Index Processor

CPI Compilation System

Consumer Price Index (CPI) Compilation System User Guide

Price Index Processor
Version II: Consumer Price Index
(Revised Version)

Acknowledgements and disclaimer

The IMF Statistics Department (STA) and the Technology and General Services Department (TGS) developed the Price Index Processor Software (PIPS) to assist countries in improving their capabilities to process collected price observations and use them to compile price indices. Acknowledgements are due to Gangti Zhu for developing the software and Paul Armknecht for advising the development of index number issues. The IMF has authorized UNECE Statistical Division, with whom the IMF has no other affiliation, to distribute, modify, and maintain the software. While the IMF retains ownership rights to the original software, the IMF assumes no responsibility to users for support or maintenance and has disclaimed all liability for any errors that may exist in the software and for any other claims relating to the software.

This user guide is no more than a “guide” and the user is strongly advised to become very familiar with the software using a set of trial data, preferably from their own country, before adopting the software for use. It is for the user to decide on the basis of such a trial whether the software suits their needs. Neither the IMF nor UNECE Statistical Division are responsible for any errors of omission or commission in this documentation.

TABLE OF CONTENTS

Chapter

1

Introduction

A price index measures the proportionate, or percentage changes in a set of prices over time. Consumer Price Indices (CPIs) typically measure the changes in the price of a given basket of goods and services representing the purchases of a typical consumer. CPIs are used to measure the real purchasing power of money on various types of commodities and services. They are an important measure of the inflation experienced by households and changes in their cost of living. CPIs have become a key statistic for purposes of economic policy-making, especially monetary policy. It is often specified in legislation and in a wide variety of private contracts as the appropriate measure of inflation for purposes of adjusting payments, such as wages, rents, interests and social security benefits, for the effects of inflation.

Price Index Processor - Consumer Price Index System (CPI System) is an application software developed for assisting countries in compiling their CPIs. The system can also be used for teaching, training, and research purposes. Its main function is to compile CPIs for all items by product, by outlet[1], and by geographic location.

The CPI System follows the guidelines of the *Manual – Consumer Price Index Manual: Theory and Practice (2004), International Labor Office, et al. URL:* http://www.ilo.org/public/english/bureau/stat/guides/cpi/, available in several languages, which describes the international standards for CPI compilation. These standards draw upon the collective experience and expertise accumulated in many countries.

The CPI System calculates indices as weighted averages of the price changes for a specified set, or basket of consumer products, the weights reflecting their relative importance in household consumption in some period.

---

[1] In "Resolution Concerning Consumer Price Indices, adopted by the Fourteenth International Conference of Labor Statisticians", "Outlet" is defined as a shop, market, service establishment, or other place, where goods and/or services are sold or provided to consumers for non-business use.

The CPI System consists of five parts: (1) *Data Entry, Editing; (2) Data Diagnosis and Validation*; (3) *Item Weights Creating, Editing and Distribution*; (4) *Index Imputation*; and (5) *Tabulation and Reporting*. They are integrated into a single database environment. The System works with COICOP-ICP - *Classification of Individual Consumption by Purpose (COICOP) - International Comparison Programme* (ICP) product classification. It allows the user to choose one of the two alternative formulas: Jevons (lower level)-Laspeyres (higher level) and Jevons(lower level)-Geometric Laspeyres(higher level) in combination with three basic compilation methods: compile by (a) product; (b) geographical region; and (c) outlet.

The System handles the consumption expenditure weights (typically obtained from household expenditure surveys) at any (COICOP-ICP) product classification) aggregation level. The detail of the levels of such a weights system can be mixed, catering to the different consumption patterns of different countries. The software has an embedded engine to distribute the item weights equally to the individual products and varieties by regions and by products that comprise them. Thus elementary “unweighted” price changes are ascribed equal weighting within each elementary index. The System also allows the user to revise the outlet, product and variety weights, after the item weights have been distributed, which will reflect the relative importance of the products and varieties. The System can calculate adjusted weights based on the outlet status and, if applicable, the probability group it belongs to, the result of which will be used in index imputation. If an outlet status is changed, the recalculation procedure can be triggered to re-distribute the item weights.

