by

$$d(\mathbf{x},\mathbf{y})=\sqrt{\sum_{i=1}^{p}(x_i-y_i)^2}=\sqrt{(X-Y)'(X-Y)}$$

The statistical distance between the same two observations is of the form

$$d(\mathbf{x},\mathbf{y})=\sqrt{(\mathbf{x}-\mathbf{y})'\mathbf{A}(\mathbf{x}-\mathbf{y})}$$

where $\mathbf{A}=\mathbf{S}^{-1}$ and $\mathbf{S}$ is the sample variance-covariance matrix. The Euclidean distances for p variables is given by

$$D_{ij}^2 = \sum_{k=1}^{p} (x_{ik} - x_{jk})^2 \tag{15}$$

Another distance measure is the Minkowski metric given by

$$D_{ij} = \left[\sum_{k=1}^{p} \left(\left| x_{ik} - y_{jk} \right| \right)^m \right]^{\frac{1}{m}} \tag{16}$$

where D_{ij} is the distance between observations *i* and *j*, and *p* is the number of variables.

2.6.2 Types of Clustering Methods

Data clustering algorithms are either Hierarchical or Nonhierarchical. We briefly enumerate the methods of hierarchical clustering and illustrate one that is relevant to this study.

Hierarchical Clustering

The Hierarchical Clustering (algorithms) find successive clusters using previously established clusters. These algorithms can be either agglomerative ("bottom-up") or divisive ("top-down"). Agglomerative algorithms begin with each element as a separate cluster and merge them into successively larger clusters. Divisive algorithms begin with the whole set and proceed to divide it into successively smaller clusters. Density-based clustering algorithms are devised to discover arbitrary-shaped clusters. In this approach, a cluster is regarded as a region in which the density of data objects exceeds a threshold.

A number of different rules or methods have been suggested for computing distances between two clusters. Some of the popular methods are

1. Centroid method
2. Nearest – Neighbor or Single – Linkage Method
3. Farthest – Neighbor or Complete – Linkage Method

4. Average – Linkage Method
5. Ward's Method

Detailed discussions of these methods are readily available in standard text such as Sharma, (1996) and Johnson and Wichern (2003).

Centroid Method

In the centroid method, each group is replaced by an ''average subject'' which is the centroid of that group. The distance is obtained by computing the squared Euclidean distance between the centroids of respective clusters. Like a typical hierarchical clustering algorithm, the centroid method forms clusters in a hierarchical fashion. That is, the number of clusters at each stage is one less than the previous one. If there are n observations, then at Step 1, Step 2, …, Step $n-1$ of the hierarchical process the number of clusters, respectively, will be $n-1$, $n-2$, …, 1. Each cluster is represented by the centroid of that cluster for computing distances between clusters. To illustrate the Centroid method, we use a small portion of the data matrix in Appendix A which covers six cases (markets) from where prices on five variables are provided.

CHAPTER THREE

METHODOLOGY AND PRELIMINARY ANALYSIS

In this chapter, we carry out an exploratory analysis of the data to identify its characteristic features and the relationship among variables. We use mainly routine methods of descriptive statistics, contingency tables, graphs and plots. We also perform covariance analysis in an attempt to determine the suitability of the data for any further analysis.

3.1 The Data

The data for the thesis is obtained from The Statistical, Research and Information Directorate (SRID). It covers the wholesale prices in 2008 (quoted in GHS and Gp) of 31 foodstuffs, out of which 15 foodstuffs were selected for the study. It covers 135 markets in the entire country, out of which 100 are selected. Out of the 135 markets, 79 are in the rural area and 56 in the urban area. Generally, the SRID collects prices on the following items: Maize, Imported Rice, White Yam, Cassava, Tomatoes, Garden Eggs, Dried Pepper, Red Groundnut, White Cowpea, Groundnut Oil, Palm Oil, Orange, Banana, Smoked Herrings and Kobi (Salted Dried Tilapia), Millet, Guinea Corn, Cocoyam, Gari, Cassava Chips (Kokonte), Plantain (Apentu), Pineapple, Mangoes, Onion, Okro, Kpakpo Shitor, Unshelled Groundnut, Local Rice, Ginger, Beef, Anchovies, Pork, Eggs Commercial and Live Births. All the foodstuffs are covered at both wholesale and retail levels except for Live Bird, Beef and Pork which are covered only at the retail level. The 15 foodstuffs/variables selected for this work are given in Table 3.2 with their variable names and the respective standard unit of sale/measurement. The figures in the data collected represent annual average wholesale prices of foodstuffs.

3.2 Distribution of Markets in the Regions of Ghana

In Table 2, we have the number of markets available in Ghana and the number actually selected for the study. As already described in Chapter One,

Table 2: Number of Selected Markets from available Markets

REGION	Number of Available Markets (*N*)				Number of Selected Markets (*n*)			
	Rural	Urban	Total	%	Rural	Urban	Total	%
Ashanti	11	4	15	11.1	8	3	11	11
Brong Ahafo	9	8	17	12.6	7	6	13	13
Central	10	4	14	10.3	7	3	10	10
Eastern	7	8	15	11.1	5	6	11	11
Greater Accra	4	12	16	11.9	3	9	12	12
Northern	10	2	12	8.9	8	1	9	9
Upper East	5	3	8	5.9	4	2	6	6
Upper West	4	3	7	5.2	3	2	5	5
Volta	8	7	15	11.1	6	5	11	11
Western	11	5	16	11.9	8	4	12	12
Total	79	56	135	100.0	59	41	100	100

data is collected by the Statistics Research and Information Directorate (SRID) of Ministry of Food and Agriculture in 2008 for the study. The table is made up of two parts. The first part involves available markets from which SRID collects market data in the entire ten regions of Ghana. The markets are categorized as rural and urban. The first column shows the ten regions. The second column shows the number of rural markets while the third column contains the number of urban markets. The fourth column shows respective percentages of markets per region. Out of this, 79 (58.5%) are rural and 56 (41.5%) are urban in type. In all, SRID collected data from 135 markets.

The second part of the table, which starts from the sixth column, involves markets actually selected for the study from the markets described in the first part. The selected markets from both the rural and urban centres are contained in column six and seven. In column eight, the total number of markets selected per region is summarized and the respective percentages are shown in column nine.

We see from Table 2 that the number of markets is selected from each of the region so that the proportion in the sample is the same as that in number of markets identified. We also note that with exception of Eastern and Greater Accra regions, more markets are located in the rural centres than the urban centres in each of the other eight regions. Markets from Upper East and West constitute the smallest numbers. Other features of the relationships between the numbers of available markets from both the rural and urban centres in the regions are better illustrated in Figure 1.

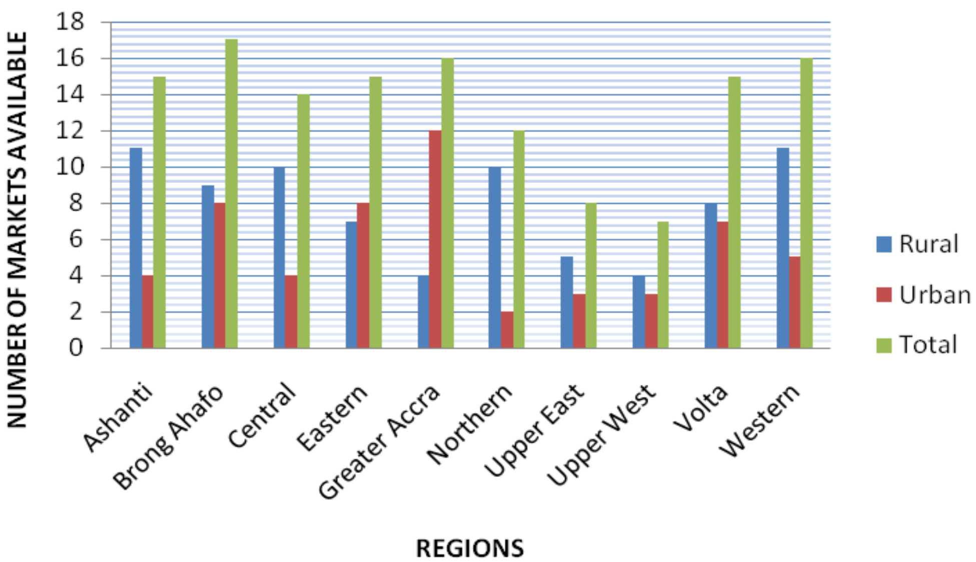


Figure 1: Distribution of Available Markets

The Greater Accra region has the highest (12) number of urban markets while Ashanti region and Brong Ahafo regions have the highest number of rural markets of 11 each as exhibited also by Figure 1 above.

The least number (2) of markets occur in the Northern region followed by Greater Accra region and Upper West region which have four rural markets each.

This study considers proportionally stratified random sampling in selection of the samples for the study. In all, 100 markets (cases) are selected for the study out of the 135. It consists of 59 out of 79 rural markets and 41 out of 56 urban markets.

Table 3: Variables for the Study

No.	VARIABLES (COMMODITIES)	VARIABLE NAME	UNIT OF SALE
1	Maize	Mz	100kg
2	Imported Rice	RiImp	50kg
3	White Yam	YmWt	250kg (100 Tubers)
4	Cassava	Cv	91kg
5	Tomatoes	Tm	52kg (Crate)
6	Garden Egg	GEg	27kg
7	Dried Pepper	PpDr	16kg
8	Red Groundnut	GnR	82kg
9	White Cowpea	CpWt	109kg
10	Groundnut Oil	GnOil	18 Litres
11	Palm Oil	PmOil	18 Litres
12	Orange	Org	20kg (100 Singles)
13	Banana	Ban	6-8kg (Average Bunch)
14	Smoked Herring	HrSmk	100 Singles
15	Kobi	Kobi	100 Singles

Source: Field Data, SRID-MoFA 2008

The foodstuffs in Table 1 are in eight categories. In Table 4, we have the categorizations of the food items. We see from Table 4 that with exception of the spices category that has only one item, each of the remaining seven categories contains two food items.

Table 4: Categorization of Foodstuff

No.	Category of Foodstuff	Food Items
1	Cereals	Maize, Imported Rice
2	Roots and Tubers	Cassava, White Yam
3	Vegetables	Tomatoes, Garden Eggs
4	Pulses	Cowpea, Red Groundnut
5	Fish	Smoked Herrings, Kobi
6	Spices	Dried Pepper
7	Oil	Palm Oil, Groundnut Oil
8	Fruits	Orange, Banana

Source: Field Data, SRID-MoFA 2008

As indicated earlier, there are in all 100 markets selected for the study. Out of the 100 markets, 59 are in the rural area and 41 in the urban area. In Table 5, we have the details of the markets selected from all the ten regions.

Table 5: Markets Selected for the Study

No	Region	Markets	Total
1	Ashanti	Ejura, Tepa, Agona, Abofour, Adugyaman, Juabeso, Agogo, Nsuta, Kumasi (Kejetia), Bekwai, Obuasi	11
2	Brong Ahafo	Atebubu, Kwame Danso, Yeji, Kukuom, Nsawkaw, D Nkwanta, Drobo, Sunyani, Techiman, Brekum, Dorman Ahenkro, Goaso, Kintampo	13
3	Central	Mankessim, Bawjiasi, Kasoa, Elmina, Assin Praso, Ajumako, Fante Nyankomase, Cape Coast, Swedru, Dunkwa	10
4	Eastern	Somanya, Akuse, Ahoman, Anyinam, Agormanya, Koforidua, Suhum, New Tafo, Asamankese, Akim Oda, Mpraeso,	11
5	Greater Accra	Kasseh, Ada Foah, Hobor, Makola, Agbogbloshie, Kaneshie, Mamprobi, Tema, Ashaiman, Madina, Malam, Dome	12
6	Northern	Bimbila, Damongo, Nalerugu, Gushiegu, Salaga, Kumbungu, Zabzugu, Saboba, Tamale	9
7	Upper East	Fumbisie, Bongo, Garu, Zebila, Bolga, Navrongo	6
8	Upper West	Babile, Jirapa, Busie, Wa, Tumu	5
9	Volta	Mafi-Kumase, Abotoase, Kpeve, Logba-Alakpeti, Kpetoe, Kute, Ho, Kpando, Hohoe, Dambai, Denu	11
10	Western	Bogoso, Asawinso, Bibiani, Sewfi Dwinase, Tikobo, Juabeso, Elubo, Dadieso, Takoradi, Trakwa, Agona Nkwanta, Sefwi Bewai	12

Source: Field Data, SRID-MoFA 2008; The SRID Market Enumerator's Manual, June 2000

Besides, these markets selected for the study are marked in their geographical locations with dots, using the respective market names and IDs in Table 6, on the map of Ghana (as shown in Figure 2).

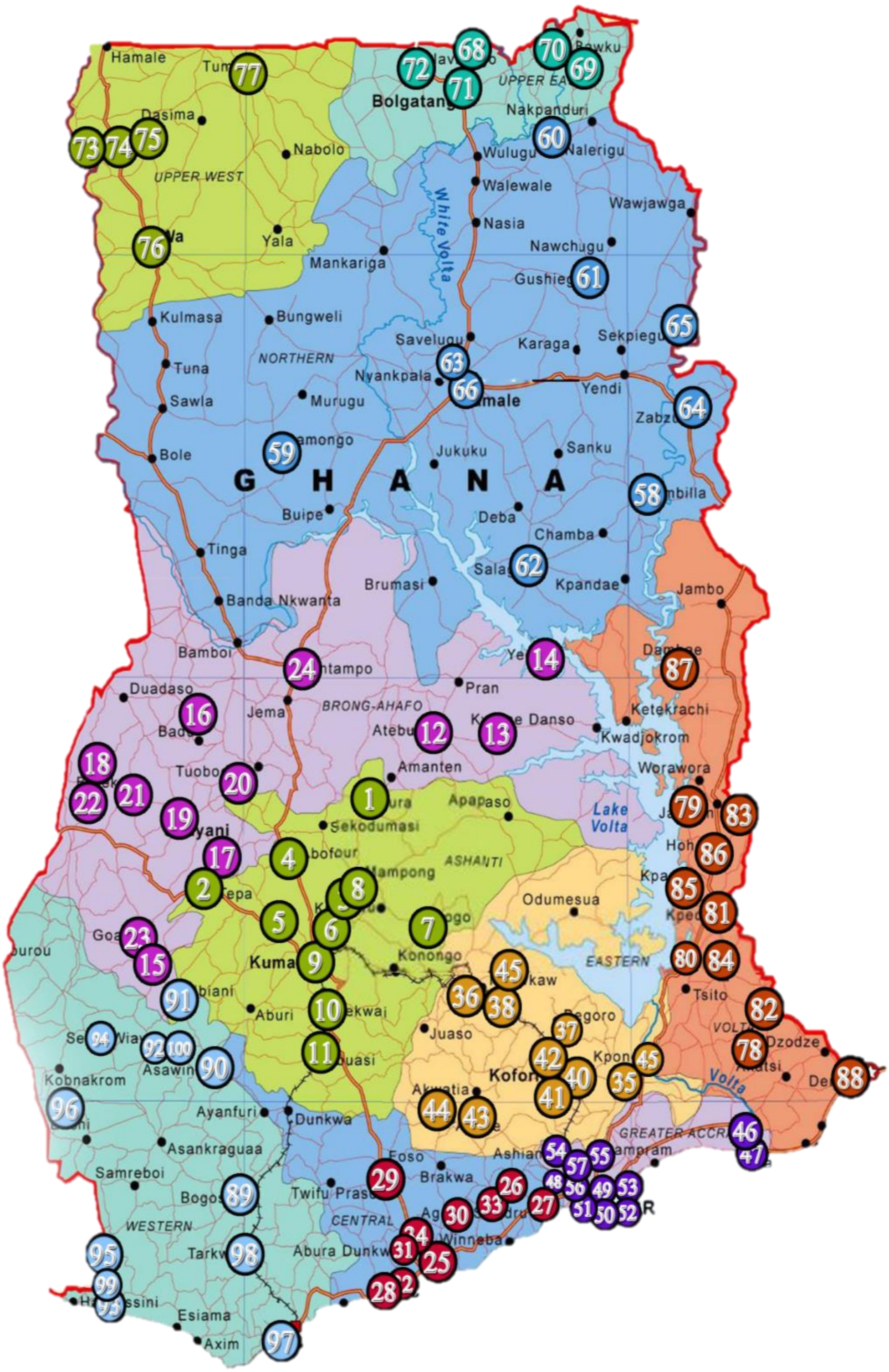


Figure 2: Map of Ghana Showing the Study Markets (Towns) with Dots

Table 6: Selected Markets Shown in the Map of Ghana by their IDs

MARKET ID	NAME OF MARKET	REGION
1	Ejura	Ashanti
2	Tepa	Ashanti
3	Agona	Ashanti
4	Abofour	Ashanti
5	Adugyama	Ashanti
6	Juaben	Ashanti
7	Agogo	Ashanti
8	Nsuta	Ashanti
9	Kumasi (Kejetia)	Ashanti
10	Bekwai	Ashanti
11	Obuasi	Ashanti
12	Atebubu	Brong Ahafo
13	Kwame Danso	Brong Ahafo
14	Yeji	Brong Ahafo
15	Kukuom	Brong Ahafo
16	Nsawkaw	Brong Ahafo
17	D Nkwanta	Brong Ahafo
18	Drobo	Brong Ahafo
19	Sunyani	Brong Ahafo
20	Techiman	Brong Ahafo
21	Brekum	Brong Ahafo
22	Dorman Ahenkro	Brong Ahafo
23	Goaso	Brong Ahafo
24	Kintampo	Brong Ahafo
25	Mankessim	Central
26	Bawjiase	Central
27	Kasoa	Central
28	Elmina	Central
29	Assin Praso	Central
30	Ajumako	Central
31	Fante Nyankomase	Central
32	Cape Coast	Central
33	Swedru	Central
34	Dunkwa	Central
35	Somanya	Eastern
36	Akoase	Eastern
37	Ahoman	Eastern
38	Anyinam	Eastern
39	Agormanya	Eastern
40	Koforidua	Eastern
41	Suhum	Eastern
42	New Tafo	Eastern
43	Asamankese	Eastern
44	Akim Oda	Eastern
45	Mpraeso	Eastern
46	Kaseh	Greater Accra
47	Adah Foah	Greater Accra
48	Hobor	Greater Accra
49	Makola	Greater Accra
50	Agbogbloshie	Greater Accra
51	Kaneshie	Greater Accra
52	Mamprobi	Greater Accra
53	Tema	Greater Accra
54	Ashaiman	Greater Accra
55	Madina	Greater Accra
56	Mallam	Greater Accra
57	Dome	Greater Accra
58	Bimbila	Northern
59	Damongo	Northern
60	Nalerugu	Northern
61	Gushiegu	Northern
62	Salaga	Northern
63	Kumbugu	Northern
64	Zabzugu	Northern
65	Saboba	Northern
66	Tamale	Northern
67	Fumbusi	Upper East
68	Bongo	Upper East
69	Garu	Upper East
70	Zebila	Upper East
71	Bolgatanga	Upper East
72	Navrongo	Upper East
73	Babile	Upper West
74	Jirapa	Upper West
75	Bussie	Upper West
76	Wa	Upper West
77	Tumu	Upper West
78	Mafi-Kumase	Volta
79	Abotoase	Volta
80	Kpeve	Volta
81	Logba Alakpeti	Volta
82	Kpetoe	Volta
83	Kute	Volta
84	Ho	Volta
85	Kpando	Volta
86	Hohoe	Volta
87	Dambai	Volta
88	Denu	Volta
89	Bogoso	Western
90	Asawinso	Western
91	Bibiani	Western
92	Sefwi Dwinase	Western
93	Tikobo	Western
94	Juabeso	Western
95	Elubo	Western
96	Dadieso	Western
97	Takoradi	Western
98	Takwa	Western
99	Agona Nkwanta	Western
100	Sefwi Bekwai	Western

Source: Field Data, SRID-MoFA 2008

3.3 The Institution for the Collection and Compilation of the Data

The SRID is one of the four line directorates of MoFA; and consists of three (3) divisions: Statistics, Research and Information Divisions. The Statistics Division is made up of the Agricultural Surveys and Censuses Unit (ASCU) and the Marketing Services Unit (MSU). The SRID performs the following functions:

i. Collects, computes and analyses prices for various agricultural produce both at the wholesale and retail levels

ii. Monitors producer prices, farm input prices and transport charges for agricultural commodities, and

iii. Publishes weekly and monthly market prices for dissemination to the public through the press houses and to interested individuals and organizations.

Preparations are underway to expand the functions of the Unit to include supply, utilization, exports, imports and the reporting of prices according to quality.

3.4 Procedures of the Market Price Information System

This section discusses the procedure for collecting market information on each of the following essential areas of price monitoring, namely:

i. Wholesale and Retail Price Information

ii. Producer Market Price Information

iii. Agricultural Input Price Information

iv. Transport Charge Information

In line with the concern of this study, emphasis is laid on only information on wholesale prices.

3.5 The Role of a Market Enumerator

The role of a Market Enumerator is to collect and record price data from pre-assigned markets on specially designed form or questionnaire on selected market days. Each Market Enumerator is provided with sufficient supply of forms and is informed of the market(s) assigned and frequency and time of data collection. Thus, they must know about the following:

i. The market(s) from which data is to be collected;
ii. Market day(s) on which the data is to be collected;
iii. Specific frequency or regularity at which, or days on which, the data is to be collected;
iv. The most suitable time of the day for data collection as may apply to a given commodity.

Each Enumerator receives this information from his/her Market Supervisor. Price data is collected through interviewing a few sellers, using the questionnaire or market enumerator's forms. Appendix B shows the enumerator's form.

Selection of Sellers for Interviewing

Since the Market Information System (MakIS) is designed to give objective and impartial information, it is important that the sellers selected for the interview are reliable and can be trusted. It is possible that some sellers may distort price data for fear of taxation or other reasons. Hence an important part of the data collection exercise is to select the right type of seller who will cooperate fully. The major qualities of sellers interviewed will include (1) Willingness to cooperate; (2) Regularity in that market; (3) Handling reasonable quantities of the produce (this will exclude very small scale traders and children selling quantities in the market); and (4) Ability to supply accurate data.