The weights inputted can be price updated to the reference period or can be actual weights. For the arithmetic option the resulting formulas would be a Lowe index (price-updated weights) and Young index (without price updating). For the geometric option the weights for a geometric Laspeyres, (also referred to as a Geometric Young) should not be price updated. **Price updated weights should only be used with arithmetic aggregation.**

The System has flexible data input and output modules—the user can batch input and upload outlet information, item weights or price data using a spreadsheet and transfer them into the CPI system. Alternatively, the user can enter data into the system directly. Reports can be generated at different aggregation levels on an *ad hoc* basis or in time series format. All the compiled results have a tabulation format and can be easily exported to Excel, HTML formats.

The System allows the user to define additional dimensions of product/variety in addition to COICOP-ICP and SPD (*Structured Product Description*) level. This

variety list reflects the particular consumption pattern of a country and can be unique to a specific country. Each variety list has its own SPD and is closely integrated with COICOP-ICP and COICOP product classification systems. They are therefore internationally comparable, even though different countries will have different consumption patterns.

A geometric mean formula is used for the lower (elementary) level[2] calculation[3]. When used as an elementary aggregate, the geometric mean has superior axiomatic properties than alternative formulas. There is a choice between the Laspeyres or geometric Laspeyres formula for the upper level compilation of the index. Both versions take a "modified" or "two-stage" approach: the (Geo) Laspeyres formula is decomposed into three components: Short-Term Price Relatives ($STPR_t$) of the current period; Long-Term Price Relatives ($LTPR_{t-1}$) of the previous period, and Base Period Weight (W_0). The modified approach has several advantages over the standard formulation: it facilitates (i) the introduction of new varieties as soon as two successive price quotations are available; (ii) data verification; and (iii) more reasonable assumptions of similar short-run price changes for imputing missing values. The CPI is calculated using the equivalent of a recursive procedure, in which last period's cost weights (or base-weighted long-term price relatives, $p_{t-1}.q_0$) are updated by the current period's price relatives (p_t/p_{t-1}) in order to obtain the corresponding cost weights ($p_t.q_0$).

The price averages are calculated on the basis of "matched observations". Whenever a particular price observation is missing from either the previous month or the current month, the corresponding price observations are eliminated from the other period. This ensures a consistent sample of price quotations in each period. However, the system will be able to impute missing indices using next level data. Once the missing indices are estimated, the missing prices will be imputed and marked as red in the system.

The user can associate detailed meta data and notes (up to about 22 pages of an MS Word document) to an outlet, product/variety, or to a specific price observation. The CPI system has a flag to show whether the user has entered any meta data using different color coding. Each price observation can be flagged or un-flagged to indicate whether the price is imputed by the CPI system.

---

[2] i.e. when information on weights is unavailable or assumption of equal weights used.

[3] The software only allows the Jevons index at this level. A ratio of arithmetic means (Dutot index) at the lower level may alternatively be used at the lower level, but the CPI Manual recommends this only for strictly homogeneous goods and services. The arithmetic mean of price relatives (Carli index) is biased.

The user can also create a new outlet by cloning an existing one. Each outlet is associated with a label indicating its current status. The outlet status will determine whether the outlet will be included in the index calculation. This feature is important in conducting sample rotation and introducing new outlets. The user can also clone a variety, through which a quality adjustment is made possible. The software contains a facility for estimating the quality-adjusted base period price.

Both non-statistical checking and statistical checking routines are available for identifying the possible errors and outliers of input data. Non-statistical checking is implemented to check whether specific price data has increased more than a designated threshold, which is definable by the user. The recorded prices will be compared to the previous period's prices of the same items. While this procedure will detect unusual price changes, it is far from certain that all errors will be detected or that all unusual prices are errors.

Three statistical methods (or filters) have been implemented to detect the possible errors and outliers of short term price relatives. The first one is based on Chebyshev's theorem which applies to all possible price relative distributions. This theorem predicts that at least 88.8 percent of all the observations in a data set will lie in the range of the mean plus or minus 3 standard deviations and at least 75 percent will fall within the mean plus or minus 2 standard deviations. Alternatively, observations with z-score greater than 3 will be potential outliers. The second method is Box Plots, in which price relative data are re-sorted in ascending order, and median and 1st and 3rd quartiles are calculated. Observations that fall outside of outer quartiles are considered as possible outliers. The third method is to assume price relatives are log-normally distributed. Therefore, by transforming them into logarithmic form, the price relatives data will exhibit the normal distribution. The intervals are calculated multiplicatively, and any price relatives that lie outside of the mean plus or minus 2 standard deviations will be identified as the possible outliers.

The System will impute price relatives for missing items and missing prices. If no price quotation is entered for any of the varieties covered by an item, its price relative is imputed using (geometric) average price relatives from the entire commodity group of the missing item. Missing prices of one or some varieties of a particular item are estimated as the previous period's price multiplied by the current period's price relative of that item.

Technical Overview

The CPI System is a Microsoft Visual Basic® application that runs in the Microsoft Windows environment. It works with Windows

2000, Windows Server 2003, Windows XP and Windows Vista operation system. The System stores both cross-sectional (e.g. outlets, products and varieties information) and time series data (e.g. price quotations and price indices) in database. The System supports three database environment Microsoft Access®, Microsoft SQL server® and Microsoft SQL Express® (formerly known as MSDE). The design makes use of a relational database architecture and object technology.

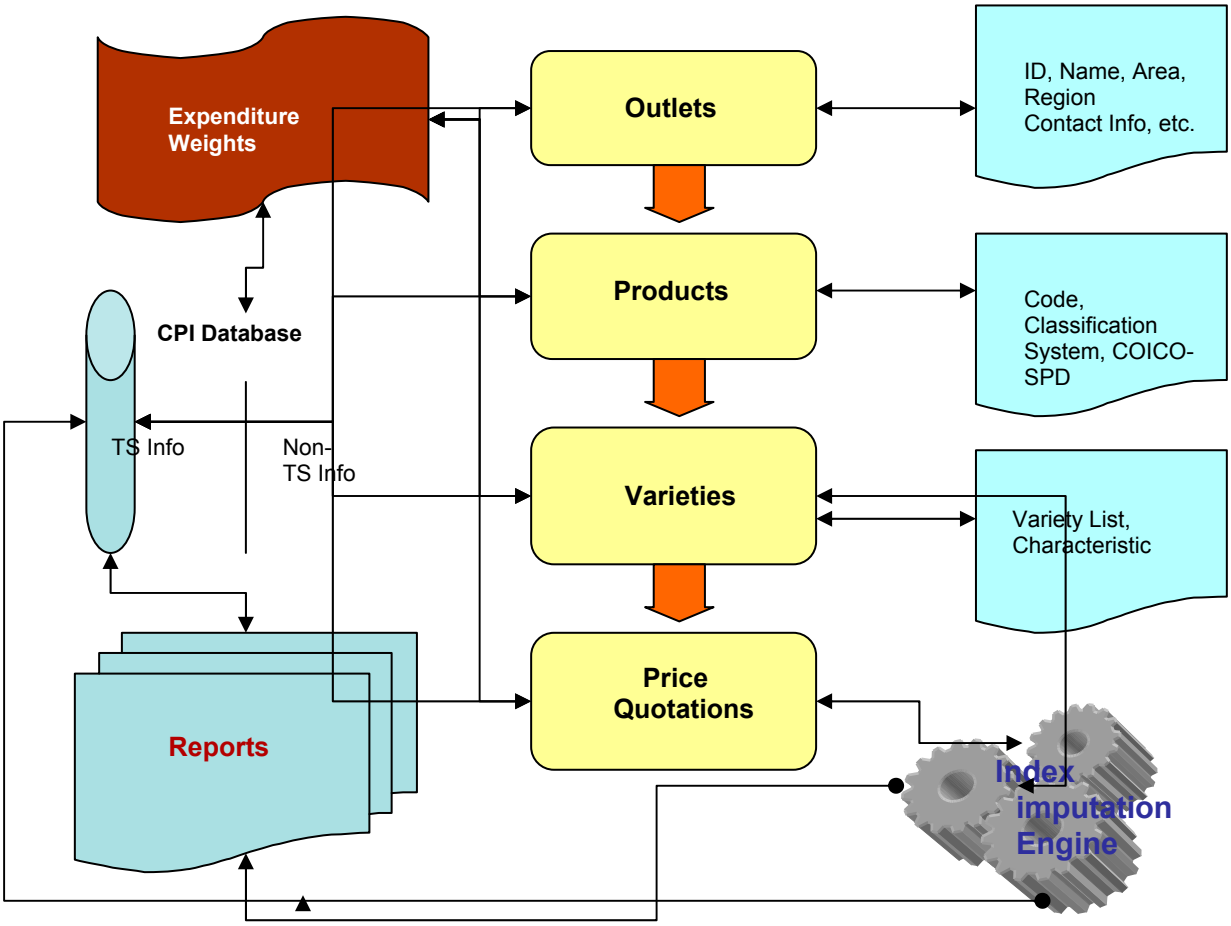
In the CPI System, several objects are generated representing outlet, product, variety and imputation. Each CPI System object contains an unique object name with its properties, attributes and methods.