3.6 Interviewing Techniques

In addition to ensuring that the selected seller possesses the above qualities, the approach to the seller is made in such a manner as to encourage willingness to provide the required information. The following are a few tips to guide the interviewers on data collection:

i. Always show the identity of being a MoFA staff.
ii. Outward appearance of interviewer – be properly dressed.
iii. Approach to the sellers should be courteous and friendly.
iv. Explain the purpose and objectives of data collection exercise to the respondents.
v. You may proceed to collect data,
vi. Be grateful to every interviewee after collecting data.

3.7 Data Collection and Processing

The wholesale and retail price data are collected by interviewing three sellers per commodity. If the data is available from less than three sellers no price should be quoted on the questionnaire. A note stating “insufficient to Quote” should be recorded in the remark column (see Appendix B).

It is noteworthy that in many markets, especially in the North, commodities originating from the South get to the markets a day or so before the actual market day. Enumerators are therefore advised to put their ears on the ground so that whenever information reaches them on such arrivals they endeavour to cover them and include the price data on the next market information schedule. In all cases a weighing scale will be required to obtain price information by weight.

Calculating Price for Standard Unit Weight

For each foodstuff there is a space on the form to fill in the unit of sale for which the price is to be recorded. For example, in the case of maize, the wholesale price is to be reported for maxi bag weighing 100 kg, whereas in the case of cassava it is measured in a maxi bag weighing 91 kg. Similarly, the retail price of maize is to be reported on a "kg" basis even though each region has its own unit of measurement as, "Olonka" or calabash bowl. For the Olonka, calabash or other units, a weighing scale should be used to estimate the average weight of the container. If this weight differs from the standard weight a conversion factor must be applied to report prices in standard units.

For example, let us assume that the standard weight for maize is 100 kg but in the market, maize is sold in bags weighing 120 kg at a wholesale price of 24,000 cedis. The price to be recorded on the form should be calculated as follows: 120 kg costs ¢24,000; Thus, 100 kg costs ¢20,000 (i.e. $\frac{100}{120} \times 24{,}000$). So $\frac{100}{120}$ becomes the Conversion Factor. This factor is used to multiply the price collected to obtain the ''True Price'' to record. Thus,

i. $\text{Conversion Factor} = \dfrac{\text{Standard Unit Weight}}{\text{Common Unit in Market}}$; and

ii. $\text{Standard Price} = \text{Price Common in the market} \times \text{Conversion Factor}$

Calculating Daily Average (Simple Mean) Price

After collecting prices from three sellers, average price is calculated and recorded for each food item. The following procedures are used:

(a) Compute a simple mean by adding all prices and dividing by the number of prices recorded.

(b) Round–off to the nearest whole number if average price has decimal attachment.

For example: the average daily price of Maize per bag of 120 kg from three sellers is provided as follows:

Table 7: Average Daily Price of Maize per Bag of 120Kg from Three Sellers

Seller	Price (GHS)/Bag
1	10.9
2	11.6
3	11
Total	**33.5**

Source: SRID-MoFA 2008

Therefore, dividing the total of GHS33.5 by 3 we obtain GH ¢11.167. Then rounding to the nearest whole number we obtain the average daily price of Maize to be GHS11.00.

Calculating Weekly Average Prices

In case prices for a particular market are collected more than once a week. We compute the weekly average price for the market as follows:

i. Compute the daily average price as explained above

Compute weekly average by taking simple average of daily averages and round it off, if necessary. For example, in Table 7, the weekly average price is obtained by simply dividing the total day's average by

Table 8: Weekly Average Price

Market Day	Seller 1 (¢)	Seller 2 (¢)	Seller 3 (¢)	Day's Average (¢)
Thursday	1000	900	1000	966.7
Friday	1200	1100	1200	1166.7
Saturday	1200	1100	1100	1133.3
Total	**3,400**	**3,100**	**3,200**	**3266.7**

Source: SRID-MoFA 2008

This gives ¢1,089.00 (¢3,266.70/3=¢1,088.90) after rounding to the nearest whole number.

Quality Control of Data

It must be emphasized that the quality of data is of utmost importance. The process of quality control begins with the Enumerators themselves. Any negligence or carelessness on the part of an Enumerator can result in serious consequences. It can jeopardize the credibility of the MarkIS. So it is important that at every stage the accuracy of the data is checked and steps taken to correct any errors. If in doubt, it is better not to report doubtful data.

3.8 Variance-Covariance Structure

The variance-covariance matrix is given in Table 9. Generally, the covariance between pairs of the variables is high and this translates into high correlations among the variables. The correlation matrix is given in Appendix C. The high covariances between pairs of variables show that the fifteen items adequately portray the variability in the prices of commodities in the selected markets. The diagonal elements are the variations within the respective items. We observe that Imported Rice has the smallest variation. This is followed by Orange, Banana and Groundnut Oil, in that order. The variations in the other items are quite large. The variation in White Cowpea (WtCp) is the largest,

Table 9: Variance-Covariance Matrix

	Mz	ImpR	WtY	Cv	Org	Ban	Tm	GEg	DrP	RdGn	WtCp	GnOil	PmOil	SmHr	Kobi
Mz	73.221														
ImpR	-2.905	**2.233**													
WtY	53.009	-6.637	217.492												
Cv	-0.920	4.909	-23.275	131.395											
Org	-1.989	0.942	-6.343	10.468	3.086										
Ban	-3.128	1.414	-7.183	18.006	2.893	4.859									
Tm	88.600	-7.250	74.231	22.742	-3.062	-4.096	244.898								
GEg	1.087	5.802	-50.540	62.188	9.868	11.937	-1.507	76.697							
DrP	33.101	-3.285	68.210	35.050	-2.262	1.126	99.647	-10.754	325.125						
RdGn	104.262	-14.129	113.109	-70.046	-15.685	-24.182	169.230	-54.207	77.185	510.554					
WtCp	94.794	-14.850	123.013	-71.381	-17.702	-23.312	172.577	-81.294	129.054	375.045	510.805				
GnOil	13.172	-1.329	18.214	23.657	2.024	2.970	45.417	12.857	22.964	10.377	0.025	41.869			
PmOil	-15.529	6.075	-18.416	64.150	8.808	14.060	-8.213	37.937	14.240	-103.818	-66.236	7.369	86.497		
SmHr	-34.109	9.857	-57.990	92.194	19.153	25.141	-46.297	62.763	-2.429	-170.815	-171.525	7.744	74.764	227.574	
Kobi	-41.545	9.347	-29.812	108.788	16.107	26.071	-25.586	60.431	45.885	-207.177	-159.431	16.714	107.158	158.745	289.808

which is closely followed by Red Groundnut (RdGn). These are followed by Dried Pepper, Kobi, Tomatoes and smoked Herring in that order. There are some three items that have generally low covariance with other items. These are Imported Rice, Orange and Groundnut Oil. It means that these three items do not adequately capture the variations in the price data.

3.9 Descriptive Statistics

The Table 10 gives the basic descriptive statistics of the fifteen commodities. The first column is the variables which are the commodities. The second column is the average price of each of the commodities from the hundred markets. The third column is the standard error of the mean (*SE* Mean), which is obtained simply by dividing the standard deviation in the prices of the particular item and by 10 (i.e. square root of 100). The standard deviation can be obtained from the variance-covariance matrix. If the standard deviation is small, then the *SE* Mean will be small since, in this case the sample size is the same for all commodities. Next is the coefficient of variation. This measure is a property of relative variation which makes it possible to judge the precision of the measurements of the respective average commodity prices. This property is particular relevant since the pricing of the various commodities depends on the manner the commodity is packaged or quantified which differs from commodity to commodity. In the table, the coefficient of variation is obtained by dividing the *SE* Mean by the mean price of the item and then multiplying by 100 (see Chapter Two for explanation). The fifth and sixth columns are the range of prices for each item and the skewness of the distribution of the prices.

Table 10: Basic Descriptive Statistics

Variable	Mean	SE Mean	Coefficient of Variation	Range	Skewness
Mz	50.73	0.86	16.87	38.23	0.04
RiImp	54.63	0.15	2.73	7.64	0.75
YmWt	79.73	1.47	18.50	53.81	-0.64
Cv	18.37	1.15	62.38	52.00	1.15
Org	4.38	0.18	40.09	7.60	0.66
Ban	3.31	0.22	66.58	7.03	0.81
Tm	51.28	1.56	30.52	70.00	0.26
GEg	16.89	0.88	51.85	43.60	1.75
PpDr	60.06	1.80	30.02	87.50	0.06
GnR	106.74	2.26	21.17	121.68	0.32
CpWt	106.37	2.26	21.25	108.30	-0.02
GnOil	39.10	0.65	16.55	36.40	-1.37
PmOil	27.15	0.93	34.25	35.98	0.72
HerSm	36.29	1.51	41.57	60.25	0.94
Kobi	44.55	1.70	38.21	72.00	0.50

As has been pointed out in Chapter Two, the standard errors and the coefficient of variations should be used concurrently to identify the real effect of the spread in the prices of the commodities. The ranges of the prices are reflected in the standard errors of the prices. An obvious observation from Table 10 is that the price of Imported Rice is the most stable across the country.

We also see from column six of the table that with exception of three commodities (White Yam, White Cowpea and Groundnut Oil) distribution of prices of all the other items are positively skewed. It means that in most of the markets, the prices of the respective items are lower than the average price. On the other hand, on White Yam, White Cowpea and Groundnut Oil, most of the markets have prices that are higher than the average prices. Generally, the coefficients of skewness are very low. Thus, the distributions of the prices of the items are quite close to normality.

3.10 Partial Relationship among the Commodity Prices

The unit average price of Banana is GHS 3.31 with a standard deviation of 2.2043 (hence a *SE* of 0.22) has a *CV* of 66.58. On the other hand, Red Groundnut has an average price of GHS 106.74 with a standard deviation of 22.595 (and hence a *SE* of 2.26) and has a *CV* of 21.17 (See Tables 9 and 10). Clearly, the dispersion in the prices of Banana is much lower than in Red Groundnut, which is reflected in their respective standard error. However, the average price of Red Groundnut which is much higher than that of Banana has drastically reduced its *CV* much more than that of Banana. The impression created here by the use of the *CV* is that price of Red Groundnut is more stable than that of Banana, which is not the case. This impression arises as a result of a large average price of one of the commodities both of which are measured in the same unit. It is therefore evident that in a situation where different variables are measured in the same unit but the scale of measurements are different; the use of the standard error may present a more realistic result than the coefficient of variation.

The correlation coefficients (see Appendix C) between pairs of commodities give the linear relationship between the prices and may be seen as the measure of the original relationship between the variables. They are therefore referred to as the zero-order correlation coefficient. For example, in Appendix C we see that the original correlation between Banana and Smoked Herring is 0.756 and it is the highest of all the correlations. Another correlation of interest is that between Palm nut Oil and Smoked Fish which have a correlation coefficient of 0.533. These correlations are influenced by the prices of other commodities in the same market. To determine the true relationship between the prices of any two commodities, it may be necessary to isolate the effect of prices of other commodities on the prices of these two in the same market. The matrix of correlation coefficients between pairs of items after controlling for the other items is the partial correlation coefficients and it is given in Appendix D. Ordinarily, since food items are used in conjunction with other food items, the zero-order correlation is expected to be higher than the partial correlation coefficient. Thus, generally a low element of a partial correlation coefficient matrix is desirable. If there is no change in the two correlations (i.e. their difference is zero), then it implies that the presence or absence of other commodities does not influence the relative prices of the two commodities under consideration. If on the other hand, there is a drastic change in the original correlation after controlling for the other items, (i.e. their difference is large), then it implies that the presence or absence of other commodities does greatly influence the relative prices of the two commodities.

In Table 11, we have the matrix of the difference between the zero-order and partial correlation coefficient matrices. A high positive value in the matrix signifies that after controlling for other items, the (partial) correlation between the two items decreases (i.e. becomes close to zero or more negative than it was). On the other hand, a high negative value signifies that after controlling for other items the (partial) correlation between the two items increases (i.e. it becomes more positive than it was). From Table 11 we see that the differenced correlations between the following pairs of commodities are close to zero: Maize and White Yam; White Yam and Garden Eggs; White Yam and Groundnut Oil; Cassava and Tomatoes; Banana and Tomatoes; Garden Eggs and Tomatoes; Garden Eggs and Dried Pepper; etc. This means that there is not much change in the correlation coefficient between Tomatoes and Cassava, for example, after the effect of the prices of other commodities has been isolated. This implies that the presence or absence of other commodities on the market will not affect the simultaneous purchases of the two commodities. The same applies to White Yam and Groundnut Oil and the other pairs of commodities. We also see that the differenced correlations between the following pairs of commodities are high:

Orange and Cassava; Garden Eggs and Smoked Herrings;

Palm Oil and Smoked Herrings; and Garden Eggs and Banana.

This means that there is drastic change in the correlation coefficient between Orange and Cassava after controlling the effect of other commodities. This implies that the presence of other commodities on the market will highly influence the simultaneous purchases of the two commodities. The same applies to the other pairs of commodities. We observe that all of these high values are positive. This suggests that after controlling for other commodities, the correlation between the respective pairs of commodities decreases.

Table 11: Difference between the Zero-Order and Partial Correlation Matrices

Variable	Mz	ImpR	WtY	Cv	Org	Ban	Tm	GEg	DrP	RdGn	WtCp	GnOil	PmOil	SmHr
ImpR	-0.307													
WtY	0.041	-0.264												
Cv	0.050	0.373	-0.170											
Org	-0.136	0.491	-0.294	0.701										
Ban	-0.292	0.419	-0.201	0.374	0.472									
Tm	0.161	-0.182	0.443	**-0.006**	-0.104	0.035								
GEg	-0.258	0.202	0.016	0.283	0.285	0.574	0.040							
DrP	0.237	-0.128	0.161	0.054	0.018	0.043	0.237	-0.034						
RdGn	0.472	-0.419	0.222	-0.334	-0.440	-0.461	0.358	-0.415	-0.037					
WtCp	0.356	-0.321	-0.422	-0.248	-0.416	-0.447	0.355	-0.202	-0.074	0.206				
GnOil	0.341	0.016	**-0.003**	0.235	0.137	0.115	0.073	0.095	0.101	0.073	0.147			
PmOil	-0.083	0.217	-0.239	0.407	0.437	0.490	-0.171	0.417	0.147	-0.205	**-0.552**	0.185		
SmHr	-0.223	0.284	-0.195	0.437	0.326	0.460	-0.231	**0.647**	0.113	-0.465	-0.407	0.152	**0.637**	
Kobi	-0.081	0.346	-0.259	0.512	0.533	0.488	-0.230	0.389	0.188	-0.353	-0.398	0.162	0.433	0.491

3.11 Markets with Extreme Prices

Table 12 is a 6×5 data matrix that involves prices of five food items from six markets. Our objective here is to illustrate how centroid method of cluster analysis identifies clusters, if any, in this set of data. From the table, it is difficult to identify which groups of markets can be clustered together as forming a distinct group.

The similarities between the markets are summarized by their euclidean distances using Equation (15) in Table 13. Initially, each market is a cluster. Therefore, there are six initial clusters.

Table 13: Data on Prices of Five Food Items from Six Markets

		Food Items				
Market No.	Market	Cassava	Orange	Tomato	Red Groundnut	Smoked Herring
1	Tepa	10.83	3.52	43.58	184.6	32.93
2	Nsuta	4.42	3.65	50.12	120.56	22.46
3	Obuasi	8.25	3.97	44.48	97.67	22.03
4	Kaneshi	21.83	4.67	85.00	134.5	38.83
5	Mamprobi	22.43	3.98	86.30	127.93	35.82
6	Gushiegu	23.10	7.84	38.80	62.92	58.97

We see from Table 13 that the distance between Kaneshi and Mamprobi is the smallest of all the distances. The first cluster containing two members is therefore made up of Kaneshi and Mamprobi (i.e. 4 and 5) at the first step. The number of clusters then reduces to five. These clusters are the markets 1, 2, 3, 6 and (the average of the prices in 4 and 5).

Table 13: Proximity Matrix of Six Selected Markets

Squared Euclidean Distance						
Market	Tepa	Nsuta	Obuasi	Kaneshi	Mamprobi	Gushiegu
Tepa	0.00					
Nsuta	4294.62	0.00				

Squared Euclidean Distance						
Market	Tepa	Nsuta	Obuasi	Kaneshi	Mamprobi	Gushiegu
Obuasi	7683.30	570.72	0.00			
Kaneshi	4382.76	1983.06	3465.47	0.00		
Mamprobi	5179.61	1866.27	3055.82	54.75	0.00	
Gushiegu	15676.17	5149.99	2839.89	7675.42	7033.82	0.00

The reduced data matrix of five clusters is therefore given in Table 14. The proximity matrix for this new set of data is given in Table 14. The average

Table 14: Reduced Data Matrix of Five Clusters

		Food Items				
Market No.	Market	Cassava	Orange	Tomato	Red Groundnut	Smoked Herring
1	Tepa	10.83	3.52	43.58	184.6	32.93
2	Nsuta	4.42	3.65	50.12	120.56	22.46
3	Obuasi	8.25	3.97	44.48	97.67	22.03
4,5	K&M	22.13	4.33	85.65	131.22	37.33
6	Gushiegu	23.10	7.84	38.80	62.92	58.97

prices for Kaneshi and Mamprobi are represented by the market K&M and numbered (4,5). To identify groupings in these five clusters we obtain the proximity matrix for the clusters in Table 15. In this table we see that the squared distance between Obuasi and Nsuta is the smallest. It means that among the five markets, prices in Nsuta and Obuasi are similar and therefore will form a cluster at the next level.

Table 15: Proximity Matrix for Five Clusters

	Squared Euclidean Distance				
Market	Tepa	Nsuta	Obuasi	Gushiegu	K&M
Tepa	0.00				
Nsuta	4295.00	0.00			
Obuasi	7683.00	570.718	0.00		
Gushiegu	15680.00	5150.00	2840.00	.000	
K&M	4767.00	1911.00	3247.00	7341.00	

The number of clusters then reduces to four. These clusters are the markets 1, 6, (the average of prices in 4 and 5) and (average of prices in 2 and 3). The reduced data matrix of four clusters is therefore given in Table 16. The average prices for Nsuta and Obuasi are represented by the market N&O and numbered (2,3).

Table 16: Reduced Data Matrix of Four Clusters

		Food Items				
Market No.	Market	Cassava	Orange	Tomato	Red Groundnut	Smoked Herring
1	Tepa	10.83	3.52	43.58	184.60	32.93
6	Gushiegu	23.1	7.84	38.80	62.92	58.97
4,5	K&M	22.13	4.33	85.65	131.22	37.33
2,3	N&O	6.34	3.81	47.3	109.12	22.25

To identify groupings in these four clusters we obtain the proximity matrix for the four clusters in Table 17. In this table we see that the squared distance between the two averages of Nsuta and Obuasi (N&O) and Kaneshie and Mamprobi (K&M) is the smallest. It means that among the four markets, prices in (N&O) and (K&M) are similar and therefore will form a cluster at the next (third) level. At this level there are three clusters available which are 1, 6, and (average of prices in (N&O) and (K&M)).

Table 17: Proximity Matrix for Four Clusters

	Squared Euclidean Distance			
Market	Tepa	Gushiegu	K&M	N&O
Tepa	0.00			
Gushiegu	15680.00	0.00		
K&M	4767.00	7341.00	0.00	
N&O	5845.00	3852.00	2436.00	0.00

The reduced data matrix of three clusters is given in Table 18. The average prices for (N&O) and (K&M) are represented by the market (NO)&(KM) and numbered (2,3) & (4,5).

Table 18: Reduced Data Matrix of Three Clusters

		Food Items				
Market No.	Market	Cassava	Orange	Tomato	Red Groundnut	Smoked Herring
1	Tepa	10.83	3.52	43.58	184.60	32.93
6	Gushiegu	23.10	7.84	38.80	62.92	58.97
(2,3)&(4,5)	(NO)&(KM)	14.24	4.07	66.48	120.17	29.79

To identify groupings in these three clusters we obtain the proximity matrix in Table 19 for the three clusters in Table 18. We see that clusters comprising of (NO)&(KM) and Tepa have the smallest distance. These two clusters are therefore combined to form a new cluster of markets 1&(2,3)&(4,5). The other cluster consists of Gushiegu.

Table 19: Proximity Matrix for Three Clusters

	Squared Euclidean Distance		
Market	Tepa	Gushiegu	(NO)&(KM)
Tepa	0.00		
Gushiegu	15680.00	0.00	
(NO)&(KM)	4697.00	4988.00	0.00

Obviously, the next step is to group all the markets into one cluster. The various steps or stages of the clustering procedure can be represented graphically in what is called a dendrogram. Figure 3 gives the dendrogram for the illustration. In Figure 3, it is easy to identify three distinct clusters among the initial six markets. Tepa market and Gushiegu market (Gushg) each forms a distinct cluster whilst the third cluster will be made up of Kaneshi (Knsh), Mamprobi (Mprb), Nsuta and Obuasi (Obua). Based on the knowledge of the subject matter (in this case the distribution of prices in the various markets) we can assign interpretation to the clusters. For example, the Tepa market cluster may be labelled as a low-priced market. Similarly, the Gushiegu market cluster may be labelled as high-priced market, and the third cluster will be labelled as moderate-priced market.

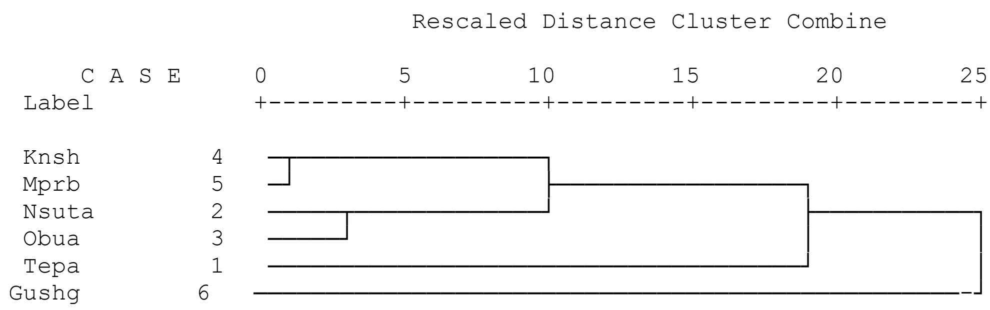


Figure 3: Dendrogram Illustrating the Centroid Method of Hierarchical Clustering

The Table 20 gives a summary of the results of box plots generated from the preliminary analysis of the data. We see from the table that Ashanti region has four markets that record the highest prices in four commodities (Maize, Red Groundnut, White Cowpea and Groundnut Oil), but has one market that records the lowest price in one item (Dried Pepper). Interestingly, all of these markets are in the rural areas. The Upper West also has three markets that record the highest prices in three items (White Yam, Banana and Kobi). Similarly, the Upper East has three markets that record highest prices in three items (Imported Rice, Garden Eggs and Palm Nut Oil). The Northern region has one market that records the highest price in an item (Orange). However, the Northern region has a market that has the lowest prices in two items (Red Groundnut and White Cowpea). It is the only region with the same market that has extreme prices in more than one item. The Eastern region has two markets with highest prices in two items (Cassava and Dried Pepper) and also has two markets with the lowest price in one item (Orange and Banana). Greater Accra has one market that has the highest price in one item (Tomatoes), but it also has a market with the lowest price in an item (Garden Eggs).

Table 20: Highest-Priced and Lowest-Priced Markets in each Food Item

A market in Brong Ahafo region records the highest price in one item (Smoked Herrings).

However, the Brong Ahafo region alone has five different markets that record lowest prices in five commodities. These commodities are (Maize, Cassava, Tomatoes, Groundnut Oil and Palm Oil). The Volta region has two

		Extreme Markets			
		Highest-Priced		Lowest-priced	
Number	Food Item	Market (Type)	Region	Market (Type)	Region
1	Maize	Obuasi (U)	Ashanti	Kintampo (U)	Brong Ahafo
2	Imported Rice	Fumbisie (R)	Upper East	Fante Nyankomase (R)	Central
3	White Yam	Tumu (U)	Upper West	Dambai (U)	Volta
4	Cassava	Akim Oda (U)	Eastern	Dormaa Ahenkro (U)	Brong Ahafo
5	Orange	Bimbila (R)	Northern	Ahoman (R)	Eastern
6	Banana	Jirapa (R)	Upper West	Dunkwa (U) & Anyinam (R)	Central & Eastern
7	Tomatoes	Tema (U)	Greater Accra	Yeji (R)	Brong Ahafo
8	Garden Eggs	Bongo (R)	Upper East	Mamprobi (U)	Greater Accra
9	Dried Pepper	Suhum (U)	Eastern	Nsuta (R)	Ashanti
10	Red Groundnut	Nsuta (R)	Ashanti	Gushiegu (R)	Northern
11	White Cowpea	Tepa (R)	Ashanti	Gushiegu (R)	Northern
12	Groundnut Oil	Agona (R)	Ashanti	Yeji (R)	Brong Ahafo
13	Palm Nut Oil	Bolgatanga(U)	Upper East	Kukuom (R)	Brong Ahafo
14	Smoked Herring	Yeji (R)	Brong Ahafo	Hohoe (U)	Volta
15	Kobi	Wa (U)	Upper West	Fante Nyankomase (R)	Central

markets with lowest prices in two items (White Yam and Smoked Herrings). There are three markets in the Central region that have the lowest prices in three items (Imported Rice, and Banana and Kobi).

In all, there are seven of the ten regions where markets have highest prices in one commodity or the other, whilst equal numbers of regions have markets with the lowest prices in commodities. Central, Western and Volta are the regions that do not record highest prices in any of the items. Western region is also among the three regions that do not record lowest prices in any of the items. The other two are Upper East and West. Thus, the Western region is the only region that does not record extreme prices in commodities.

In the clustered bar chart in Figure 4, the Pricetype represents highest (*H*) or lowest (*L*) prices recorded for a commodity in the market in each region. The Marketype represents the Rural or Urban area in which the market is located in the Region. We see from Figure 4 that, for example, the Greater Accra has no rural market that records any extreme prices. The same can be said of the Volta and Western regions. Similarly, there is no urban market in the Northern region that record extreme prices, and the same is true for the Western. The Western region is conspicuously one with no market where extreme prices are recorded. It is clear that the Brong Ahafo region appears to be one region that is notable for extremes in prices in various commodities: it has four rural markets three of which record lowest prices in three items and one recording a highest price in an item. However, it has two urban markets both of which record lowest prices in two items. The number of markets in the rural and urban areas in the other regions can be read in a similar manner from the graph.

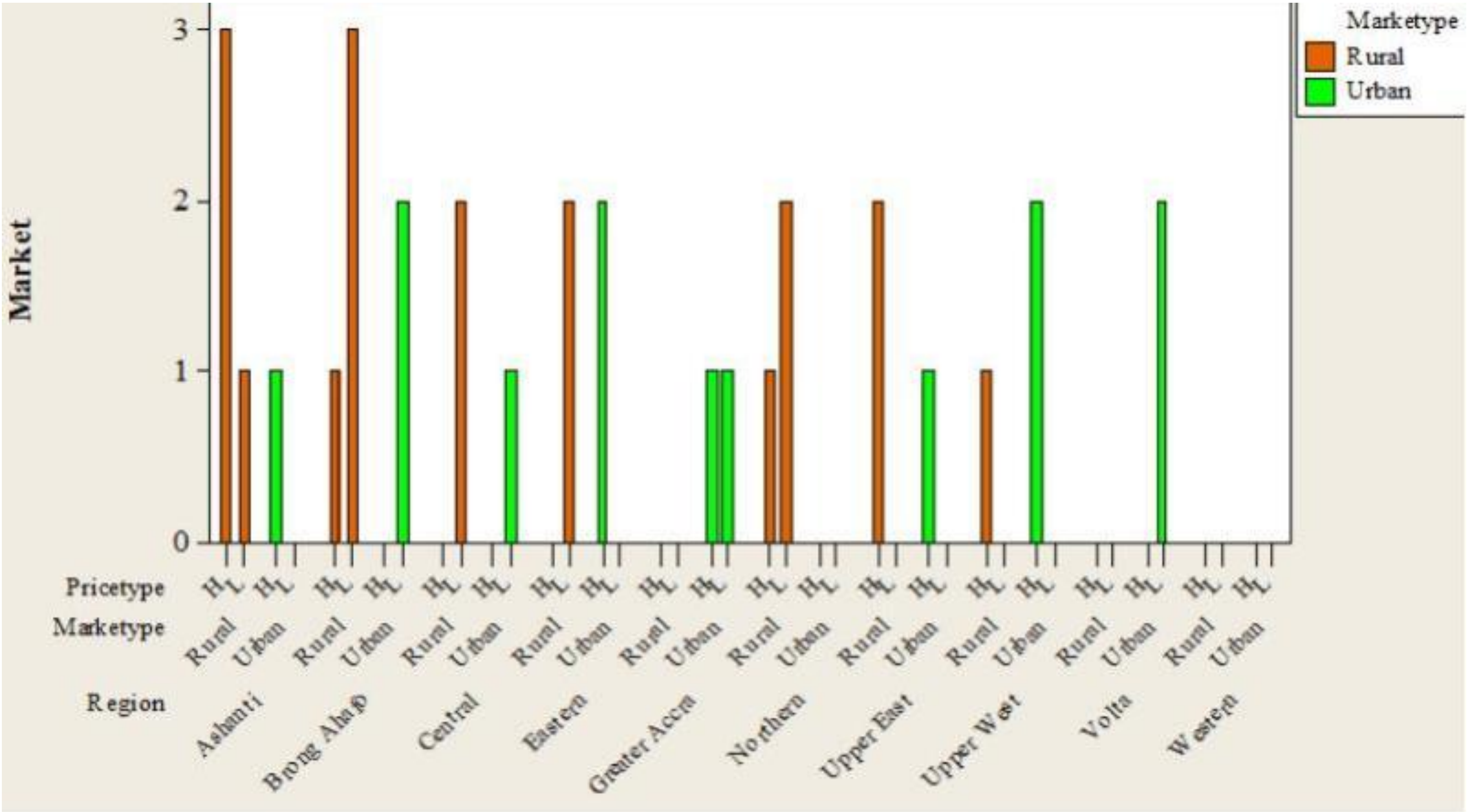


Figure 4: Distribution of Markets by Location, Type of Market and Type of Extreme Price of Commodities

3.12 Summary of Preliminary Analysis

This chapter made use of some routine multivariate exploratory techniques to analyze the data on prices of fifteen food items selected from various markets across the country. The number of markets selected from the various regions is in proportion to the actual number of markets in the regions. These selected markets were identified by whether they are located in a rural or an urban area of the region. Out of the 100 selected markets, 59 are located in a rural area and 41 in urban area.

The variance-covariance matrix of the price data was examined. It has been observed that there are generally high covariances among the items. This suggests that the selected items adequately sum up the variation in prices across the country. Within the individual items, Imported Rice is the item with the least variation in prices. Other items that have less price variability are Orange, Banana, and Groundnut Oil. On the other hand, White Cowpea and Red Groundnut have the highest variability in prices. Other items that have high variation in prices are Kobi and Dried Pepper.

The chapter has also examined the zero order and partial correlation coefficients matrices as well as the differences between the two matrices. There are a number of very low and high differences in the two correlations coefficients among some items. It is observed that the correlation differences are low for the following pairs of items: Maize and White Yam; White Yam and Garden Eggs; White Yam and Groundnut Oil; Cassava and Tomatoes; Banana and Tomatoes; Garden Eggs and Tomatoes; Garden Eggs and Dried Pepper. Palm Oil and Smoked Herrings; and between Garden Eggs and Banana. This means that in a market with a variety of other food items, the prices of such items are likely to reduce. Thus the simultaneous purchases of a combination of these items are likely to reduce as the consumer may switch to other items. Generally, there are low partial correlations among the items which buttress the fact that the selected items adequately represent the variability in the prices of food items across the country.

Further exploratory analysis shows that there are markets in the various regions in which prices of the fifteen items are in the extremes. It has come to light that the Brong Ahafo region has the leading number of six markets that record extreme prices in six different items, with only one recording the highest price in an item (Smoked Herring) and the other five recording lowest prices. It thus appears that the Brong Ahafo is the region with generally low food prices. In contrast, the Ashanti region has five markets recording extreme prices, mostly highest prices, with only one of the five markets (Nsuta) recording lowest price in an item (Dried Pepper). All regions have markets that record extreme prices in one commodity or another with exception of only one, the Western region, which has no market that record extreme prices. The Upper East and West regions have markets (six in all) that record only highest prices but not lowest prices. Thus it appears that these two regions have markets that are consistently the highest-priced. The Northern region is the only region that has a single market that record extreme prices (lowest prices) in more than one item (White Cowpea and Red Groundnut). It therefore appears that markets in the Western region have the least varied prices, whilst markets in the Ashanti and Brong Ahafo regions have the most varied prices. This means that the prices of these items are not influenced by the presence or absence of a variety of other items in the same market.

CHAPTER FOUR

FURTHER ANALYSIS

A number of results in the previous chapter appear to inform the use of few dimensions along which to evaluate the level of prices of the selected food items in order to identify specific markets in which items are high-priced and those that are low-priced. For example, it has been observed that the covariances between pairs of items are generally high. It has also been observed that markets in the Northern, Upper East and West regions are consistently high-priced, whilst markets in the Brong Ahafo and the Ashanti regions generally low-priced. However, the extreme price levels in these markets have been identified for single food items. It is necessary therefore to identify those markets with specific price-levels for certain categories of items. This can only be done by means of a multivariate approach.

In the next section we will further examine the suitability of the data for dimensionality reduction. We will examine, in particular the eigen-analysis of the data that will hopefully inform the use of principal component analysis in the subsequent sections. After extracting the relevant components, we will proceed to perform cluster analysis of the component scores in order to identify specific markets with notable price levels.

4.1 Suitability of the Data for Dimensionality-Reduction

In this section we make use of procedures that are primarily aimed at examining the variance-covariance structure of the data under study. These will include the eigen-analysis of the variance-covariance matrix and the Kaiser-Meyer-Olkin's measure of sampling adequacy.

The Table 21 gives measures for computing the KMO which is the ratio of sum of squared correlation coefficients to the same and sum of squared partial correlation coefficients. The Table also gives the value of the Bartlett's statistic. The matrices of zero-order and partial correlation coefficients are provided in Appendix C and D.

Table 21: KMO and Bartlett's Test of Sphericity

Measures	Value	Significance
Bartlett's Statistic	867.907	0.0000
Sum of squared correlation coefficients $\left(\sum r_{ij}^2\right)$	16.3955	
Sum of squared partial correlation coefficients $\left(\sum pr_{ij}^2\right)$	3.4778	
Total $= \sum r_{ij}^2 + \sum pr_{ij}^2$	19.8733	
Ratio of $\sum r_{ij}^2$ to the Total	**0.8250**	

The value of the Bartlett's test statistic is obtained for the test of the hypothesis:

H

$\boldsymbol{H_o}$**:** There are no specific dimensions along which to evaluate price levels of markets.

Clearly, this value is very large, so large that there is enough evidence to reject the null hypothesis. This is confirmed by the small *p*-value ($p - value << 0.05$) associated with the test. Therefore, there is reason to believe that there are clear dimensions along which the markets can be evaluated using the prices of the commodities.

Using Equation (4), we obtain the KMO measure of sampling adequacy as 0.8250, which is high. This index indicates that the fifteen variables adequately summarize the degree of common variation in food prices.

4.2 Principal Components Analysis of the Data

We then carry out Principal Component Analysis on the fifteen predictors in the data set to repackage the variance from p correlated variables into p uncorrelated components. The principal components are therefore weighted sums of the original variables. As a result, PCA may be used to find weighted sums or index of prices of food items.

Eigenvalues and Proportions of Variance Explained by each Component

The Table 22 gives the eigenvalues of each component along with the percentage (and cumulative totals) of the total variance explained by that component. As expected, the percentage of variation accounted for by the components decreased sequentially from the first through to the last component. The first component, (Y_1) accounts for almost 40% of variability in the set of fifteen predictor variables while the second (Y_2) and third components respectively account for about 19% and 8% of the total variation. The high percentage of variance explained by the first component alone shows that the first component might be a very influential single factor in the determination of the commodity-price ratings of the various markets under study. The second component accounts for a percentage variance that is less than half of what is accounted for by the first component. This shows that the dimension along which Y_2 evaluates the price levels of the markets may be

Table 22: Eigenvalues and Proportion of Variance Explained by Components

Component	Eigenvalues	Percent Variance	Cumulative percent Variance
1	5.985	39.899	39.899
2	2.784	18.562	58.461
3	1.157	7.716	66.176
4	0.901	6.008	72.184
5	0.757	5.044	77.228
6	0.646	4.304	81.532
7	0.546	3.640	85.172
8	0.436	2.906	88.078
9	0.421	2.807	90.885
10	0.345	2.300	93.185
11	0.291	1.938	95.123
12	0.234	1.561	96.684
13	0.177	1.183	97.867
14	0.174	1.161	99.028
15	0.146	0.972	100.000
Total	15.000	100.000	

much more restricted than in the case of the first. The third component also accounts for percentage variation that is less than half of that explained by the second. This also shows that the dimension along which we can evaluate price levels of the markets may be more restricted than in the case of the second component.

As has been pointed out, the components provide increasingly restricted dimensions for assessing the price levels in the various markets. It may therefore not be expedient to continue extracting components beyond component(s) that will not be in line with the objectives of the study.

From Table 22, only the first three components have eigenvalues greater than one. This suggests that the only components that will adequately explain variations in the prices of commodities are these first three components. The effects of these three components are better appreciated with a graphical illustration in Figure 5.

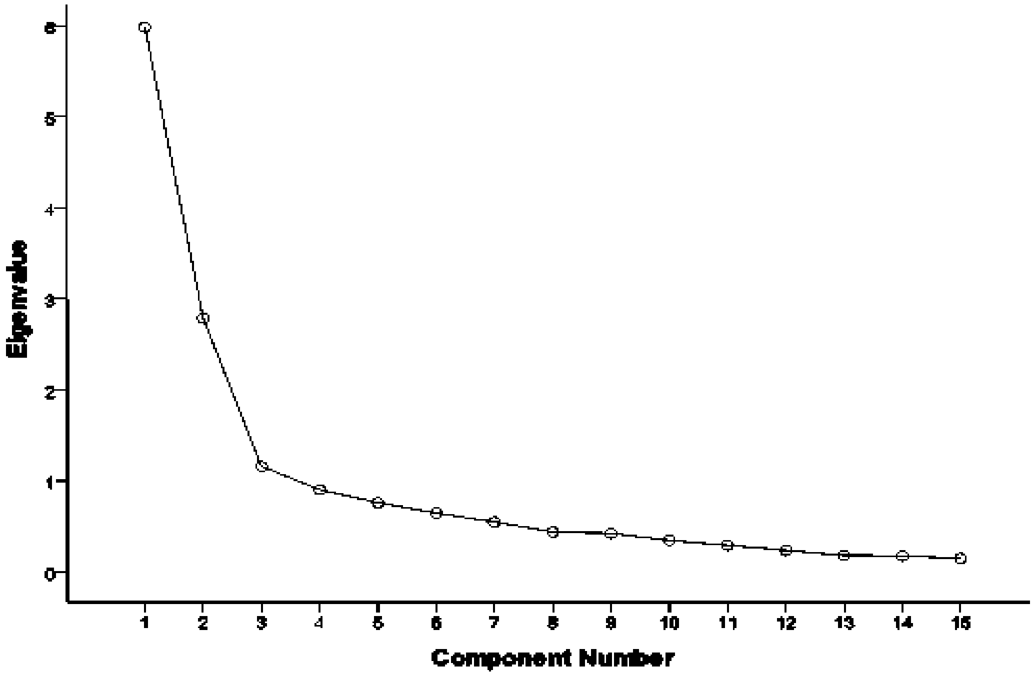


Figure 5: Scree Plot

The Figure 5 shows the plot of the eigenvalues against corresponding components. This plot is what is often referred to as the scree plot. The ‘‘elbow’’ point of the curve is clearly observed to be at the third component. Beyond this point the plot begins to level off. The line is nearly horizontal from the third point because after this point all the components explain approximately the same amount of variability, which is not significant. By this criterion, it means that it may not be necessary to consider components that are beyond three. Even though the location of the elbow point in the scree plot does not border so much on subjectivity, it may be necessary to use some confirmatory procedure to help determine the actual number of components that are likely to provide meaningful dimensions for evaluating markets. One of such procedures is the Monte Carlo Principal Component Analysis for Parallel Analysis. Figure 6 gives the plot of the criterion values (or the random eigenvalue) against component overlaid on the scree plot. From the Figure 6, we observe that the parallel analysis plot cuts the scree plot just above the third component. Thus, the number of eigenvalues that are greater than the criterion values are two. This shows that there are possibly only two components that may be used to meaningfully assess market prices.

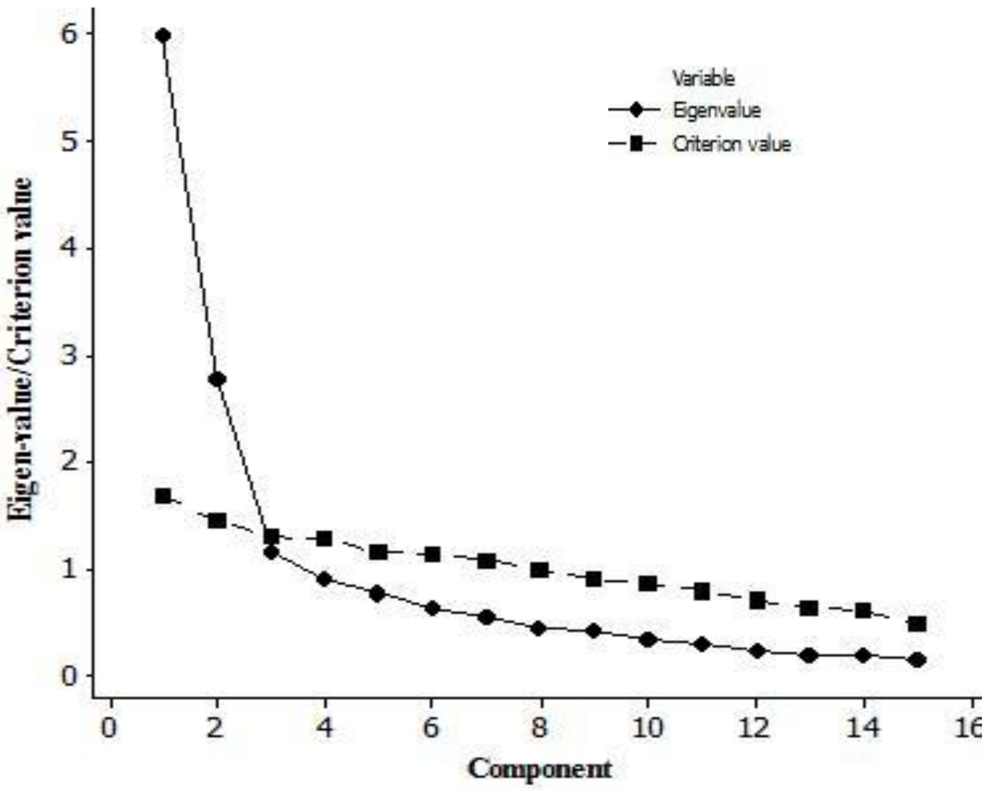


Figure 6: Plot of Eigen-value/criterion value against component

4.3 Extraction and Interpretation of Principal Components

In this section we make use of the Principal Component technique to extract component that may be considered as reflecting the general prices of all or some selected groups of food commodities. In the extraction and interpretation of components, some researchers (Larose, 2006; Steven, 1996) have suggested that the loading or weight of the variable on the component should exceed a magnitude (absolute) of 0.4 in order for the variable to have a practical significance in the formation of the component. In this study, we make use of a loading of 0.4 as a cut-off value in order to aid the interpretation of the components.