Depending on the situation, the CPI System engine creates links between the database where information is stored and an Excel spreadsheet where information can be inputted, edited and reported. Almost all data can be entered through either a Windows form screen or an Excel spreadsheet. Using Excel, data can be uploaded in batch mode.

The CPI System brings together the familiarity of Excel with a powerful statistical and index compilation tool.

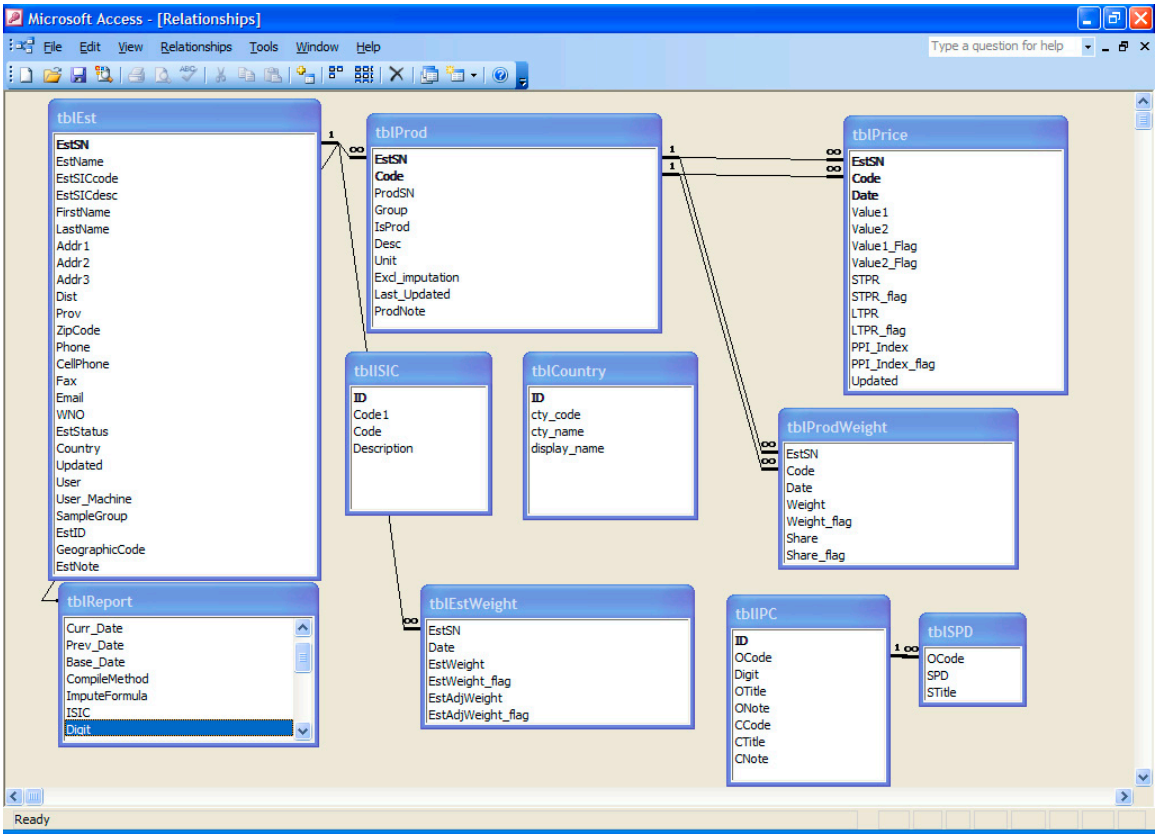
The CPI System Architecture

The diagram below shows the CPI System design and architect.



The CPI System Database Design

The CPI system allows the user to define, construct, and manipulate the underlying database using Microsoft ActiveX® Data Objects (ADO), the same type of interface and library used by Access to programmatically access data. The database table design and relationships are illustrated in the following diagram.



Chapter 2

Main Features

Outlet, Product, and Variety Model

The System is designed and developed based on outlet, product, and variety model. It stores the outlet information as the key element. Each outlet can carry multiple products; each product can have multiple varieties. The weights for outlets are derived form the weights for the products and varieties sampled within them. Such weights information is typically obtained from a household expenditure survey. The user can define and manipulate the weights. Every outlet is also classified into a specific geographic area for compilation purpose. Each outlet is associated with a list of properties including contact information for both outlet and data collector.

The user can easily add, edit, and delete an outlet. One can also input and edit the information on many outlets in a spreadsheet and upload it to the CPI database. The System also keeps track of the outlet's status and sample group information. The outlet's status indicates the current standing of the outlet as one of the following: (1) refuse to participate; (2) out of business; (3) could not locate; (4) initiated and reported; (5) resending for initiation; (6) no relevant product; (7) not yet initiated; and (8) unknown. Every outlet is classified into a sample group, either with certainty or probability.

The user can quickly find a specific outlet by using the search feature. An outlet can be searched by its ID, name, keyword, area it is located, or alphabetically.

Product and Variety Structure