The Table 23 shows the unrotated loadings of the components on each of the fifteen commodity variables. We see from the table that using the cut-off value of magnitude 0.4, the first principal component loads highly on all the variables with exception of three items which are Tomatoes, Dried pepper and Groundnut Oil. Thus, we will label the first principal component as a general price index. Thus, almost all the variables are influential in the formation of the first component. We will then label the first component as the *weighted sum of all the commodities*. This means that the dimension provided by the first component enables us to obtain a general assessment of the price levels of commodities in the markets using all the commodities under consideration. This is in line with the observation made from the percentage of variation accounted for by this component in Table 22. That is, using the prices of all the commodities in evaluating the price-levels in the markets appears to be the most effective monitoring mechanism of prices.

Table 23: Unrotated Component Matrix

No.	Variables	Components		
		1	2	3
1	Maize	-0.422	0.657	-0.312
2	Imported Rice	0.613	-0.169	-0.113
3	White Yam	-0.429	0.411	0.431
4	Cassava	0.673	0.493	-0.041
5	Orange	0.779	0.215	-0.188
6	Banana	0.865	0.280	0.016
7	Tomatoes	-0.338	0.787	-0.131
8	Garden Egg	0.692	0.270	-0.502
9	Dried Pepper	-0.107	0.544	0.525
10	Red Groundnut	-0.725	0.366	-0.246
11	White Cowpea	-0.711	0.382	0.058
12	Groundnut Oil	0.095	0.633	-0.094
13	Palm Nut Oil	0.743	0.251	0.238
14	Smoked Herring	0.814	0.102	0.055
15	Kobi (Salted Tilapia)	0.761	0.210	0.384

The formation of the second component is highly influenced by six of the fifteen variables. These are Maize, White Yam, Cassava, Tomatoes, Dried Pepper and Groundnut Oil. We observe that these items constitute a typical Ghanaian primary diet. We will therefore label this component as a *local diet component*. Thus, the second component provides a dimension that evaluates the general price-levels of the markets by considering only those commodities that are main constituents of the diet of typical Ghanaian folks.

The formation of the third component is influenced by White Yam, Garden Eggs, and Dried Pepper. It is not clear what these items collectively represent. Moreover, all of these items have been identified on the two earlier components. As a result, it will not be necessary to attempt to assign a label to this component. Thus, we are able to identify only two components that provide meaningful dimensions for evaluating price levels in the markets. This observation is very much in line with the result of the parallel analysis in Figure 22. The table of all principal components is given Appendix E.

It may be necessary to examine the rotation of the components in Table 23. However, we have the Varimax rotation of Table 23 in Appendix F. The principal components provided in that Table in Appendix F do not provide meaningful indices for evaluating the markets in line with the objectives of the study. Consequently, the unrotated components are preferred to the rotated.

4.4 Identification of Market Groups Based on Price Levels

In this section we will make use of the two components that have been extracted in the previous section to categorize the markets into five main groupings. These groups are the *Highest, High, Moderate, Low and Lowest - Priced markets*. We make use of scatter plot of the principal component to identify the categories of the markets. In the first part of the section we examine how the classifications are done using the first principal component. Thus, the categorization using the first principal component will identify markets groups with reference to the general price levels of the commodities. The categorization using the second principal component will identify markets groups with reference to the few commodities that constitute the local diet in a typical Ghanaian society. The use of the principal component scores in identifying groupings among subjects have been adequately reviewed in Chapter Two. In that chapter we have learnt that scores that are close to zero reflect original markets in which prices are just about the average in all the markets. Thus, all the markets on the origin and between the intervals from -1 to 1 are markets that have moderate (average) prices. To help us identify the locations of the markets in the various categories we reproduce summary of Appendix A in Table 24.

Table 4.2 Location of Markets in Various Regions

Market Number	Region
1 – 11	Ashanti
12 – 24	Brong Ahafo
25 – 34	Central
35 – 45	Eastern
46 – 57	Greater Accra
58– 66	Northern
67 – 72	Upper East
73 – 77	Upper West
78 – 88	Volta
89 – 100	Western

4.4.1 Identifying Market Groups Based on the First Principal Component

The plot of the first principal component scores (see Appendix G) is given in Figure 23. In the figure, there are six zones demarcated by five horizontal lines. Using Table 24, we see from Figure 23 that markets that are moderate in prices on all commodities are markets that are generally located in the Ashanti, Brong Ahafo, Central, Eastern, Greater Accra, Volta and Western regions. Thus, out of the ten regions, markets in seven of them are considered as having moderate prices.

The markets that are in the zone specified by the interval from -1 to -2 are considered as *Low-Priced* since their prices are a little less than the average prices. These low-priced markets are summarized in Table 25. There are eight markets in all in this category out of which half of them are located in the Western region. It is interesting to note that half of these markets are urban markets.

Table 25: Low-priced Markets and their Locations

Market Number	Town	Region	Type of Market
5	Adugyama	Ashanti	Rural
42	Ntafo	Eastern	Urban
51	Kaneshie	Greater Accra	Urban
52	Mamprobi	Greater Accra	Urban
89	Bogoso	Western	Rural
90	Asawanse	Western	Rural
91	Bibiani	Western	Rural
98	Tarkwa	Western	Urban

Markets that are in the zone specified by the interval from -2 and below are considered as the *Lowest-Priced* markets since their prices are much less than the average prices. Only one market is located in this category. This is identified as Tepa market which is a rural market in the Ashanti region.

Markets that are in the zone specified by the interval from 1 to 2 are considered as *High-Priced* since their prices are a little higher than the average prices. These are predominantly markets that are located in the Northern region and the Upper East region. A few markets in the Upper West are also *High-Priced*. One of the high-priced markets is located in the Brong Ahafo region. This market is Yeji which is a rural market. Markets that are in the zone within the interval from 2 and above are considered as the *Highest-Priced* markets since their prices are much higher than the average prices. There are four markets in this category. These are provided in Table 4.6.

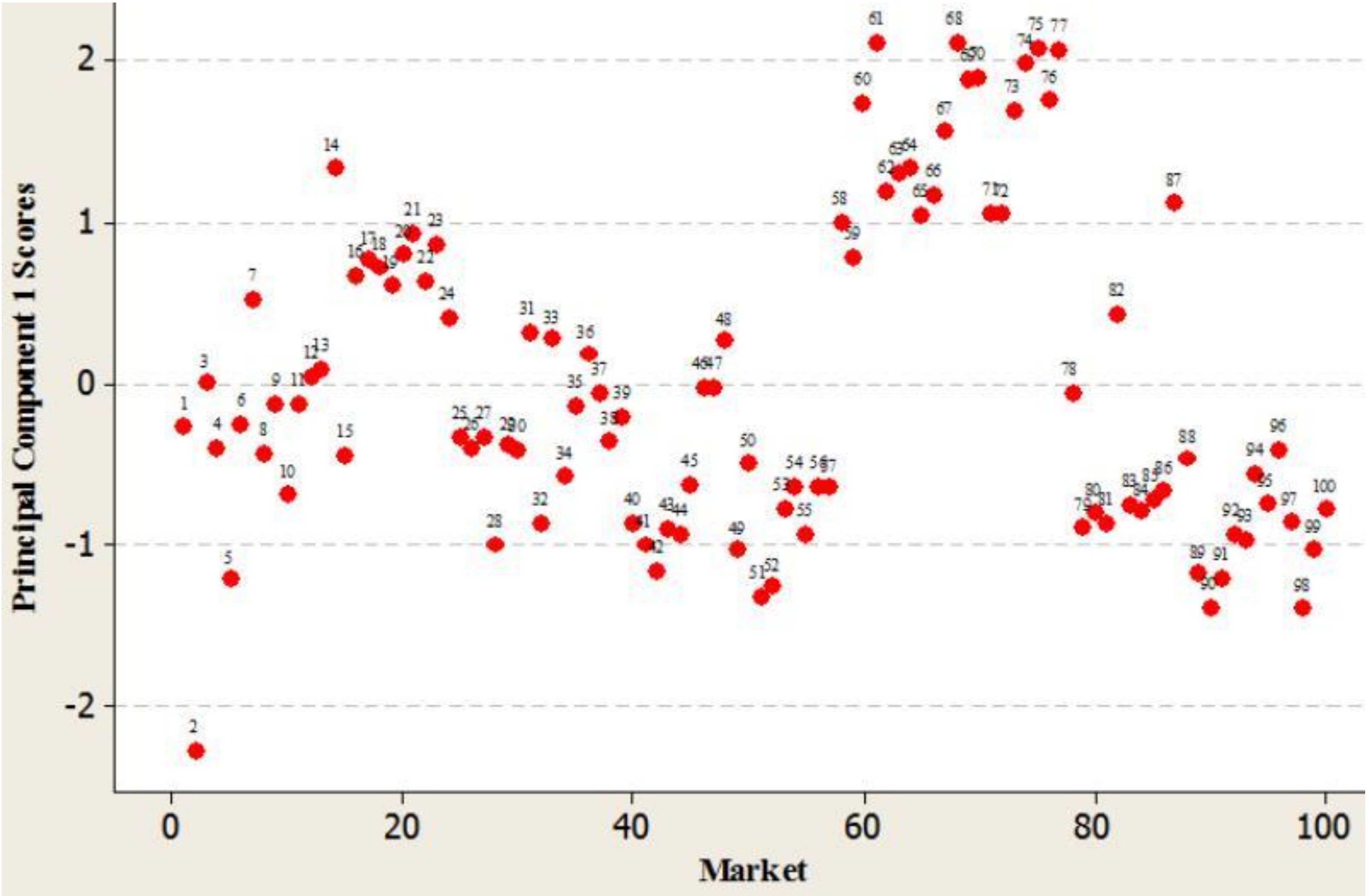


Figure 7: Plot of the First Principal Component Scores

From the table we see that the markets in which the commodities are highest-priced are all located in the upper part of the country.

Table 26: Highest-Priced Markets and their Locations

Market Number	Town	Region	Type of Market
61	Gushiegu	Northern	Rural
68	Bongo	Upper East	Rural
74	Jirapa	Upper West	Rural
75	Busie	Upper West	Rural
77	Tumu	Upper West	Urban

4.4.2 Identifying Market Groups Based on the Second Principal Component

The plot of the second principal component scores (see Appendix G) is given in Figure 24. As in Figure 23, there are six zones demarcated by five horizontal lines. Markets that are moderate in prices of the commodities represented by the second component are those around the origin within the interval -1 to 1. Using Table 24, we see from Figure 24 that these markets include all the markets in the Volta Western, Central and Northern regions, and some markets in the Upper East and West, Greater Accra, Eastern regions, Brong Ahafo and Ashanti regions. Thus, with reference to the principal component two (PC 2), there are markets in all the regions that are *Moderately-Priced*.

Markets that are low-priced on PC 2 are within the area specified by the interval from -1 to -2. These are markets that are predominantly located in the Ashanti and Brong Ahafo regions.

Markets that are in the area specified by the interval from -2 and below are considered as the lowest-priced markets on PC 2. Only two markets are located in this category. These are identified as Agogo and Nsuta markets both in the rural areas of the Ashanti region.

Markets that are in the area specified by the interval from 1 to 2 are considered as high-priced with respect to PC 2. These include some markets in the Eastern, Greater Accra, Upper East and West regions.

We observe that there is no market located beyond a PC value of 2. This means that with respect to PC 2, there is no market that may be considered as highest-priced.

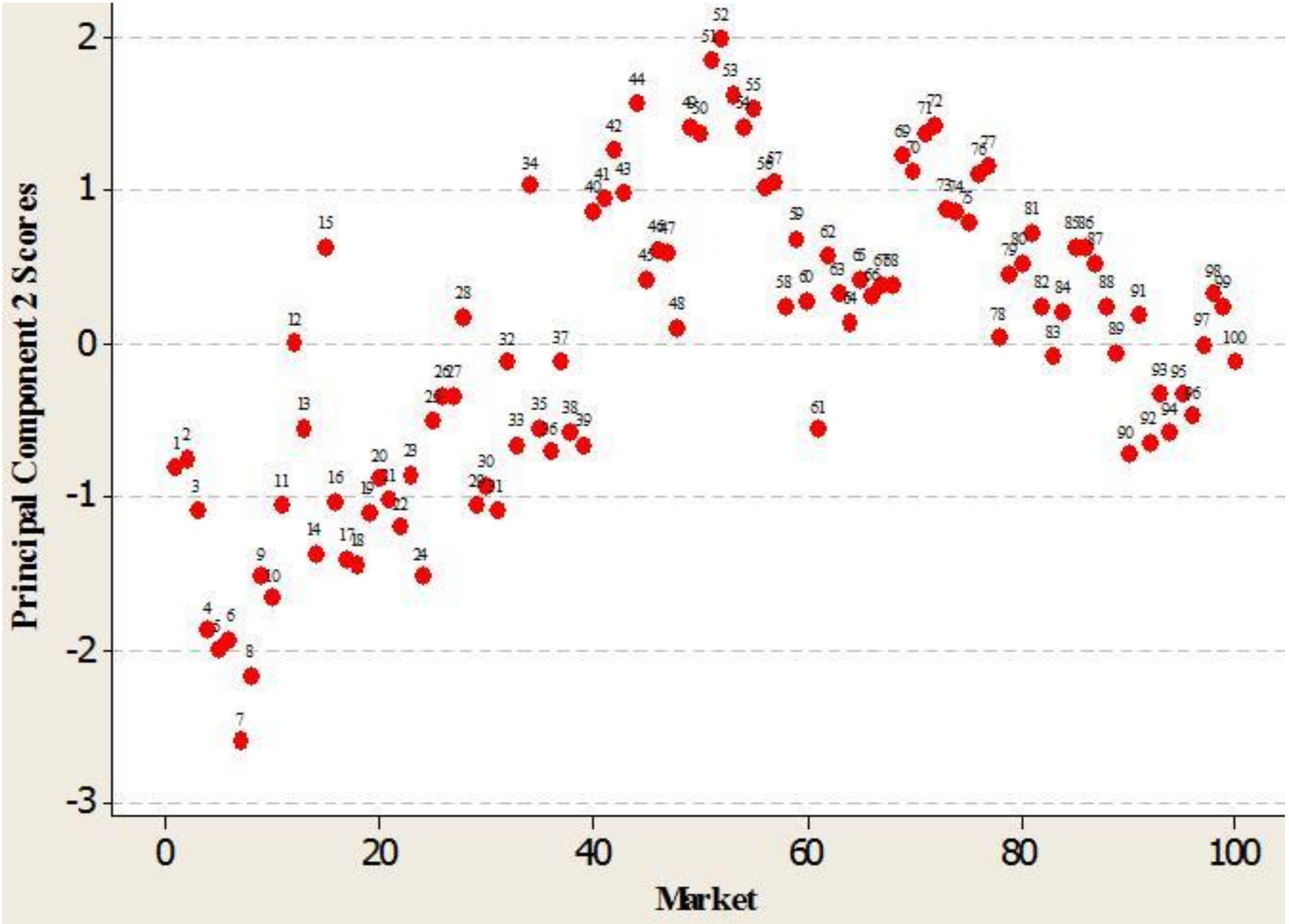


Figure 8: Plot of the Second Principal Component Scores

4.5 Cluster Analysis

In this section we present the result of cluster analysis of the first and second principal component scores using the Centroid method. As we know, the summary of the cluster procedure is given by the dendrogram. However, for large data set such as the one used in this study, the dendrogram is too cumbersome to present here. Consequently, we will present a graphical presentation of the market cluster membership. The dendrograms for both the first and second components are given in Appendix H and Appendix I. The cluster memberships that have been illustrated in the graphs in this section have been provided in Appendix J and K.

4.5.1 Market Clustering Using the First Principal Component

The dendrogram showing all the possible clusters using the first principal component is given in Appendix H. From that diagram five distinct clusters can be identified. A plot of the five market-cluster membership is given in Figure 9. The first cluster is one of the clusters with large membership. It is made up of markets in seven regions, with only one market from the Western region. This market from the Western region is numbered 96. Other noticeable members are markets number 78 and 88, which are located in the Volta region. Other members are eight of the eleven markets located in the Ashanti region. From the previous section we can use these identified markets to label this cluster as *Moderate-Priced market*.

The second cluster has the smallest membership which is made up of a single market located in the Ashanti region. This market is numbered 2 and it identified from the previous section as Tepa, a rural market. This cluster can be labelled as *Lowest-Priced market*.

The third cluster is one of the clusters with large membership. It is made up of markets in six regions, and involves almost all the markets located in the Western region. Other members that can help identify this cluster are the two markets from the Ashanti region. From the previous section, we can label this cluster as a *Low-Priced market*.

The membership of the fourth cluster is made up of market four regions, with only one market (numbered 87) from the Volta region and two from the Upper East region. The other two regions with large representations are Brong Ahafo and the Northern regions. We will label this cluster as *High-Priced market*.

The membership of the fifth cluster is second smallest. Its membership is made up of markets from only three regions, which includes all the markets from the Upper West region. The other markets are from the Northern and Upper East regions. This cluster may be labelled as the *Highest-Priced market*.

The dendrogram showing all the possible clusters using the second principal component is given in Appendix I. From that diagram only three distinct clusters can be identified. A plot of the three markets cluster membership is given in Figure 10. The first cluster is made up of markets in six regions, with only one market from the Northern region. This market from the Northern region is numbered 61, and from the previous section it is identified as Gushiegu, a rural market. We will label this cluster as *Moderate-Priced market*.

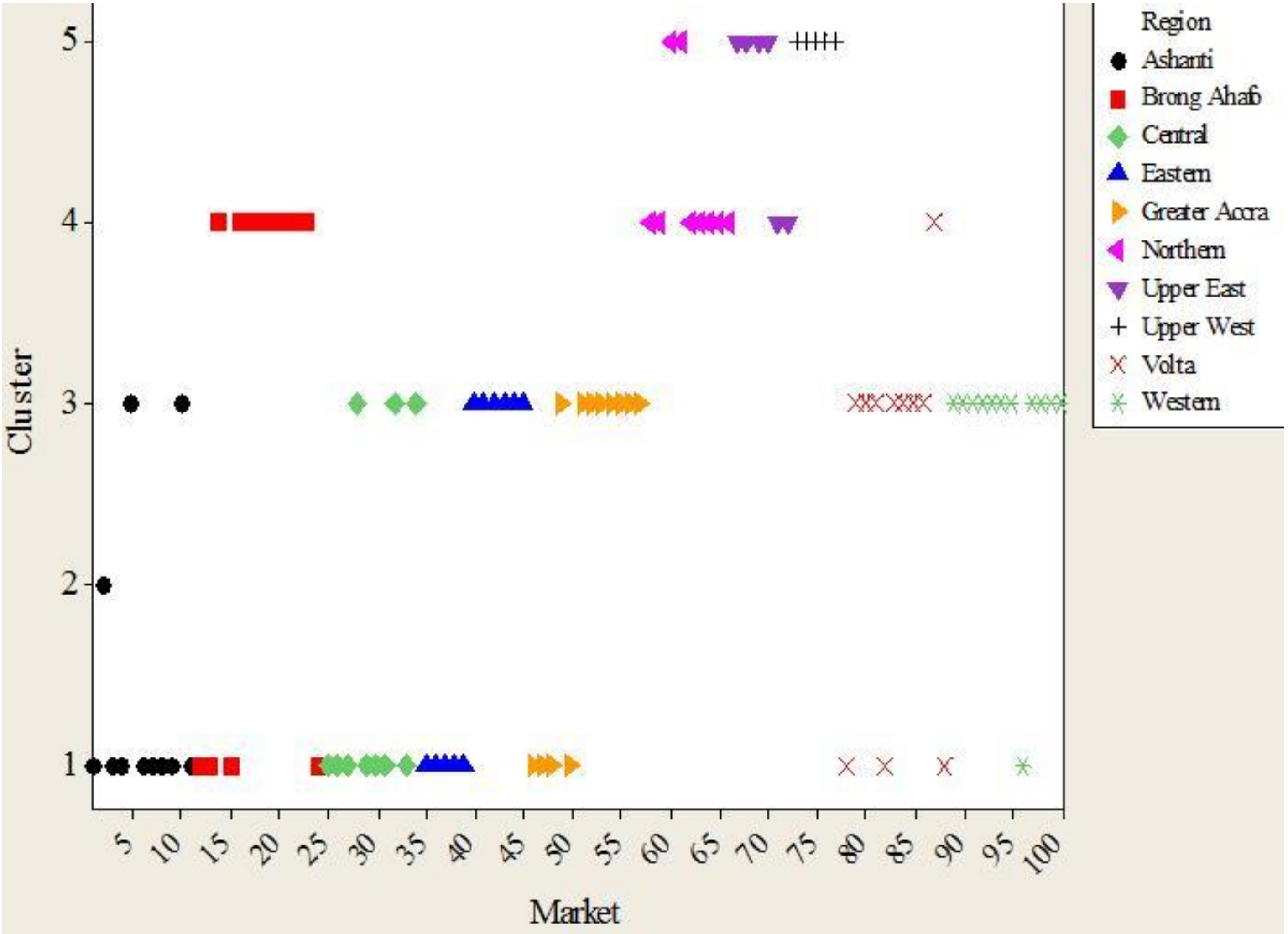


Figure 9: Market Cluster Membership by Region Using the First Principal Component

4.5.2 Market Clustering Using the Second Principal Component

The dendrogram showing all the possible clusters using the second principal component is given in Appendix I. From that diagram only three distinct clusters can be identified. A plot of the three markets cluster membership is given in Figure 10. The first cluster is made up of markets in six regions, with only one market from the Northern region. This market from the Northern region is numbered 61, and from the previous section it is identified as Gushiegu, a rural market. We will label this cluster as *Moderate-Priced market*.