The System allows the user to create product list for each outlet by selecting from a COICOP-ICP product classification scheme, which is displayed in a hierarchical tree view structure, outlined below. The user can not only search for a particular product by keywords in description field, but also search for detail notes associated with the product. The product list contains 7-digit detail level –

COICOP-ICP at the 6-digit level and an additional extended level for Structured Product Descriptions (SPDs).

After the product is selected, the user can create a variety list which is defined under this product. At this 8-digit level the user has the freedom to create their own variety list, which can be specific to her country. The user is first given a list that has previously been created. If she finds that the variety that needs to be entered is in the list, she can select it. Alternatively, she can add a new variety that will be stored in the database and shared across the outlets for that product. In this way, the system provides an additional dimension for the user to reflect her country's own unique consumption pattern. At the same time, each variety is well integrated into a standard product classification system through COICOP .

Multiple Ways of Entering the Price Quotations

Once the products and varieties are defined, the user is ready for the price observation input. The system offers multiple ways for entering the price quotes. The user can enter the prices for (a) all varieties in one outlet for three periods (base, previous and current period); or (b) one variety for multiple periods in time series format; or (c) one variety across all outlets in one year. The user can also generate an Excel spreadsheet for price editing and inputting. This spreadsheet can be sent to a remote user for entering price and validation and is reusable. Once validated data has been entered, it can be uploaded into the database.

For both time series and cross-outlet price entry, the system provides threshold validation feature, in which the user can define a threshold value. The user defined threshold value will apply globally until next time the program re-starts, which will reset to default value of 20%. When the price entered is greater or equal to the specified threshold, that particular number will be shown in red to warn the user of a possible data entry error for the price quotation.

The system can also estimate the base period price of any particular variety using another calculated time series index as the deflator. The user can specify an imputed index series, select a particular period, and its index will be used to estimate the base period price. This feature is very useful in the case when a new product or variety is introduced to the System where base period price is not observable.