The second cluster has the smallest membership and it is made up of markets that are located in only two regions, Ashanti and Brong Ahafo. From the scatter plot of the component scores in the previous section, we see that this cluster can be labelled as *Low-Priced market*. All markets from 4 to 10 are in this cluster. These include the 7th and 8th markets, which are already identified in the previous section as Agogo and Nsuta, rural markets in the Ashanti region.

The third cluster has the largest membership. The membership involves markets from all regions with exception of the Ashanti region. Another characteristic of this cluster is that it includes all markets from four region; the Greater Accra, Upper East and West and Volta regions. This cluster may be labelled as *High-Priced market*.

A noticeable feature of the cluster membership on the second component is that in all three market clusters one can find markets from the Brong Ahafo region.

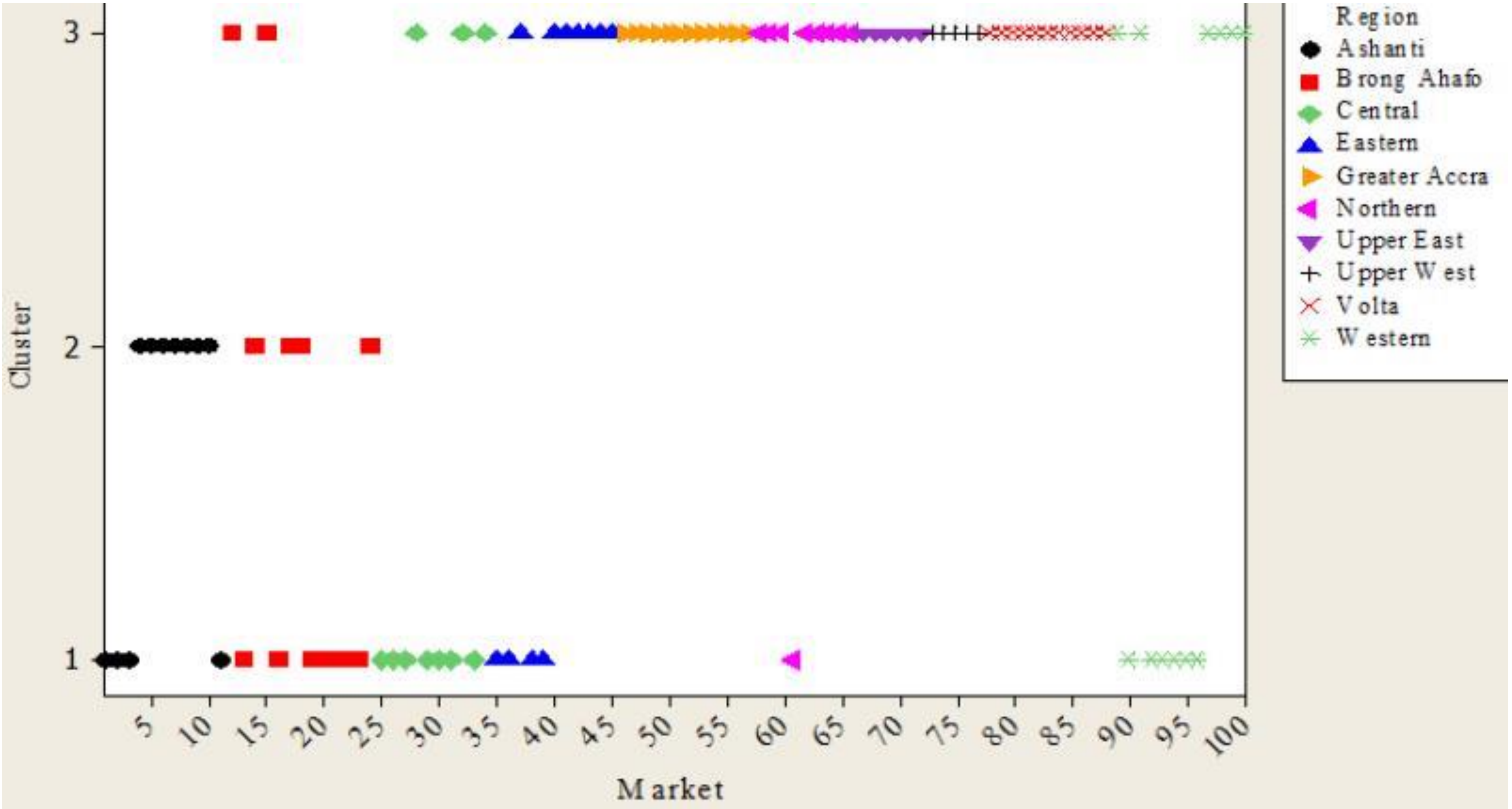


Figure 10: Market Cluster Membership by Region using the Second Principal Component

4.6 Summary of Further Analysis

A number of procedures for examining the covariance structure are carried out to determine the suitability of the data for principal component analysis. The procedures include the examination of the variance-covariance matrix, the eigen-analysis and the Kaiser-Meyer-Olkin's measure of sampling adequacy. The KMO measure of sampling adequacy is obtained to be 0.8250, which is high. This index indicates that the fifteen variables adequately summarize the degree of common variation in food prices. The examination of the variance-covariance matrix involves basically the eigen-analysis. It is observed that the first principal component alone accounts for about 40% of total variation in the data, whilst the second and third principal components accounts for about 19% and 8%, respectively, of the total variation. Thus, the first three components, which are the only components that account for more than 1% of the variation, together explain about two-thirds of the total variation in the data. The high percentage of variation explained by the first component alone shows that the first component might be a very influential single dimension along which commodity-price levels of the various markets could be evaluated. The use of the scree plot also confirmed that the number of influential components for assessing price levels could not exceed three. The results of all of these procedures informed the use of the principal component analysis of the data set.

Using a cut-off value of loading of 0.4, it is found that only the first three principal components can be extracted. However, it is only the first two components for which meaningful interpretation could be assigned. The first principal component is highly influenced by almost all the items.

Consequently, the first principal component is labelled as the *weighted sum of all the food items*. This component alone explains about 40% of the variation in the prices of the fifteen items. Thus, in evaluating the price levels of the hundred markets, the first most important dimension is the general prices of all the items.

The formation of the second principal component is highly influenced by items that are generally considered to be the main constituents of the staple food of a typical local diet. These commodities are Maize, White Yam, Cassava, Tomatoes, Dried Pepper and Groundnut Oil. Thus, the second principal component is labelled as the *local diet components*. The second component alone explains about 19% of the variation in the prices. Thus, to evaluate the price levels of the markets, another rule that can be used is the prices of only the items that are considered as constituting the main local diet of the Ghanaian folks. Therefore, the first two components explain about 59% of the variation in the commodity prices.

By means of a scatter plot of the first and second component scores, the markets have been categorized as *Lowest-priced, low-Priced, Moderate, High-Priced and Highest-Priced markets*. With reference to the first principal component, only one market is identified as *Lowest-Priced*. This market is Tepa market, a rural market in the Ashanti region. On the other hand, there are about five markets that may be regarded as *Highest-Priced markets*. Typical among them are Gushiegu, a rural market in the Northern region; Bongo, a rural market in the Upper East region; Jirapa and Busie, both rural markets in the Upper West region; and Tumu, an urban market in the Upper West region. All the markets in the Northern, Upper East and West Regions are typically

High-Priced markets, whist most markets in the Eastern, Western, Ashanti, Volta and Greater Accra Regions are *Low-Priced markets*. Markets in the Central Region are typically *Moderate-Priced*.

Using the centroid method of hierarchical clustering of the first principal component, it is observed that five distinct market clusters can be identified. These categories mostly coincide with those obtained from the scatter plot of the component scores. Notably, Tepa market, a rural market in the Ashanti Region is identified as forming one distinct cluster, which may be interpreted as the *Lowest-Priced market*.

Using the centroid method of hierarchical clustering of the second principal component, it is observed that three distinct market clusters can be identified. These clusters are interpreted as *Low-Priced, Moderate-Priced and High-Priced markets*. Typically on the second principal component, a *Low-Priced market* is Agogo and Nsuta, both rural markets in the Ashanti region; a *Moderate-Price market* is Takoradi, an urban market in the Western region and Atebubu, a rural market in the Brong Ahafo region; and *High-Priced markets* are Kaneshie and Mamprobi, both urban markets in the Greater Accra region.

CHAPTER FIVE

SUMMARY, DISCUSSION, CONCLUSION AND RECOMMENDATION

In this chapter, we take a second look at the summary of the study which will involve mainly the summaries from the preliminary as well as the further analysis. We will then provide a discussion of some relevant findings that have emerged from the analysis. On the basis of the findings, we will draw some conclusions and provide relevant recommendations.

5.1 Summary

This thesis has been concerned with the study of general prices of some selected food items from various markets across the ten regions of Ghana. The number of markets selected from the various regions is in proportion to the actual number of markets in the regions. In each region markets were selected from both the rural and urban areas. Out of the 100 selected markets, 59 were located in rural areas while 41 in urban areas. Data on prices of fifteen food items from each of the selected markets were obtained from the Statistical, Research and Information Directorate of the Ministry of Food and Agriculture that covered prices of the items in 2008.

In order to analyze this high dimensional data set, a number of multivariate techniques were used, of which some were routine. The main objective was to determine the major dimensions along which to evaluate the level of prices in the various markets in order to identify the markets and regions which are high-priced and those that are low-priced.

The examination of the variance-covariance matrix of the price data set found that there were generally high covariations among the items. This suggested that the selected items adequately sum up the variation in prices across the country. Within the individual items, Imported Rice was the item with the least variation in prices. Other items that exhibited less price variability were Orange, Banana, and Groundnut Oil. On the other hand, White Cowpea and Red Groundnut had the highest variability in prices.

Another multivariate procedure that was used examined the zero-order and partial correlation coefficients matrices as well as the differences between the two matrices. There were a number of very low and high differences in the two correlations coefficients among some items. It was observed that the differences in the two types of correlation coefficients were low for some pairs of items, notably: Maize and White Yam; White Yam and Garden Eggs; White Yam and Groundnut Oil; Cassava and Tomatoes; Banana and Tomatoes; Garden Eggs and Tomatoes; Garden Eggs and Dried Pepper. The implication was that the prices of these items were not influenced by the presence or absence of a variety of other items in the same market. However, the correlation differences were very high among some other pairs of items, notably: Orange and Cassava; Garden Eggs and Smoked Herrings; Palm Oil and Smoked Herrings; and Garden Eggs and Banana. This suggested that in a market with a variety of other food items, the prices of such items are likely to reduce. Thus, the simultaneous purchases of a combination of these items are likely to reduce as the consumer may switch to other items. Generally, there were low partial correlations among the items which buttress the fact that the selected items adequately represent the variability in the prices of food items across the country.

Further exploratory analysis showed that there were markets in the various regions in which prices of food items were in the extremes. It came to light that the Brong Ahafo Region has the leading number of six markets that recorded extreme prices in six different items. Out of these six, five recorded lowest prices with only one recording the highest price in an item (Smoked Herring). It thus appears that the Brong Ahafo is the region with generally low food prices. In contrast, the Ashanti region has five markets recording extreme prices, mostly highest prices, with only one of the five markets (Nsuta) recording lowest price in an item (Dried Pepper). All regions were observed to have markets that recorded extreme prices in one commodity or another with exception of the Western region that had no market that recorded extreme prices. The Upper East and Upper West regions had markets (six in all) that record only highest prices but not lowest prices. Thus, it appears that these two regions have markets that are consistently the highest-priced. The Northern region is the only region that has a single market that record extreme prices (lowest prices) in more than one item (White cowpea and Red Groundnut). It therefore appears that markets in the Western region have the least varied prices, whilst markets in the Ashanti and Brong Ahafo regions have the most varied prices.

A number of more rigorous procedures were used to examine the covariance structure of the data set in order to determine the suitability of analysing the data by means of principal component analysis, which is dimensionality-reduction technique. The procedures include the study of the the Kaiser-Meyer-Olkin's measure of sampling adequacy (KMO) and the eigen-analysis of the variance-covariance matrix. The study of the KMO involved the examination of both the zero-order correlations as well as the partial correlations. It was found that the partial correlations were generally low which resulted in a high KMO value of 0.8250. This index indicated that the fifteen variables adequately summarize the degree of common variation in food prices. The eigen-analysis revealed that only the first three components accounted for more than 1% of the variation in the data. The first principal component alone accounted for about 40% of total variation in the data, whilst the second and third principal components accounts for about 19% and 8%, respectively, of the total variation. Thus, together the three components explained about two-thirds of the total variation in the data. The high percentage of variation explained by the first component alone showed that the first component might be a very influential single dimension along which levels of prices of commodities in the various markets could be evaluated. The use of the scree plot also confirmed that the number of influential components for assessing price levels could not exceed three. The results of all of these procedures served as a motivation for the use of the principal component analysis of the data set.

Using a cut-off value of loading of 0.4, it is found that only the first three principal components can be extracted. However, it is only the first two components for which meaningful interpretation could be assigned. The first principal component is highly influenced by almost all the items. Consequently, the first principal component is labelled as the *weighted sum of all the food items*. This component alone explains about 40% of the variation in the prices of the fifteen items. Thus, in evaluating the price levels of the hundred markets, the first most important dimension is the general prices of all the items.

The formation of the second principal component is highly influenced by items that are generally considered to be the main constituents of the staple food of a typical local diet. These commodities are Maize, White Yam, Cassava, Tomatoes, Dried Pepper and Groundnut Oil. Thus, the second principal component is labelled as the *local diet components*. The second component alone explains about 19% of the variation in the prices. Thus, to evaluate the price levels of the markets, another rule that can be used is the prices of only the items that are considered as constituting the main local diet of the Ghanaian folks. Therefore, the first two components explain about 59% of the variation in the commodity prices.

By means of a scatter plot of the first and second component scores, the markets have been categorized as Lowest-priced, low-priced, moderate, high-priced and highest-priced markets. With reference to the first principal component, only one market is identified as lowest-priced. This market is Tepa market, a rural market in the Ashanti region. On the other hand, there are about five markets that may be regarded as highest-priced markets. Typical among them are Gushiegu, a rural market in the Northern region; Bongo, a rural market in the Upper East region; Jirapa and Busie, both rural markets in the Upper West region; and Tumu, an urban market in the Upper West region. All the markets in the Northern, Upper East and West regions are typically high-priced markets, whist most markets in the Eastern, Western, Ashanti, Volta and Greater Accra are low-priced markets. Markets in the Central region are typically moderate.

Using the centroid method of hierarchical clustering of the first principal component, it is observed that five distinct market clusters can be identified. These categories mostly coincide with those obtained from the scatter plot of the component scores. Notably, Tepa market, a rural market in the Ashanti region is identified as forming one distinct cluster, which may be interpreted as the lowest-priced market.

Using the centroid method of hierarchical clustering of the second principal component, it was observed that three distinct market clusters could be identified. These clusters are interpreted as low-priced, moderate-priced and high-priced markets. Typically on the second principal component, a low-priced market is Agogo and Nsuta, both rural markets in the Ashanti region; a moderate-price market is Takoradi, an urban market in the Western region and Atebubu, a rural market in the Brong Ahafo region; and high-priced markets are Kaneshie and Mamprobi, both urban markets in the Greater Accra region.

5.2 Discussion

In this section we provide discussion on the some of the findings made in this study. The discussion basically will bother on the variance-covariances among the food items, the correlation analysis and the choice of the techniques used in the study.

On the Variation in the Prices of Individual Items

The variability in the prices of particular commodities is one that is worth discussing. It is found that the prices of White Cowpea are the most dispersed across the country. White Yam is also one of the items with large variability in prices. The third item of interest is Groundnut Oil, even though the variability in prices is among the lowest. In each of these three items the coefficient of skewness is found to be negative unlike the remaining twelve items. This means that in most of the markets the prices of these items are higher than the average. It is therefore necessary for the consumer to be very selective of the market to purchase these items. This is because it is more likely that one may be found in a market where the price of these items is on the higher side.

The two items with the highest variation in prices are White Cowpea and Red Groundnut. Interestingly, these are the two items that have their lowest prices in a single market in Gushiegu in the Northern region. This is the only market among all the 100 hundred markets that record extreme (lowest) prices in more than one item. It does appear that this is the only market (or region) in which the two items are predominantly grown but much less grown in other regions. Consequently, prices keep changing drastically as the few quantities get distributed across the country. This is also true for the other commodities that have high variation in prices (like Kobi which is predominantly farmed in the coastal areas). This shows that the price of a commodity is likely to have less variation if it can be traded in various parts of the country.

Closely related to the above point is the relevance of the difference between the correlation coefficient and the partial correlation coefficient of a given item. It has been revealed in this study that the differences in correlations are close to zero for mainly the following pairs of commodities: Maize and White Yam; White Yam and Garden Eggs; White Yam and Groundnut Oil; Cassava and Tomatoes; Banana and Tomatoes; Garden Eggs and Tomatoes; and Garden Eggs and Dried Pepper. It was explained that after isolating the effect of the prices of all other commodities, there is not much change in the correlation coefficient between the particular pair of items, Tomatoes and Cassava, for example. Thus, the presence or absence of other commodities on the market will not affect the simultaneous purchases of the two commodities. It is interesting to note that the items that mostly influence the formation of the second principal component (which is labelled as *local diet component*) are Maize, White Yam, Cassava, Tomatoes, Dried Pepper and Groundnut Oil. Incidentally, these items coincide with those items that have been identified as being beneficial for pairwise simultaneous purchases. Thus, it now appears that the prices of commodities that we classify as forming our local diet are less likely to change whether in the presence or absence of a variety of other commodities. This finding is in line with the observation by Roache (May 2010) in his study on 'What Explains the Rise in Food Price Volatility?' He found that Ghanaian households appear to be relatively insulated from swings in international food markets, because a large share of their diet is based on non-tradable staples such as cassava and sorghum. It has been found in this work that there were generally no extreme prices on items that we consider to constitute staple food.

The examination of the differences in the zero-order and partial correlation coefficient matrix further revealed that the correlation coefficient between some pairs of commodities change drastically after controlling the effect of other commodities. These pairs of commodities are: Orange and Cassava; Garden Eggs and Smoked Herrings; Palm Oil and Smoked Herrings; and Garden Eggs and Banana. The implication is that in the presence of other commodities, the relative prices of these items are highly likely to reduce. Thus, it may not be economical to purchase some of these items when there are no varieties of food items in the same market. This pattern is largely in support of the findings by Claro and Monteiro (2010). We recall that in the literature, these researchers found that the participation of fruit and vegetable in total food purchases increased as the price of these foods decreased, or as income increased.

On the suitability of Selected Commodities for Assessing Price Levels of the Markets

The low value of the sum of squares of the partial correlations (3.4778) as compared to the sum of squares of the ordinary correlation coefficients (16.3955) has a number of implications. One of these is that the prices of the selected commodities vary in a market of varying food items. In this study we see that, generally, the prices are more likely to reduce in the presence of a variety of other food items. This shows that the items selected for periodic market evaluation by the Statistics, Research and Information Directorate of the MoFA are a good set of items that generally gives a good reflection of the price variability across the entire country.

On the Use of Hierarchical Clustering Procedure

This study made use of the Centroid method of the hierarchical cluster analysis. The result of this method was compared with those of other hierarchical methods such as the Average Linkage and the Ward's methods. It was realised that the clustering result obtained from the Centroid method generally coincides with all the results of the other methods with exception of the Ward's method. The difference is that in the case of the Ward's method, it did not produce distinct clusters that contained a single or very few markets like the Tepa cluster. However, all the other methods produced the same single

Tepa market cluster and other small cluster-membership sizes. The Centroid method was therefore selected as one of the typical cluster methods that produced a cluster solution that will be consistent with the subject matter and the objective of the study.

5.3 Conclusion and Recommendation

This study considered a multivariate analysis of the prices of fifteen food items in hundred selected markets from all the regions of Ghana. The data covered prices in 2008 and was obtained from the Statistical, Research and Information Directorate of the Ministry of Food and Agriculture. The objective was to identify the various dimensions, if any, along which to evaluate the level of prices of the food items in order to identify specific markets in which items are high-priced and those that are low-priced. At the end of the analysis, two major components that explain variations in food item prices were found.

The first two components that can be used to determine the price levels of markets are, in order of importance: *weighted sum of all the food items*; and the *local diet components*. Thus, in evaluating the price levels of the hundred markets, the relevant dimensions along which this can be done is to consider the general level of prices of all the items as well as the prices of those commodities that are considered to be the main constituents of a typical local diet. These commodities are Maize, White Yam, Cassava, Tomatoes, Dried Pepper and Groundnut Oil. The first two components explain about 60% of the variations in the commodity prices in all markets across the country.

Cluster formation of the first principal component scores revealed five distinct market clusters each of which was labelled if the markets membership were Lowest-priced, low-priced, moderate, high-priced or highest-priced. Notable among the five clusters was one that had only one membership, which was *Tepa*, a rural market in the Ashanti region. This was identified as the lowest-priced market in the country with respect to prices of general food items. On the other hand, there are about five markets that may be regarded as highest-priced markets. These are predominantly located in the Northern, Upper East and Upper West regions. Typical among them are Gushiegu, a rural market in the Northern region; Bongo, a rural market in the Upper East region; Jirapa and Busie, both rural markets in the Upper West region; and Tumu, an urban market in the Upper West region. Overall, all the markets in the Northern, Upper East and West regions are typically high-priced markets, whilst most markets in the Eastern, Western, Ashanti, Volta and Greater Accra are low-priced markets. Markets in the Central region are typically moderate.

Cluster formation of the second principal component scores revealed only three distinct market clusters each of which was labelled as being Low-priced, moderate, or high-priced. Typically on the second principal component, a low-priced market is *Agogo* and *Nsuta*, both rural markets in the Ashanti region; a moderate-price market is *Takoradi*, an urban market in the Western region and *Atebubu*, a rural market in the Brong Ahafo region. Kaneshie and Mamprobi, both of which are urban markets in the Greater Accra region were found to be high-priced markets with respect to food items that are regarded as constituting the local diet.