Item Weights Design and Distribution

Item weights information are typically obtained from a household expenditure survey. The weights are used to reflect the relative importance of the goods and services as measured by their shares in the total consumption expenditure of

households. The weight attached to each good or service determines the impact that its price change will have on the overall index.

The system requires weight data by area and by product. The user first has to determine the level of details of the weights that is available to the country, usually at 6 (COICOP-ICP) or 7 (SPD) product classification level. Then the user has to select from a list the relevant items for that particular country. A spreadsheet template will then be generated with different items/products in rows and weight areas in columns. The user has to key in the weight information in this matrix table and save to the CPI database.

Item weight by area of a particular item/product will be distributed to all varieties under this item equally. But the user can modified the weights after distribution to specify the relative importance of varieties if necessary. Manually defined weights will replace/overwrite the auto-distributed weights and stored in the database for imputation.

Data Diagnosis and Error Checks

The System has a built-in module for the user to diagnose data errors and check for possible missing, critical variables. The System then generates a detail report if errors are detected. By double clicking on the error item, the user goes directly to the screen where such errors can be corrected. The user can also select the error item, right click, edit outlet, product, variety, weights or prices where the problem can be examined and corrected.

Data Status Indicator

The System has a graphical indicator to show the availability of all the outlet data for the compiling period. Indicators inform the compiler whether the price data are currently (1) available; (2) partially available; or (3) not yet ready. Data Status Indicator will also show outlet status (whether the outlet is initiated and reported or out of business, etc). By clicking on the outlet, the user will be prompted a screen where a price can be entered or updated. This provides a useful and intuitive tool for the statistical staff at central office who can follow up any missing outlet price data and prepare for final compilation of the indices.

Product Classification, SPD Extension and Varieties

The System uses COICOP-ICP as its default product classification scheme and displayed in hierarchical tree view structure for the user to select. The most detail level of COICOP-ICP has been extended by an additional level of SPD for details

of structured product. The user can select a SPD which carries a code that combines the COICOP-ICP code and SPD extended code. The SPD is the parent code for the varieties, which can extend to one more SPD level. The user has complete freedom to build her own variety list with the constraint that each variety must have a parent who is a member of SPD.

Calculating Average Prices

The CPI System has a separate module to calculate the average prices for the varieties across the outlets. The user has an option of using a geometric mean or arithmetic mean in such calculation. Since the Jevons index (geometric mean) is used for CPI compilation at the elementary level, the geometric option is recommended. Missing prices are imputed based on the average prices of identical varieties in the same unit of measurement. Such varieties are maintained in the variety list by the System.

Compilation Method and Elementary Formulae

The CPI System can compile price indices by (1) product; (2) outlet; and (3) geographical area. Missing price indices will be imputed. The compilation results are stored in the database in time-series format.

Tabulation and Reporting

The System generates both *ad hoc* and time-series reports in a tabular format. *Short-term price relatives* (STPR), *long-term price relatives* (CPI indices), and *updated cost weights* are the three key indices produced by the System. It reports the most detailed indices for all components. The user can select the report detail at different aggregation levels (from level 1 to level 8). The reports can be exported to Microsoft Excel spreadsheets, HTML for web publication, and other formats.

Chapter

3

Methodology

Many countries use the standard Laspeyres-type arithmetic mean of price relatives to compile their consumer price indices. Although the term ”Laspeyres” is often used to describe the formula, three points should be noted. First, Laspeyres requires that the weights reference period is the same as the price reference period, which is generally not the case. Weights may be from some prior survey period, say 2008, it taking some time to compile the weights for use with a price reference period of, say, January 2010. As noted below, the resulting index may more formally be a Young or Lowe index, depending on whether the weights are price-updated. The term “Laspeyres” or “Laspeyres-type” is used hereafter with this in mind. Second, the CPI is generally considered as being compiled in two stages, the elementary level using an equally-weighted geometric mean (Jevons) index and the weighted higher level using a Laspeyres or geometric Laspeyres.[4] Third, a modified/two-stage formulation will be used, as outlined below.

Elementary Index Formulas

The *CPI Manual* (Chapters 1 and 20) favors the use of geometric mean formula (Jevons index) on axiomatic grounds. Jevons index is defined as followings:

$$(1) P_J = \prod_{i=1}^{N} \left(\frac{p_i^t}{p_i^0} \right)^{1/N}$$

The arithmetic mean of price relatives (Carli index) is biased, especially in a chained form, and the ratio of arithmetic means of prices (Dutot index) is only suitable for strictly homogeneous varieties.

The Standard Laspeyres-type Formula

The standard Laspeyres-type formula, applicable up to the most detailed level of weighted items in the CPI basket, compares the current period cost of the base period market basket (the numerator) with the cost of the base period market basket (the denominator) and can be written as:

[4] Users should note that the distinction between the two levels in the software is dictated by the classification structure used which is explained below. Aggregation at level 8 is elementary level aggregation at which weights are equally distributed from some higher level, generally level 6 or 7. If the option is selected to use a geometric aggregator (Jevons) at the elementary level, and an arithmetic aggregator (Laspeyres-type) at upper levels, the software will define the elementary level as level 8 and use Jevons and the higher levels as levels 7 to the overall CPI, and use a Laspeyres-type aggregator. The user should define products at different levels with their requirements for the use of these different formulas in mind.

$$(2) I_{0\to t} = \frac{\sum_{i=1}^{N} q_i^0 \; p_i^t}{\sum_{i=1}^{N} q_i^0 \; p_i^0} = \sum_{i=1}^{N} \left[\frac{q_i^0 \; p_i^0}{\sum_{i=1}^{N} q_i^0 \; p_i^0} \times \left(\frac{p_i^t}{p_i^0} \right) \right]$$

where *i = 1, ... , n* stands for the products comprising the consumption basket and symbols *0,t* respectively designate the price reference period (or the base price period and the current price period). The symbols *p* and *q* designate the prices and the quantities of the products in question, respectively. The ratio ($p_{t,i}$ $/p_{0,i}$) is the price relative to the base period for item *i* (sometimes called the long-term price relative).

By expressing the consumption expenditure's share for the item *i* as a ratio of the total expenditure during the base period as: [5]

$$(3)\; w_i^0 = \frac{q_i^0 \; p_i^0}{\sum_{i=1}^{N} q_i^0 \; p_i^0}$$

Using the preceding expression formula (2) can be written in a slightly different form as:

$$(4) I_{t\to 0} = \sum_{i=1}^{N} w_i^0 \left(\frac{p_i^t}{p_i^0} \right)$$

However, these versions of the Laspeyres formula do not provide the flexibility required for economies that are going through significant and rapid changes.

The Modified or Two-stage Laspeyres Approach

There are several reasons why the Modified Laspeyres Approach is superior to the standard formula. First, in the standard formula, we are comparing price relatives for the current period to the base period. In practice, the editing of the current period's price data is done by comparing the prices for the collection period for an item with those charged for the same item in the previous period.

[5] The consumption expenditures should be adjusted for price changes occurring between the date of the household budget survey and that of the CPI rebasing, the first time these expenditures are used for the CPI compilation.

Any large variations falling outside a predetermined range checks (e.g. 0.8000 to 1.1000) might indicate either the wrong item has been priced or some kind of error has been made in recording the price. With formula (4), this comparison cannot easily be made as it uses, for each item *i*, the price relatives of current period to the price reference period $\left(\frac{p_i^t}{p_i^0}\right)$.

Second, the standard formula involves a comparison of changes in prices for each item over long time periods, requiring the continuity of priced item specifications. In practice varieties become permanently missing or unrepresentative and need to be replaced with new varieties for which there is no price in the reference period 0 to compare with. In these circumstances, it is advisable to apply a modified version of the Laspeyres formula that makes use of $\left(\frac{p_i^t}{p_i^0}\right)$, the price relative to the previous period so that a new variety can be introduced as soon as two successive price quotes are available.