The structure of the variance-covariance matrix of the price data suggested that generally, the food items selected by the Statistical, Research and Information Directorate of the Ministry of Food and Agriculture adequately sum up the variation in food item prices. It is therefore appropriate to continue the use of the prices of these items for monitoring price levels across the country. The study has also shown that it will be economically beneficial to make simultaneous purchases of certain groups of items only when there are varieties of other food items in the market. These items are Orange, Cassava, Palm Oil and Smoked Herrings. It is thus expedient to make purchases of these items from a large market. However, for such items as Maize, White Yam, Groundnut Oil, and Tomatoes, there may not be any economic benefit to make simultaneous purchases of them when there are a variety of other items. It may be beneficial to purchase these items from a small market.

The findings of this study show clearly that it is necessary for the consumer to be mindful of the location and the size of the market from which one intends to make purchases of certain categories of items.

REFERENCES

African Development Fund (2006), Burkina Faso - *Decentralized Rural Development Support Project in the Provinces of Gnagna and Kourittenga.* Appraisal Report. African Development Bank, Tunis.

Barnett, E. (2008). "*HighFood Prices: Challenges and solutions*" www.worldchanging.com/archives/007762.html

Bartlett, M. S. (1954). *A note on the multiplying factors for various chi-squared approximations.* Journal of the Royal Statistical Society, **16** (Series B), *296-298.*

Bailey, T. C. & Gatrell, A. C. (1995). *Interactive Spatial Data Analysis*. London: Longman Scientific and Technical.

Begg, D., Fischer, S., Dornbusch, R. (2003); *Economics* (7 ed). McGraw Hill Education Publication. Berkshire. ISBN 13-978-007710775-8

Berven, N. L. (1985). "Reliability and validity of standardized case management simulations". *Journal of Counseling Psychology*, **32**, *397-409.*

Cattell, R.B (1966). *The Scree Test for the number of factors.* Multivariate Behavioral Research, ***1***, *245-276.*University of Illinois, Urbana – Champaig. IL)

Claro, R. M., Monteiro C. A. (2010). *Family income, food prices, and household purchases of fruits and vegetables in Brazil*. Rev Saude Publica. **44** (*6*):1014-20.

Elliott, R. (1985). "Helpful and non-helpful Events in Brief Counseling Interviews: An empirical taxonomy". *Journal of Counseling Psychology*, **32**, *307-322*.

FASDEP II Document, (2009). http://mofa.gov.gh/site/?page_id=598. ISBN: 798-9988-1-2704. Published in collaboration with Agriculture Sector Development Partners Working Group

Ferris, R. S. B., Legg, J., Bua, A., Agona, A. & Whyte, J. (2000). *Dissemination and Utilisation of Mosaic Resistant Cassava in Uganda,* 8th Quarterly USAID Report. **159**.

Filmer, D. & Lant, H. P. (2001). *Estimating Wealth Effect Without Expenditure Data – or Tears: An application to Educational Enrollments in States of India.* Demography **38,** *115-132*.

Gordon, A.D. (1999). *How Many Clusters? An Investigation of Five Procedures for Detecting Nested Cluster Structure Classification.* New York: Chapman and Hall/CRC.

Gruber, P. and Gruber, T. and Fabian J. '*Grassmann Clustering*' 14th European Signal Processing Conference (EUSIPCO 2006). Institute of Biophysics, University of Regensburg.

Green, P.E (1976). *Analyzing Multivariate Data*, Dryden Hinsdale, Illinoiis.

Guttman, L. 1954. "The Principal Components of Scalable Attitudes." *216-257* in Paul F. Lazarsfeld (ed.) *Mathematical Thinking in the Social Sciences*. New York: Dryden Press.

Hair, J.F., Anderson, R.E. and Tatham, R.L (1987). *Multivariate Data Analysis*, Macmillan, New York.

Johnson, R.A. and Wichern, D.W. (2003). *Applied Multivariate Statistical Analysis*. Prentice-Hall of India Private Limited.

Johnson, J.H (1981). *The Q-Analysis of Road Traffic Systems*. Environment and Planning B **8(2)**, *141-189.*

Kaiser, H. (1970). *A Second Generation Little Jiffy*. Psychometrika, **35**, *401-415*.

Kaiser, H. (1974). *An Index of Factorial Simplicity*. Psychometrica, **39**, *31-36.*

Larose, D.T (2006): '*Data Mining: Methods and Models'*, John Willey & Sons Inc. Publications. ISBN – 13978-00471-66656-1.

Levine, N. (1999a). CrimeStat: *A Spatial Statistics Program for the Analysis of Crime Incident Locations*, Version 1.1. Washington DC: Ned Levine & Associates/National Institute of Justice.

Nunnally, J.C. (1978). *Psychometric Theory*. 2 ed. New York: McGraw Hill

Pallant, J. (2001). *SPSS Survival Manual.* Open University Press, Buckingham.

Roache, S. K. (2009). *What Explains the Rise in Food Price Volatility*? IMF Working Paper, WP/10/129 – Authorized for distribution by Thomas Helbling Research Department.

Robbins, P. and Ferris, R.S.B. (1999). *A preliminary Study of the Maize Marketing System in Uganda and the Design of a Market Information System.* CTA/IITA. Contract No. 4-1-06-215-9. *23*.

Robbins, P., Ferris, R.S.B. and Muganga, A. K. (2000). *Market information Services in Post-liberalized Uganda and Eastern Africa*. Phaction News.

Robbins, P. and Ferris, R.S.B. (2000). *Co-ordination of a Preliminary Study of the Maize Marketing System in Uganda and the Design of a Market Data Information System.* Contract No. 4-1-06-215-9. Design of a Pilot Scheme for Testing of a Market Information System. *16*.

Sharma, S. (1996). *Applied Multivariate Techniques.* John Wiley & Sons, New York.

Shepherd, A. W. (1997). *Market Information Services: Theory and Practices*, FAO.

Shepherd, A. W. (2001). *Farm Radio as a Medium for Market Information Dissemination*, FAO.

SRID (2000). *Market Enumerator's Manual*. http://mofa.gov.gh. Visited: 9th May 2010.

Stevens, J. (1996). *Applied Multivariate Statistics for the Social Sciences* (3 Ed). Mahwah, New Jersey; Lawrence Erlbaum.

Tabachnick, R.G & Fidell, L.S (1996). *Using Multivariate Statistics* (3 Ed), New York; Harper Collins.

The Newsletter of the Global Post-Harvest Forum. Number 2, June 2000.

United Nations, *Food and Agriculture Organization Report* (2008).

Warde, A. (1997). *Consumption, Food and Taste.* London, Sage Publications.

www.ifad.org/operations/food/. Higher and Volatile Food Prices and Poor Rural People.

IFAD. Rural Poverty Report (2011). www.ruralpovertyportal.org. Policy Report by G8 in June 2011. *Price Volatility in Food and Agricultural Markets: Policy Responses.* A Report Co-ordinated by FAO and OECD

www.support.sas.com/publishing/pubcat/chaps/55129.pdf. *Introduction: The Basics of Principal Component Analysis.* Chp 1. Visited on 10th May 2011.

www.younglives.org.uk (2008). "*Children and the Food Price Crisis*" Young lives Policy, Brief 5; September 2008 Edition.

APPENDICES

APPENDIX A

The Data for the Study

ID	MARKETS	2008 PRICE OF COMMODITIES (GHS) – DATA FOR THE STUDY														
		Mz	RiImp	YmWt	Cv	Org	Ban	Tm	GEg	PpDr	GnR	CpWt	GnOil	PmOil	HerSm	Kobi
1	RA_Ejura	40.88	55.11	92.70	7.17	2.12	3.33	27.59	8.42	59.78	116.05	110.21	39.00	19.54	35.08	34.90
2	RA_Tepa	49.25	53.18	61.46	10.83	3.52	1.00	43.58	10.33	56.07	184.60	166.90	18.88	17.07	32.93	26.75
3	RA_Agona	48.33	54.25	65.00	5.00	5.33	1.72	51.17	28.67	40.00	105.00	107.50	50.40	23.08	32.00	31.00
4	RA_Abofo	42.21	54.26	75.56	10.24	4.01	1.48	25.11	11.04	35.71	128.11	110.30	40.36	24.30	25.50	22.50
5	RA_Adugy	46.21	53.90	91.46	9.05	3.36	2.28	31.40	14.97	18.10	138.33	143.40	44.60	25.20	21.15	18.14
6	RA_Juab	55.52	54.75	86.98	10.79	4.01	2.18	28.08	17.35	33.80	112.89	106.82	21.60	18.61	26.7	18.28
7	RA_Agogo	57.82	56.53	74.58	5.82	4.03	3.11	29.06	24.03	19.58	90.00	85.00	24.00	17.43	20.93	19.20
8	RA_Nsuta	46.75	54.97	54.17	4.42	3.65	1.98	50.12	13.56	18.00	120.56	116.93	38.24	25.42	22.46	21.33
9	UA_Ksi	49.11	54.52	86.68	9.62	2.50	2.30	28.94	15.58	50.64	116.57	91.97	30.52	19.44	23.18	23.19
10	UA_Bekw	63.26	55.63	89.38	6.89	3.27	2.11	46.96	15.68	31.00	125.28	109.58	34.17	18.47	25.14	18.2
11	UA_Obua	68.96	54.40	93.50	8.25	3.97	1.20	44.48	12.71	33.02	97.67	111.33	37.23	17.01	22.03	45.00
12	RB_Ateb	42.98	53.53	93.31	5.27	6.79	3.13	36.72	12.80	76.49	109.20	99.07	33.50	15.84	46.92	44.33
13	RB_KDns	44.00	56.25	66.00	5.76	6.03	2.19	31.67	12.30	83.34	90.00	108.00	30.00	19.25	28.00	28.33
14	RB_Yeji	46.00	54.53	78.75	6.00	6.00	2.00	18.67	14.67	47.50	80.00	73.75	14.00	19.00	80.00	25.00
15	RB_Kukuo	45.94	55.5	80.75	15.50	2.58	2.08	28.50	10.25	99.00	110.00	127.00	29.50	12.00	28.25	54.00
16	RB_Nsawk	40.00	56.00	46.00	14.17	2.11	2.20	38.50	17.07	57.00	103.31	79.00	31.00	18.77	21.33	49.03
17	RB_DNkw	37.80	55.67	62.60	11.50	3.04	1.55	32.50	11.00	35.20	85.55	97.00	29.00	18.94	23.00	55.00
18	RB_Drobo	43.72	56.01	50.00	4.00	2.00	1.18	50.00	12.33	38.32	93.76	82.00	27.11	27.00	26.50	43.56
19	UB_Suny	32.84	52.82	76.20	16.25	3.90	0.77	32.43	19.88	28.90	99.55	99.10	40.13	30.38	21.81	61.00
20	UB_Techi	34.82	53.88	65.59	11.28	4.26	1.35	33.10	9.42	54.60	92.07	83.03	36.00	18.00	20.00	60.00
21	UB_Brkm	38.40	53.60	66.60	4.80	4.80	0.95	35.20	17.00	42.60	90.25	84.50	33.00	28.00	31.67	58.00
22	UB_DAhnk	41.00	55.00	92.00	3.60	3.67	1.90	25.20	8.00	59.00	85.20	80.00	35.44	30.00	21.67	34.25
23	UB_Goas	49.13	55.46	62.50	7.00	1.45	0.83	21.88	11.63	65.00	83.93	85.00	37.30	23.25	28.50	48.75
24	UB_Kinta	30.73	52.50	73.75	9.25	5.55	2.42	29.46	6.50	36.25	91.25	107.20	42.00	17.25	33.75	38.75

ID	MARKETS	2008 PRICE OF COMMODITIES (GHS) – DATA FOR THE STUDY														
		Mz	RiImp	YmWt	Cv	Org	Ban	Tm	GEg	PpDr	GnR	CpWt	GnOil	PmOil	HerSm	Kobi
25	RC_Manks	50.40	53.85	91.75	14.36	3.62	1.95	62.50	12.30	55.83	108.33	100.81	43.90	21.00	32.50	22.40
26	RC_Bawj	46.33	54.00	60.00	14.58	3.17	1.53	59.33	11.50	61.17	126.17	107.67	39.04	21.83	37.33	34.00
27	RC_Kasoa	41.83	53.17	66.67	16.33	2.75	4.65	53.10	16.83	65.33	118.00	109.83	35.49	22.00	30.65	33.50
28	RC_Elmna	55.60	54.00	99.17	16.62	3.38	3.13	60.33	13.00	65.33	126.17	125.00	41.09	23.00	28.33	34.00
29	RC_APrso	47.20	53.25	81.67	14.67	1.80	1.02	64.58	8.15	32.40	89.45	131.7	33.22	15.25	31.00	25.80
30	RC_Ajmk	42.50	54.00	88.00	15.62	2.25	1.50	63.32	15.58	33.07	117.83	109.89	40.11	15.00	32.17	33.41
31	RC_FNks	50.21	51.67	85.00	20.50	2.82	1.92	52.84	10.55	50.00	89.81	81.88	41.32	22.00	25.67	18.00
32	UC_Cape	49.57	54.15	90.83	13.85	3.42	3.50	60.31	11.00	63.33	125.83	118.50	45.16	24.83	29.33	28.50
33	UC_Swed	46.40	53.14	66.67	12.25	3.30	1.08	53.00	14.51	52.17	95.50	98.17	44.50	21.33	34.40	36.67
34	UC_Dunkw	54.30	54.00	99.32	11.40	2.83	0.92	51.08	13.10	88.53	117.50	119.17	40.98	17.33	41.04	53.50
35	RE_Soma	54.14	53.09	87.14	10.06	2.74	1.83	55.83	10.31	53.44	98.72	103.00	44.81	19.86	21.32	38.60
36	RE_Akoas	48.57	55.14	84.57	9.57	2.56	1.23	52.57	8.47	58.67	88.50	92.57	46.11	20.63	22.50	31.64
37	RE_Ahom	48.60	53.1	87.77	8.00	1.40	0.80	43.00	7.00	81.25	93.25	102.60	40.35	21.40	24.75	35.50
38	RE_Anynm	49.00	53.22	89.43	7.36	2.25	0.92	50.50	9.64	59.67	107.14	107.86	41.23	22.71	22.57	33.57
39	RE_Agorm	48.00	53.20	79.00	10.80	2.30	1.82	50.40	6.55	62.50	106.22	100.93	40.33	21.40	24.00	30.00
40	UE_Kdua	53.31	54.10	86.47	17.16	3.48	2.75	52.33	10.83	93.57	117.33	130.83	42.88	24.83	37.50	33.17
41	UE_Suhum	50.67	53.77	85.00	13.17	2.82	1.83	46.68	12.50	105.50	125.00	129.17	43.23	25.00	34.33	32.16
42	UE_NTafo	55.20	54.06	96.55	10.60	3.10	1.76	70.20	11.00	92.20	130.00	129.20	42.29	23.60	39.40	42.20
43	UE_Asnmk	46.60	54.18	86.25	10.60	3.07	1.73	53.80	13.50	103.61	118.43	125.82	42.77	24.00	23.00	40.81
44	UE_AOda	56.00	53.04	96.62	55.60	2.98	2.06	61.40	12.90	85.00	127.20	133.00	41.09	19.82	39.20	42.80
45	UE_Mprso	49.83	53.83	76.67	45.33	2.92	2.08	65.50	11.67	64.18	115.00	124.14	44.00	27.50	23.67	33.60
46	RG_Kaseh	58.23	54.23	78.66	45.22	4.81	6.18	53.50	29.88	68.17	119.00	97.00	45.31	30.83	30.83	42.35
47	RG_AdFoa	58.40	53.80	78.60	45.29	4.83	6.16	53.22	29.88	68.14	118.80	96.61	45.44	30.50	30.83	42.30
48	RG_Hobor	58.90	54.01	80.32	31.19	3.50	4.53	49.80	18.38	59.77	106.67	92.00	39.99	23.75	38.83	44.52
49	UG_Mkl	60.31	52.83	83.11	19.82	4.50	4.00	78.16	18.27	74.19	130.80	137.21	43.15	24.00	37.5	59.33
50	UG_Agbls	58.11	52.73	94.03	10.67	4.30	3.60	67.17	16.16	77.50	105.15	130.83	42.96	22.55	36.80	68.23
51	UG_Knsh	60.31	53.16	91.12	21.83	4.67	4.50	85.00	20.87	88.20	134.50	139.32	43.88	25.67	38.83	50.00

ID	MARKETS	2008 PRICE OF COMMODITIES (GHS) – DATA FOR THE STUDY														
		Mz	RiImp	YmWt	Cv	Org	Ban	Tm	GEg	PpDr	GnR	CpWt	GnOil	PmOil	HerSm	Kobi
52	UG_Mprb	59.80	52.80	99.62	22.43	3.98	3.40	86.30	2.77	89.70	127.93	137.88	43.78	26.82	35.82	57.40
53	UG_Tema	61.00	53.22	96.00	24.17	4.82	4.53	88.67	20.67	78.41	122.23	117.38	45.50	27.83	38.03	51.34
54	UG_Ashm	61.16	53.30	96.02	27.50	5.00	3.22	88.61	21.54	67.95	119.50	116.11	45.50	27.80	38.61	52.50
55	UG_Madn	68.68	53.67	96.21	20.18	5.50	3.41	86.50	17.50	68.44	116.95	134.67	45.00	36.83	31.50	57.32
56	UG-Malm	63.50	54.05	95.50	21.67	5.00	4.00	74.67	20.80	70.11	128.50	106.66	41.58	26.86	25.39	57.40
57	UG_Dome	63.31	54.88	98.09	19.30	4.16	2.20	73.66	22.46	71.20	131.66	104.11	41.53	29.13	25.05	59.16
58	RN_Bimbi	44.25	53.56	51.00	23.70	9.00	6.40	35.75	22.94	62.00	94.00	96.00	43.54	34.20	59.50	57.17
59	RN_Dmgo	44.00	55.33	68.00	28.42	8.00	6.51	46.00	30.81	59.20	89.21	108.80	43.20	35.71	59.71	58.16
60	RN_Nalgu	38.50	55.58	87.20	22.78	7.92	6.24	49.00	21.3	51.18	67.90	63.00	46.30	37.15	59.61	60.10
61	RN_Gushg	34.40	55.25	54.40	23.10	7.84	6.80	38.80	21.93	46.40	62.92	58.60	27.56	24.60	58.97	57.81
62	RN_Salga	40.00	56.00	53.30	25.81	6.20	7.35	38.32	23.00	66.76	92.00	89.33	42.40	28.00	59.10	69.10
63	RN_Kmbgu	49.00	55.87	57.00	24.95	5.00	7.64	42.50	22.00	56.93	88.50	82.00	37.25	29.00	59.75	67.11
64	RN_Zabzg	49.80	55.59	59.00	25.80	4.50	6.85	42.50	20.42	49.50	80.00	87.00	44.10	37.46	58.97	59.35
65	RN_Sboba	47.97	54.99	69.75	25.70	4.50	6.03	52.40	20.91	55.50	88.00	86.40	36.65	35.00	58.07	58.70
66	UN_Tmle	48.88	55.02	69.30	25.43	4.04	6.61	50.17	21.55	43.47	80.80	93.46	41.67	31.50	58.00	68.00
67	RUE_Fumb	47.49	59.31	47.90	40.00	7.81	7.00	48.00	46.25	45.55	94.41	72.76	44.29	34.28	58.32	67.11
68	RUE_Bngo	45.72	56.98	48.00	40.39	7.80	7.13	43.22	46.37	45.61	69.22	69.42	36.61	46.13	58.73	67.81
69	RUE_Garu	49.69	56.03	66.00	43.00	8.00	7.00	42.00	43.08	71.00	69.44	69.43	46.00	46.22	58.66	66.90
70	RUE_Zebl	47.75	53.89	61.17	41.33	7.00	7.05	41.99	43.77	68.92	70.11	69.88	46.83	46.75	58.89	66.87
71	UUE_Bolg	46.07	58.17	72.00	41.44	7.01	7.22	53.50	42.00	65.75	94.59	91.50	37.03	47.98	58.23	66.30
72	UUE_Navg	52.00	58.76	71.94	45.00	8.33	7.09	48.31	41.98	67.89	93.22	93.66	41.04	45.88	58.34	66.12
73	RUW_Babl	38.64	57.87	82.22	30.90	6.83	7.00	34.69	14.11	58.60	77.53	77.64	41.81	42.80	65.00	80.00
74	RUW_Jirp	37.39	57.50	88.90	31.00	6.00	7.80	36.33	14.12	56.86	66.06	68.93	42.93	43.18	65.00	80.11
75	RUW_Bussie	43.20	56.88	86.43	33.70	7.25	7.50	36.00	18.27	53.00	65.88	64.67	41.33	44.22	64.26	80.20
76	UUE_Wa	42.00	56.06	91.72	33.00	7.33	7.52	33.07	17.17	52.71	76.13	81.43	40.22	46.61	66.01	90.00
77	UUW_Tumu	44.23	56.00	99.81	33.14	7.00	7.12	37.14	18.00	52.66	65.86	67.40	41.87	46.88	66.34	89.56
78	RV_MKse	40.10	55.70	97.94	26.12	3.75	2.63	48.07	10.33	49.36	103.77	117.38	28.67	43.80	25.55	53.00

ID	MARKETS	2008 PRICE OF COMMODITIES (GHS) – DATA FOR THE STUDY														
		Mz	RiImp	YmWt	Cv	Org	Ban	Tm	GEg	PpDr	GnR	CpWt	GnOil	PmOil	HerSm	Kobi
79	RV_Abot	52.18	55.00	81.00	13.50	3.66	1.83	65.00	10.34	69.92	112.20	140.09	31.00	33.68	24.00	42.60
80	RV_Kpve	54.20	55.76	94.40	13.45	3.00	1.52	58.00	9.53	74.11	109.70	133.71	33.45	35.14	21.88	43.67
81	RV_LAlkp	53.60	55.82	83.25	13.62	2.05	1.81	60.00	13.00	83.85	113.80	133.41	33.66	36.03	20.00	42.99
82	RV_Kptoe	43.89	54.90	86.55	12.00	3.43	2.03	57.13	11.25	80.33	99.25	72.78	34.54	39.12	26.70	42.65
83	RV_Kute	50.70	56.23	85.00	8.93	1.93	1.06	54.00	13.10	55.32	115.62	138.38	35.08	32.85	34.16	40.53
84	UV_Ho	57.16	53.44	59.00	18.16	4.08	2.20	55.50	15.00	69.50	116.17	141.17	32.77	34.18	23.00	40.99
85	UV_Kpndo	52.33	53.49	94.60	22.29	5.08	6.72	46.09	8.80	77.21	115.92	131.51	34.55	40.46	23.09	45.00
86	UV_Hohoe	57.58	53.00	81.60	11.56	2.17	1.24	57.61	9.42	77.30	94.44	143.23	34.81	39.74	19.75	45.00
87	UV_Dmbai	52.50	57.32	46.00	28.36	6.39	4.88	65.77	14.32	60.50	67.42	98.50	41.23	40.80	67.21	40.71
88	UV_Denu	62.40	53.46	98.55	11.11	5.11	5.66	46.47	9.97	46.50	110.07	139.75	32.40	36.86	46.63	55.71
89	RW_Bgoas	60.00	53.22	70.00	13.00	5.67	2.33	70.00	16.21	60.00	160.00	112.67	40.11	26.50	39.28	30.00
90	RW_Aswns	66.31	53.40	93.00	10.00	4.00	1.02	65.00	6.00	40.00	160.00	114.69	43.80	14.00	34.00	30.00
91	RW_Bibni	56.65	56.13	95.43	12.62	3.50	1.30	60.00	20.00	67.63	154.22	114.56	45.00	16.83	39.36	33.88
92	RW_SDwn	61.01	54.66	76.00	12.67	4.83	1.13	63.33	15.12	50.00	136.56	115.31	43.26	18.50	39.33	19.47
93	RW_Tikbo	67.11	55.43	90.97	15.44	5.20	2.30	62.00	17.00	56.71	131.22	114.95	44.35	18.77	37.05	19.00
94	RW_Juabs	62.66	54.88	87.16	13.33	4.22	1.81	61.12	12.00	46.94	116.00	109.87	43.30	16.00	25.00	35.00
95	RW_Elubo	66.39	53.49	98.94	13.03	5.00	2.34	70.00	20.00	49.11	119.05	107.78	44.94	17.99	20.00	32.95
96	RW_Dades	56.93	53.66	83.35	9.00	4.50	3.00	61.84	18.01	56.32	112.39	103.38	45.22	21.67	23.00	31.54
97	UW_Tdi	58.05	53.11	71.19	22.32	4.81	2.42	70.10	21.64	58.45	126.12	122.00	44.13	21.22	21.00	35.08
98	UW_Takwa	61.10	52.00	93.00	14.00	5.11	2.50	68.19	13.65	77.53	127.00	133.04	44.38	17.89	20.88	25.00
99	UW_AgnNk	58.5	56.31	74.58	13.00	4.00	2.40	69.74	20.00	75.40	118.55	130.00	40.95	18.15	23.57	28.63
100	UW_SBekw	58.91	57.35	89.39	11.20	3.40	2.13	55.00	19.00	67.23	116.44	121.12	41.33	21.22	24.30	31.00

APPENDIX B

QUESTIONNAIRE USED BY SRID-MoFA TO COLLECT THE DATA

Weekly Wholesale and Retail Price Questionnaire

Region:………………… District:………..…… Name of Market:………………. Reporting Month & Year:……………… Date of Submission:……………...

	Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7
		WHOLESALE PRICES			RETAIL PRICES		
No	*Commodity*	*Unit of Sale*	*Price*	*Remarks*	*Unit of Sale*	*Price*	*Remarks*
1	Maize	100kg			1kg		
2	Millet	93kg			1kg		
3	Guinea Corn	109kg			1kg		
4	Rice (Local)	100kg			1kg		
5	Rice (Imported)	50kg			1kg		
6	Yam (White)	100 tubers (250 kg)			1kg		
7	Cocoyam	91kg			1kg		
8	Cassava	91kg			1kg		
9	Gari	68kg			1kg		
10	Cassava chips (Kokonte)	40kg			1kg		
11	Plantain (Apentu)	9-11 kg			1kg		
12	Orange	100 singles (20kg)			1kg		
13	Banana	Av Bunch(6-8kg)			1kg		
14	Pineapple	100 singles(150kg)			1kg		
15	Mangoes	Crate (100kg)			1kg		
16	Tomatoes	Crate (52kg)			1kg		
17	Garden eggs	27kg			1kg		
18	Onion	73kg			1kg		
19	Ginger	48kg			1kg		
20	Dried Pepper	16kg			1kg		
21	Unshelled Groundnut				1kg		
22	Groundnut (Red)	82kg			1kg		
23	Cowpea (white)	109kg			1kg		
24	Groundnut Oil	18 litres			1 Litre		
25	Palm Oil	18litres			1 Litre		
26	Beef				1kg		
27	Pork				1kg		
28	Smoked Herrings	100 Singles			1kg		
29	Salted Dried Fish (Kobi)	100 Singles			1kg		
30	Eggs (Commercial)	Crate			Single		
31	Live Birds*				Single		
32	Anchovies	Average Basket					
33	Kpakpo Shitor	Average Basket					
34	Okro	Average Basket					

**Price not covered at the wholesale level*

APPENDIX C

Correlation Matrix

Variable	Mz	ImpR	WtY	Cv	Org	Ban	Tm	GEg	DrP	RdGn	WtCp	GnOil	PmOil	SmHr
ImpR	-0.227													
WtY	0.420	-0.301												
Cv	-0.009	0.287	-0.138											
Org	-0.132	0.359	-0.245	0.520										
Ban	-0.166	0.429	-0.221	0.713	0.747									
Tm	0.662	-0.310	0.322	0.127	-0.111	-0.119								
GEg	0.015	0.443	-0.391	0.619	0.641	0.618	-0.011							
DrP	0.215	-0.122	0.257	0.170	-0.071	0.028	0.353	-0.068						
RdGn	0.539	-0.419	0.339	-0.270	-0.395	-0.485	0.479	-0.274	0.189					
WtCp	0.490	-0.440	0.369	-0.276	-0.446	-0.468	0.488	-0.411	0.317	0.734				
GnOil	0.238	-0.137	0.191	0.319	0.178	0.208	0.449	0.227	0.197	0.071	0.000			
PmOil	-0.195	0.437	-0.134	0.602	0.539	0.686	-0.056	0.466	0.085	-0.494	-0.315	0.122		
SmHr	-0.264	0.437	-0.261	0.533	0.723	0.756	-0.196	0.475	-0.009	-0.501	-0.503	0.079	0.533	
Kobi	-0.285	0.367	-0.119	0.557	0.539	0.695	-0.096	0.405	0.149	-0.539	-0.414	0.152	0.677	0.618

APPENDIX D

Partial Correlation Matrix

Variable	Mz	ImpR	WtY	Cv	Org	Ban	Tm	GEg	DrP	RdGn	WtCp	GnOil	PmOil	SmHr
ImpR	0.080													
WtY	0.379	-0.037												
Cv	-0.059	-0.086	0.032											
Org	0.004	-0.132	0.049	-0.181										
Ban	0.126	0.010	-0.020	0.339	0.275									
Tm	0.501	-0.128	-0.121	0.133	-0.007	-0.154								
GEg	0.273	0.241	-0.407	0.336	0.356	0.044	-0.051							
DrP	-0.022	0.006	0.096	0.116	-0.089	-0.015	0.116	-0.034						
RdGn	0.067	0.000	0.117	0.064	0.045	-0.024	0.121	0.141	0.001					
WtCp	0.134	-0.119	-0.053	-0.028	-0.030	-0.021	0.133	-0.209	0.204	0.528				
GnOil	-0.103	-0.153	0.194	0.084	0.041	0.093	0.376	0.132	0.050	-0.002	-0.147			
PmOil	-0.112	0.220	0.105	0.195	0.102	0.196	0.115	0.049	-0.016	-0.289	0.237	-0.063		
SmHr	-0.041	0.153	-0.066	0.096	0.397	0.296	0.035	-0.172	0.065	-0.036	-0.096	-0.073	-0.104	
Kobi	-0.204	0.021	0.140	0.045	0.006	0.207	0.134	0.016	0.186	-0.186	-0.016	-0.010	0.244	0.127

APPENDIX E

Principal Components

	Components														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Mz	-.422	.657	-.312	.154	.313	-.052	.066	-.194	-.226	.022	-.043	.240	.043	-.018	-.113
RiImp	.613	-.169	-.113	.430	.232	-.474	.198	.019	.251	-.095	.003	-.022	-.076	-.037	-.010
YmWt	-.429	.411	.431	-.165	.591	-.030	.082	.216	-.076	.032	.040	-.115	-.020	.065	.066
Cv	.673	.493	-.041	.043	-.135	-.023	-.182	.291	-.212	-.288	-.064	-.054	-.115	.002	-.123
Org	.779	.215	-.188	-.039	.131	.354	.154	-.040	.097	.253	-.022	-.128	-.168	-.091	-.120
Ban	.865	.280	.016	.012	.071	.164	-.016	.069	-.006	-.059	-.052	.169	-.004	-.207	.241
Tm	-.338	.787	-.131	-.022	-.061	-.087	-.122	-.361	.032	-.126	.021	-.247	-.042	-.035	.104
GEg	.692	.270	-.502	.103	-.100	-.095	.046	.141	-.159	.204	.151	-.048	.051	.185	.124
PpDr	-.107	.544	.525	.288	-.397	-.059	.379	.008	-.114	.112	-.058	-.011	.018	-.025	-.002
GnR	-.725	.366	-.246	.218	-.020	.187	.030	.249	.203	-.061	.151	-.077	.204	-.128	-.053
CpWt	-.711	.382	.058	.334	-.063	.181	-.206	.071	.257	.049	-.035	.141	-.178	.156	.053
GnOil	.095	.633	-.094	-.619	-.152	-.251	.106	.090	.272	.035	-.051	.114	.009	.020	-.035
PmOil	.743	.251	.238	.157	.076	-.118	-.392	-.018	.084	.212	-.203	-.050	.179	.000	-.050
SmHer	.814	.102	.055	.020	.096	.317	.226	-.130	.130	-.247	-.098	.000	.135	.200	-.006
Kobi	.761	.210	.384	-.007	-.004	.001	-.135	-.134	.047	-.015	.419	.099	-.002	002	-.070

APPENDIX F

Varimax Rotated Components

No.	Variables	Component		
		1	2	3
1	Maize	-0.151	0.823	0.087
2	Imported Rice	0.484	-0.284	-0.320
3	White Yam	-0.192	0.280	0.650
4	Cassava	0.808	0.214	-0.013
5	Orange	0.782	0.016	-0.279
6	Banana	0.903	-0.051	-0.094
7	Tomatoes	-0.007	0.821	0.277
8	Garden Egg	0.697	0.233	-0.513
9	Dried Pepper	0.163	0.245	0.704
10	Red Groundnut	-0.537	0.647	0.108
11	White Cowpea	-0.493	0.517	0.381
12	Groundnut Oil	0.330	0.539	0.141
13	Palm Nut Oil	0.798	-0.137	0.122
14	Smoked Herring	0.789	-0.200	-0.117
15	Kobi (Salted Tilapia)	0.812	-0.244	0.230

APPENDIX G

Ranking of Markets on the First and Second Components

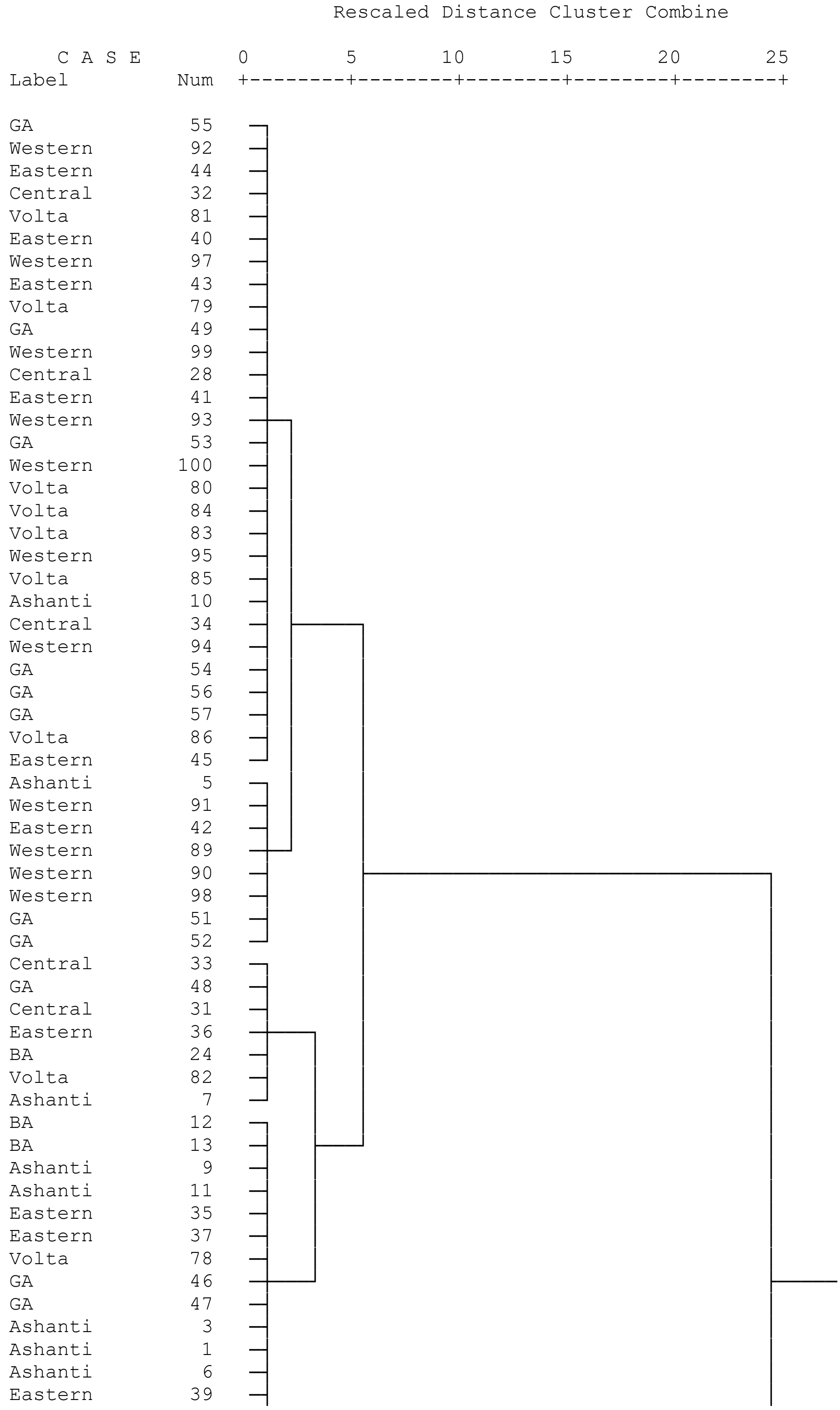
POSITION (Highest-priced First)	PRINCIPAL COMPONENT 1			PRINCIPAL COMPONENT 2		
	ID	MARKET AND ITS VARIABLE NAME	SCORES	ID	MARKET AND ITS VARIABLE NAME	SCORES
1	61	RN_Gushg	2.10552	52	UG_Mprb	2.00561
2	68	RUE_Bngo	2.10365	51	UG_Knsh	1.85147
3	75	RUW_Bussie	2.07900	53	UG_Tema	1.62330
4	77	UUW_Tumu	2.05541	44	UE_AOda	1.57816
5	74	RUW_Jirp	1.97885	55	UG_Madn	1.53098
6	70	RUE_Zebl	1.90273	72	UUE_Navg	1.43903
7	69	RUE_Garu	1.88066	54	UG_Ashm	1.41450
8	76	UUE_Wa	1.75114	49	UG_Mkl	1.41118
9	60	RN_Nalgu	1.74114	50	UG_Agbls	1.38284
10	73	RUW_Babl	1.69347	71	UUE_Bolg	1.37343
11	67	RUE_Fumb	1.55582	42	UE_NTafo	1.27674
12	14	RB_Yeji	1.33413	69	RUE_Garu	1.24185
13	64	RN_Zabzg	1.32798	77	UUW_Tumu	1.15983
14	63	RN_Kmbgu	1.29898	70	RUE_Zebl	1.13540
15	62	RN_Salga	1.18622	76	UUE_Wa	1.10639
16	66	UN_Tmle	1.17854	57	UG_Dome	1.05659
17	87	UV_Dmbai	1.11601	34	UC_Dunkw	1.03317
18	71	UUE_Bolg	1.05531	56	UG_Malm	1.01983
19	72	UUE_Navg	1.05045	43	UE_Asmk	0.98390
20	65	RN_Sboba	1.04862	41	UE_Suhum	0.94905
21	58	RN_Bimbi	1.00120	73	RUW_Babl	0.88813
22	21	UB_Brkm	0.93305	74	RUW_Jirp	0.87072
23	23	UB_Goas	0.87104	40	UE_Kdua	0.86178
24	20	UB_Techi	0.79687	75	RUW_Bussie	0.79543
25	59	RN_Dmgo	0.78644	81	RV_LAlkp	0.72994
26	17	RB_DNkw	0.76193	59	RN_Dmgo	0.68597
27	18	RB_Drobo	0.71916	85	UV_Kpndo	0.63084
28	16	RB_Nsawk	0.67830	15	RB_Kukuo	0.62807
29	22	UB_DAhnk	0.63531	86	UV_Hohoe	0.62775
30	19	UB_Suny	0.60613	46	RG_Kaseh	0.61355
31	7	RA_Agogo	0.52479	47	RG_AdFoa	0.60347
32	82	RV_Kptoe	0.42664	62	RN_Salga	0.57510
33	24	UB_Kinta	0.41336	87	UV_Dmbai	0.53393
34	31	RC_FNks	0.31106	80	RV_Kpve	0.53175
35	33	UC_Swed	0.27479	79	RV_Abot	0.44954
36	48	RG_Hobor	0.25882	65	RN_Sboba	0.42683
37	36	RE_Akoas	0.17788	45	UE_Mprso	0.42387
38	13	RB_KDns	0.08999	68	RUE_Bngo	0.38285
39	12	RB_Ateb	0.04553	67	RUE_Fumb	0.38001
40	3	RA_Agona	0.00193	63	RN_Kmbgu	0.32906

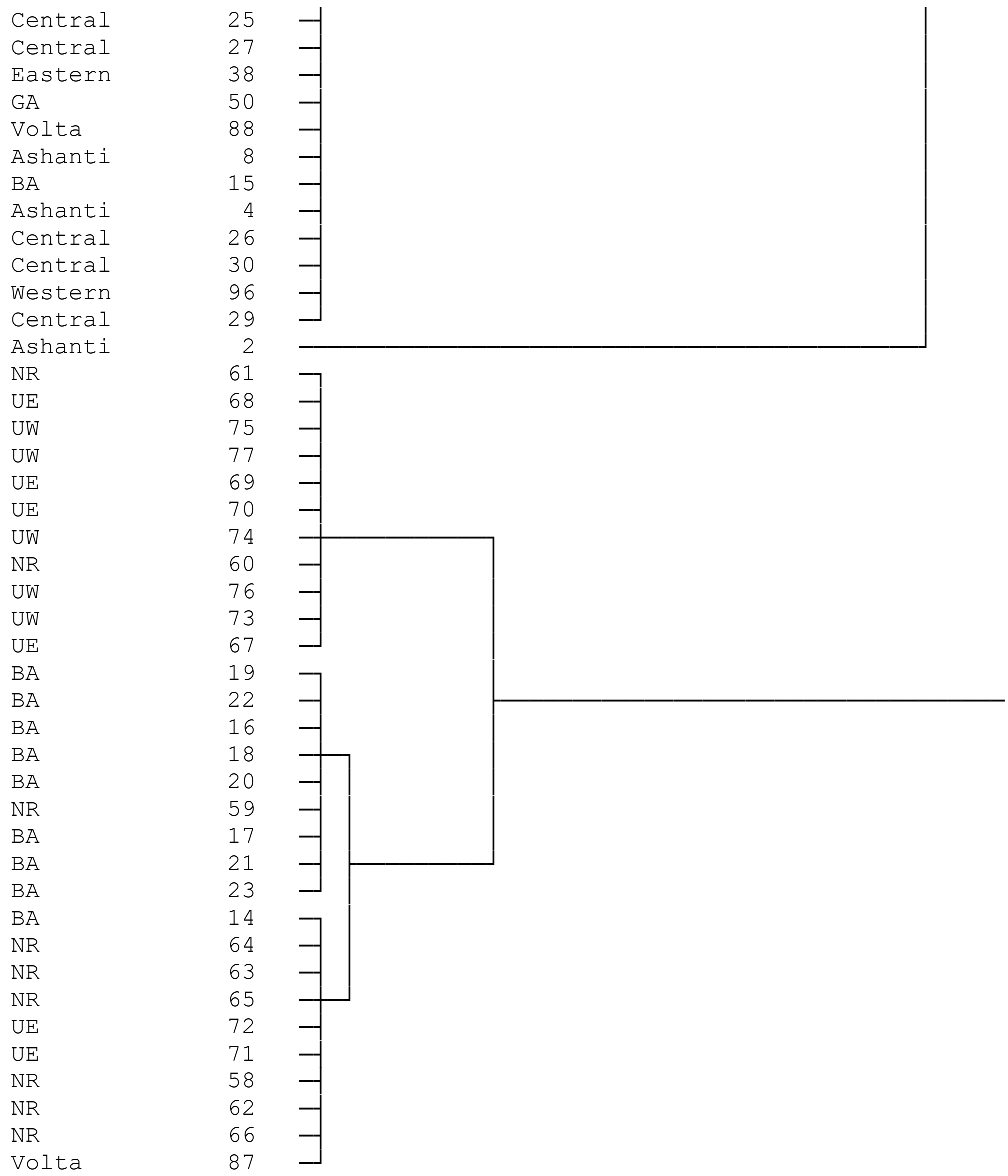
POSITION (Highest-priced First)	PRINCIPAL COMPONENT 1			PRINCIPAL COMPONENT 2		
	ID	MARKET AND ITS VARIABLE NAME	SCORES	ID	MARKET AND ITS VARIABLE NAME	SCORES
41	47	RG_AdFoa	-0.01819	98	UW_Takwa	0.32476
42	46	RG_Kaseh	-0.02736	66	UN_Tmle	0.30514
43	78	RV_MKse	-0.05474	60	RN_Nalgu	0.27683
44	37	RE_Ahom	-0.05638	88	UV_Denu	0.25074
45	9	UA_Ksi	-0.11956	58	RN_Bimbi	0.24744
46	11	UA_Obua	-0.12492	82	RV_Kptoe	0.24702
47	35	RE_Soma	-0.13043	99	UW_AgnNk	0.23959
48	39	RE_Agorm	-0.20172	84	UV_Ho	0.21390
49	6	RA_Juab	-0.24987	91	RW_Bibni	0.19752
50	1	RA_Ejura	-0.26470	28	RC_Elmna	0.17743
51	27	RC_Kasoa	-0.32512	64	RN_Zabzg	0.13396
52	25	RC_Manks	-0.33317	48	RG_Hobor	0.10272
53	38	RE_Anynm	-0.35398	78	RV_MKse	0.04226
54	29	RC_APrso	-0.38759	12	RB_Ateb	0.00529
55	4	RA_Abofo	-0.39868	97	UW_Tdi	-0.00514
56	26	RC_Bawj	-0.39881	89	RW_Bgoas	-0.05607
57	30	RC_Ajmk	-0.40636	83	RV_Kute	-0.06834
58	96	RW_Dades	-0.41147	100	UW_SBekw	-0.10672
59	8	RA_Nsuta	-0.4253	37	RE_Ahom	-0.11181
60	15	RB_Kukuo	-0.43843	32	UC_Cape	-0.11762
61	88	UV_Denu	-0.46586	93	RW_Tikbo	-0.32756
62	50	UG_Agbls	-0.49012	95	RW_Elubo	-0.33084
63	94	RW_Juabs	-0.55159	27	RC_Kasoa	-0.33737
64	34	UC_Dunkw	-0.57210	26	RC_Bawj	-0.34505
65	45	UE_Mprso	-0.61605	96	RW_Dades	-0.47214
66	54	UG_Ashm	-0.63633	25	RC_Manks	-0.50492
67	56	UG_Malm	-0.63904	35	RE_Soma	-0.54962
68	57	UG_Dome	-0.64425	13	RB_KDns	-0.55720
69	86	UV_Hohoe	-0.66195	61	RN_Gushg	-0.56183
70	10	UA_Bekw	-0.68635	38	RE_Anynm	-0.56977
71	85	UV_Kpndo	-0.71433	94	RW_Juabs	-0.58083
72	95	RW_Elubo	-0.73904	92	RW_SDwn	-0.64400
73	83	RV_Kute	-0.74556	39	RE_Agorm	-0.66110
74	53	UG_Tema	-0.76922	33	UC_Swed	-0.66321
75	100	UW_SBekw	-0.77479	36	RE_Akoas	-0.70339
76	84	UV_Ho	-0.78955	90	RW_Aswns	-0.72317
77	80	RV_Kpve	-0.80161	2	RA_Tepa	-0.75029
78	97	UW_Tdi	-0.85292	1	RA_Ejura	-0.80196
79	32	UC_Cape	-0.85778	23	UB_Goas	-0.84889
80	81	RV_LAlkp	-0.85852	20	UB_Techi	-0.88037
81	40	UE_Kdua	-0.86150	30	RC_Ajmk	-0.92919
82	79	RV_Abot	-0.88575	21	UB_Brkm	-1.00951

POSITION (Highest-priced First)	PRINCIPAL COMPONENT 1			PRINCIPAL COMPONENT 2		
	ID	MARKET AND ITS VARIABLE NAME	SCORES	ID	MARKET AND ITS VARIABLE NAME	SCORES
83	43	UE_Asmk	-0.89319	16	RB_Nsawk	-1.03122
84	92	RW_SDwn	-0.92358	11	UA_Obua	-1.04823
85	55	UG_Madn	-0.92359	29	RC_APrso	-1.05118
86	44	UE_AOda	-0.92541	31	RC_FNks	-1.08552
87	93	RW_Tikbo	-0.95949	3	RA_Agona	-1.08632
88	41	UE_Suhum	-0.99219	19	UB_Suny	-1.10522
89	28	RC_Elmna	-0.99389	22	UB_DAhnk	-1.20158
90	99	UW_AgnNk	-1.02710	14	RB_Yeji	-1.36969
91	49	UG_Mkl	-1.03054	17	RB_DNkw	-1.40653
92	42	UE_NTafo	-1.15241	18	RB_Drobo	-1.43967
93	89	RW_Bgoas	-1.17706	9	UA_Ksi	-1.51572
94	91	RW_Bibni	-1.19777	24	UB_Kinta	-1.51590
95	5	RA_Adugy	-1.21130	10	UA_Bekw	-1.66248
96	52	UG_Mprb	-1.25990	4	RA_Abofo	-1.87237
97	51	UG_Knsh	-1.31359	6	RA_Juab	-1.93614
98	90	RW_Aswns	-1.37946	5	RA_Adugy	-1.99962
99	98	UW_Takwa	-1.38763	8	RA_Nsuta	-2.16244
100	2	RA_Tepa	-2.26877	7	RA_Agogo	-2.59868

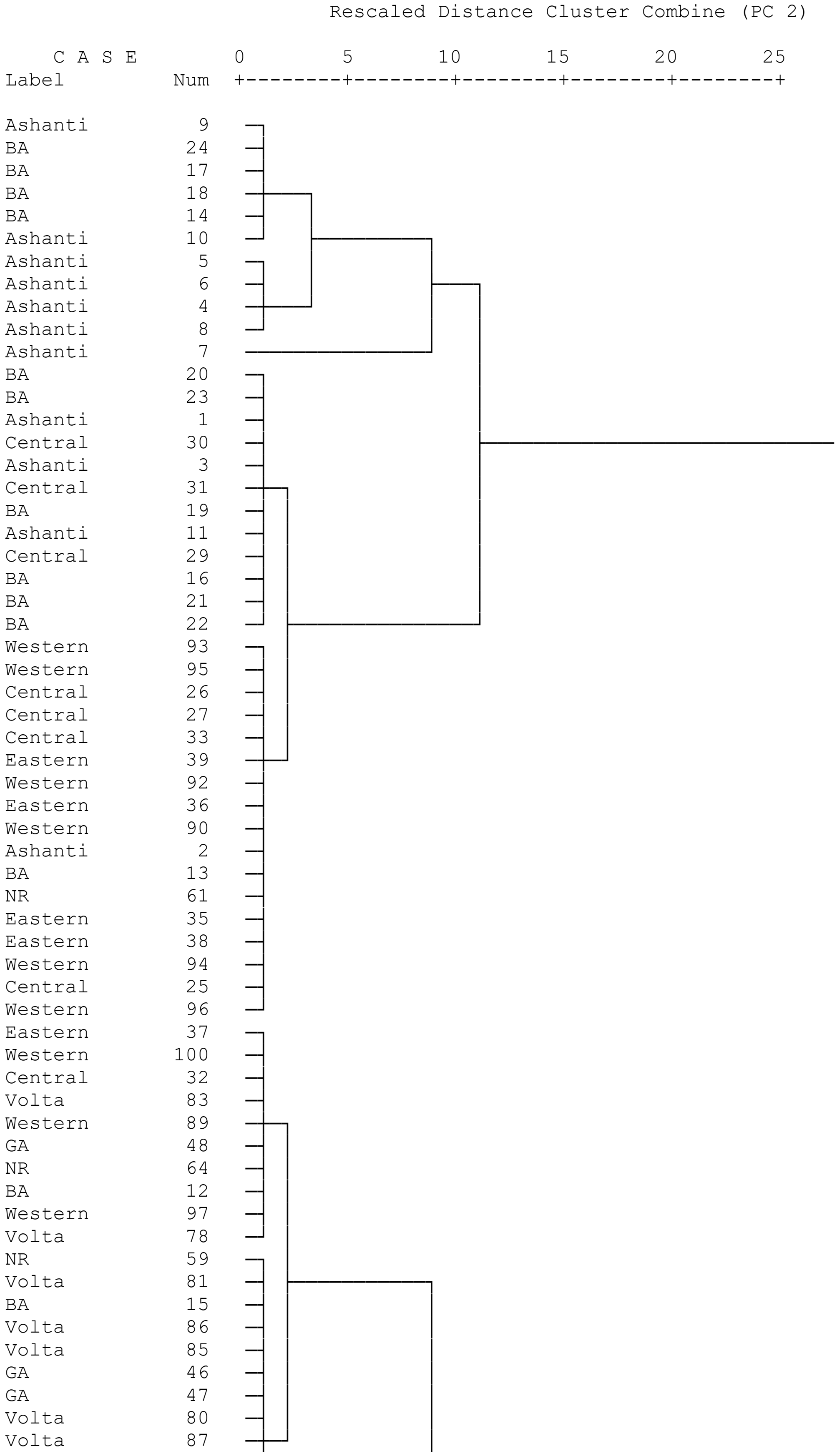
APPENDIX H
DENDROGRAM OF CLUSTER ANALYSIS OF THE FIRST PRINCIPAL COMPONENT

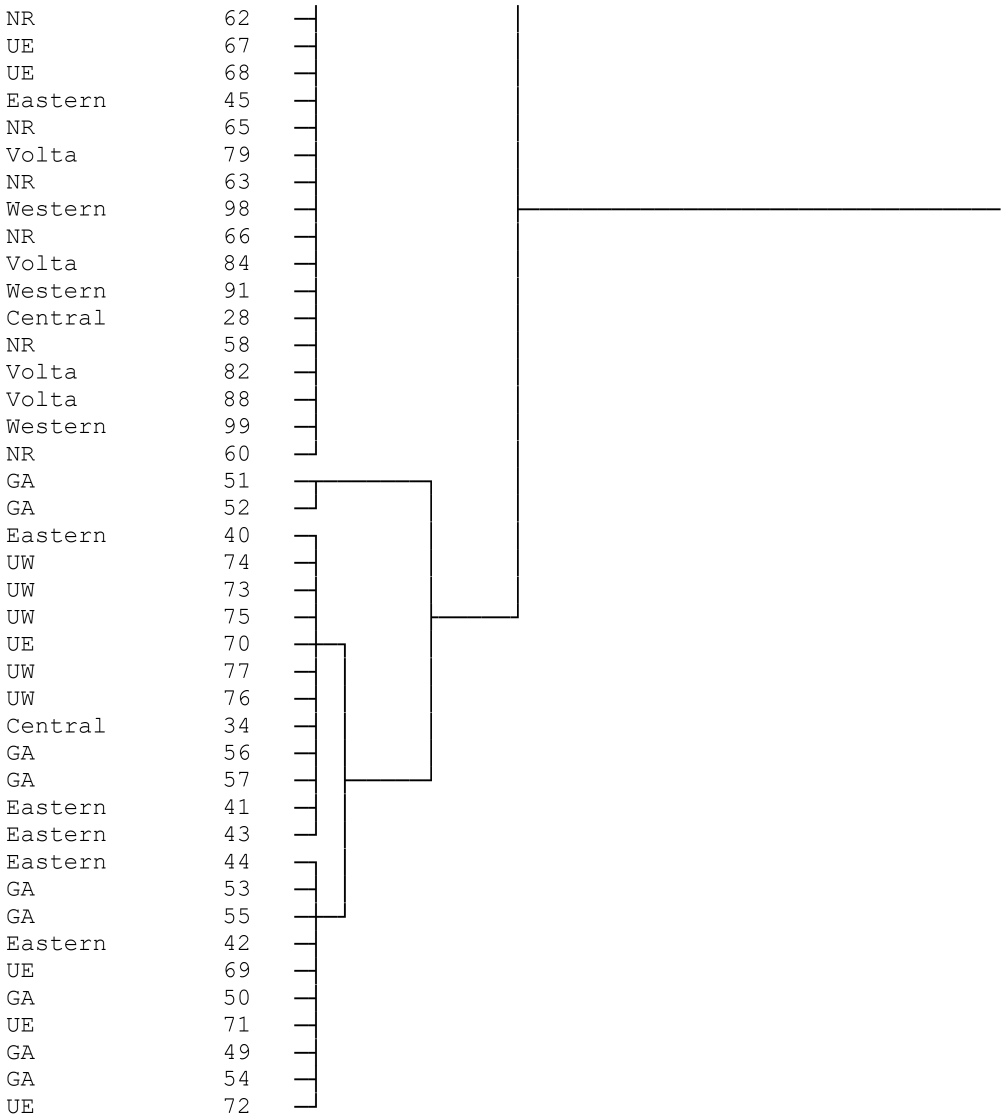
Dendrogram using Centroid Method (PC 1)




Central 25
Central 27
Eastern 38
GA 50
Volta 88
Ashanti 8
BA 15
Ashanti 4
Central 26
Central 30
Western 96
Central 29
Ashanti 2
NR 61
UE 68
UW 75
UW 77
UE 69
UE 70
UW 74
NR 60
UW 76
UW 73
UE 67
BA 19
BA 22
BA 16
BA 18
BA 20
NR 59
BA 17
BA 21
BA 23
BA 14
NR 64
NR 63
NR 65
UE 72
UE 71
NR 58
NR 62
NR 66
Volta 87

APPENDIX I
DENDROGRAM OF CLUSTER ANALYSIS OF THE SECOND PRINCIPAL COMPONENT




NR 62
UE 67
UE 68
Eastern 45
NR 65
Volta 79
NR 63
Western 98
NR 66
Volta 84
Western 91
Central 28
NR 58
Volta 82
Volta 88
Western 99
NR 60
GA 51
GA 52
Eastern 40
UW 74
UW 73
UW 75
UE 70
UW 77
UW 76
Central 34
GA 56
GA 57
Eastern 41
Eastern 43
Eastern 44
GA 53
GA 55
Eastern 42
UE 69
GA 50
UE 71
GA 49
GA 54
UE 72

APPENDIX J:

CATEGORIZATION OF MARKETS USING THE FIRST PRINCIPAL COMPONENT SCORES

1. HIGHEST-PRICED MARKETS

MARKET ID	MARKET	REGION
60	Nalerugu	Northern
61	Gushiegu	Northern
67	Fumbisie	Upper East
68	Bongo	Upper East
69	Garu	Upper East
70	Zebila	Upper East
73	Babile	Upper West
74	Jirapa	Upper West
75	Bussie	Upper West
76	Wa	Upper West
77	Tumu	Upper West

2. HIGH-PRICED MARKETS

MARKET ID	MARKET	REGION
14	Yeji	Brong Ahafo
16	Nsawkwa	Brong Ahafo
17	D/Nkwanta	Brong Ahafo
18	Drobo	Brong Ahafo
19	Sunyani	Brong Ahafo
21	Brekum	Brong Ahafo
22	Dorman Ahenkro	Brong Ahafo
23	Goaso	Brong Ahafo
58	Bimbila	Northern
59	Damongo	Northern
62	Salaga	Northern
63	Kumbungu	Northern
64	Zabzugu	Northern
65	Saboba	Northern
66	Tamale	Northern
71	Bolgatanga	Upper East
72	Navrongo	Upper East
87	Dambai	Volta

3. MODERATE PRICED MARKETS

MARKET ID	MARKET	REGION
1	Ejura	Ashanti
3	Agona	Ashanti
4	Abofour	Ashanti
6	Juaben	Ashanti
7	Agogo	Ashanti
8	Nsuta	Ashanti
9	Kumasi (Kejetia)	Ashanti
11	Obuasi	Ashanti
12	Atebubu	Brong Ahafo
13	K. Danso	Brong Ahafo
15	Kukuom	Brong Ahafo
24	Kintampo	Brong Ahafo
25	Mankessim	Central
26	Bawjiase	Central
27	Kasoa	Central
29	Assin Praso	Central
30	Ajumako	Central
31	Fante Nyankomase	Central
33	Swedru	Central
35	Somanya	Eastern
36	Akoase	Eastern
37	Ahoman	Eastern
38	Anyinam	Eastern
39	Agormanya	Eastern
46	Kasseh	Greater Accra
47	Ada Foah	Greater Accra
48	Hobor	Greater Accra
50	Agbogbloshie	Greater Accra
78	Mafi Kumase	Volta
82	Kpetoe	Volta
88	Denu	Volta
96	Dadieso	Western

4. LOW-PRICED MARKETS

MARKET ID	MARKET	REGION
5	Adugyama	Ashanti
10	Bekwai	Ashanti
28	Elmina	Central
32	Cape Coast	Central
34	Dunkwa	Central

MARKET ID	MARKET	REGION
40	Koforidua	Eastern
41	Suhum	Eastern
42	New Tafo	Eastern
43	Asamankese	Eastern
44	Akim Oda	Eastern
45	Mpraeso	Eastern
49	Makola	Greater Accra
51	Kaneshie	Greater Accra
52	Mamprobi	Greater Accra
53	Tema	Greater Accra
54	Ashaiman	Greater Accra
55	Madina	Greater Accra
56	Mallam	Greater Accra
57	Dome	Greater Accra
79	Abotoase	Volta
80	Kpeve	Volta
81	Logba Alakpeti	Volta
83	Kute	Volta
84	Ho	Volta
85	Kpando	Volta
86	Hohoe	Volta
89	Bogoso	Western
90	Asawinso	Western
91	Bibiani	Western
92	Sefwi Dwinase	Western
93	Tikobo	Western
94	Juabeso	Western
95	Elubo	Western
97	Takoradi	Western
98	Tarkwa	Western
99	Agona Nkwanta	Western
100	Sefwi Bekwai	Western

5. LOWEST-PRICED MARKET

MARKET ID	MARKET	REGION
2	Tepa	Ashanti

APPENDIX K

CATEGORIZATION OF MARKETS USING THE SECOND PRINCIPAL COMPONENT SCORES

1. HIGH PRICED MARKET

MARKET ID	MARKET	REGION
12	Atebubu	Brong Ahafo
15	Kukuom	Brong Ahafo
28	Elmina	Central
32	Cape Coast	Central
34	Dunkwa	Central
37	Ahoman	Eastern
40	Koforidua	Eastern
41	Suhum	Eastern
42	New Tafo	Eastern
43	Asamankese	Eastern
44	Akim Oda	Eastern
45	Mpraeso	Eastern
46	Kasseh	Greater Accra
47	Ada Foah	Greater Accra
48	Hobor	Greater Accra
49	Makola	Greater Accra
50	Agbogbloshie	Greater Accra
51	Kaneshie	Greater Accra
52	Mamprobi	Greater Accra
53	Tema	Greater Accra
54	Ashaiman	Greater Accra
55	Madina	Greater Accra
56	Mallam	Greater Accra
57	Dome	Greater Accra
58	Bimbila	Northern
59	Damongo	Northern
60	Nalerugu	Northern
62	Salaga	Northern
63	Kumbungu	Northern
64	Zabzugu	Northern
65	Saboba	Northern
66	Tamale	Northern
67	Fumbisie	Upper East
68	Bongo	Upper East
69	Garu	Upper East
70	Zebila	Upper East

MARKET ID	MARKET	REGION
71	Bolgatanga	Upper East
72	Navrongo	Upper East
73	Babile	Upper West
74	Jirapa	Upper West
75	Bussie	Upper West
76	Wa	Upper West
77	Tumu	Upper West
78	Mafi Kumase	Volta
79	Abotoase	Volta
80	Kpeve	Volta
81	Logba Alakpeti	Volta
82	Kpetoe	Volta
83	Kute	Volta
84	Ho	Volta
85	Kpando	Volta
86	Hohoe	Volta
87	Dambai	Volta
88	Denu	Volta
89	Bogoso	Western
91	Bibiani	Western
97	Takoradi	Western
98	Tarkwa	Western
99	Agona Nkwanta	Western
100	Sefwi Bekwai	Western

2. MODERATE MARKETS

MARKET ID	MARKET	REGION
1	Ejura	Ashanti
2	Tepa	Ashanti
3	Agona	Ashanti
11	Obuasi	Ashanti
13	K. Danso	Brong Ahafo
16	Nsawkaw	Brong Ahafo
19	Sunyani	Brong Ahafo
20	Techiman	Brong Ahafo
21	Brekum	Brong Ahafo
22	Dorman Ahenkro	Brong Ahafo
23	Goaso	Brong Ahafo
25	Mankessim	Central
26	Bawjiase	Central
27	Kasoa	Central

MARKET ID	MARKET	REGION
29	Assin Praso	Central
30	Ajumako	Central
31	Fante Nyankomase	Central
33	Swedru	Central
35	Somanya	Eastern
36	Akoase	Eastern
38	Anyinam	Eastern
39	Agormanya	Eastern
61	Gushiegu	Northern
90	Asawinso	Western
92	Sefwi Dwinaso	Western
93	Tikobo	Western
95	Elubo	Western
96	Dadieso	Western

3. LOW-PRICED MARKETS

MARKET ID	MARKET	REGION
4	Abofour	Ashanti
5	Adugyama	Ashanti
6	Juaben	Ashanti
7	Agogo	Ashanti
8	Nsuta	Ashanti
9	Kumasi (Kejetia)	Ashanti
10	Bekwai	Ashanti
14	Yeji	Brong Ahafo
17	D/Nkwanta	Brong Ahafo
18	Drobo	Brong Ahafo
24	Kintampo	Brong Ahafo