Third, when varieties are temporarily missing imputed prices may be used based on the overall price change of the product group in question. Imputations over the short run are likely to be more reasonable than long-run ones...

The basic formula for computing the CPIs can be written as:

$$(5)\quad I_{0\to t} = \frac{\sum_{i=1}^{N}\left(\frac{p_i^t}{p_i^{t-1}}\right)\times p_{t-1_i}\, q_{0_i}}{\sum_{i=1}^{N} p_{0_i}\, q_{0_i}} \times 100$$

$$\text{where}\quad p_{t-1,i}q_{0,i} = p_{0_i}\, q_{0_i} \times \frac{p_{1_i}}{p_{0_i}} \times \frac{p_{2_i}}{p_{1_i}} \times \ldots \times \frac{p_{t-1_i}}{p_{t-2_i}}$$

Formula (5), which is arithmetically equivalent to formulas (2) and (4), is considered more versatile than the formula using long-term price relative to the base period, as the linking process used facilitates the introduction of new varieties and/or items or substitution when the need arises and enables more reasonable imputations.

Formula (5) can also be rewritten as:

$$(6) \quad I_{0\to t} = \sum_{i=1}^{N} w_i^0 \times \left(\frac{p_i^t}{p_i^{t-1}} \right) \times \left(\frac{p_i^{t-1}}{p_i^0} \right)$$

which can be interpreted as:

$$(7) \quad I_{0\to t} = \sum_{i=1}^{N} w_i^{t-1} \times \frac{p_i^t}{p_i^{t-1}}$$

where $w_i^{t-1} = w_i^0 \times \frac{p_i^{t-1}}{p_i^0}$ is an updated weight sometimes referred to as a "cost weight" of item *i*.

In other words, to obtain the index for the current period *t*, the Modified Laspeyres Approach involves multiplying individual price relatives of the latest price compared period ($\frac{p_i^t}{p_i^{t-1}}$) by the previous period's updated weight (w_i^{t-1}), and then summing them.[6]

The Modified Laspeyres formula has obvious advantages over the standard Laspeyres formula when we consider the problems arising from permanently unobservable varieties, and the need in due course to the bring in a new variety to replace the missing one. There is a need to impute a base period price if the standard Laspeyres formula is used. Such imputation is unnecessary while using the Modified Laspeyres formula, in which case the current period weight for the replacement item is obtained by simply multiplying the last updated weight for the replaced item by the current period's short-term price relative of the replacement item.

The system uses the modified Laspeyres approach to calculate CPI based on monthly price quotations (or monthly average price quotations) and weights information. The price index is assigned a value of 100 in the base period and value of the index for other periods of time, which indicate the average proportionate, or percentage, change in price levels.

---

[6] Formula (6) can also be interpreted as:

$$I_{0\to t} = \sum_{i=1}^{n} w_{0,i} \times STPR_{t\to t\text{-}1,i} \times LTPR_{t\text{-}1\to 0,i}$$

where $STPR_{t\to t\text{-}1,i}$ is the short-term price relative of item *i* for current period (= p_i^t / p_i^{t-1}) and $LTPR_{t\text{-}1\to 0,i}$ is the long-term price relative of item *i* for previous period (= p_i^t / p_i^0).

Instead of holding expenditure weight reference period at 0, The CPI System allows the user to compile a CPI as a weighted geometric/arithmetic average of the individual price relatives holding constant the expenditure shares at period *b*[7]. The resulting index is called a *Young* index. In this case, the formula should be revised as:

$$(8) \quad I_{0\to t} = \sum_{i=1}^{N} w_{b,i} \times \left(\frac{p_i^t}{p_i^{t-1}} \right) \times \left(\frac{p_i^{t-1}}{p_i^0} \right)$$

The weight reference period *b* is likely to precede price reference period 0 because it takes time to collect and process the expenditure data. For example, a monthly CPI may run from January 2010 onwards, with January 2010=100, but the quantities may be derived from the annual expenditure survey made in, say, 2008 or June 2008 to July 2009.

In that case, we have the choice of assuming that *either* the quantities of period *b* remain constant *or* the expenditure shares in period *b* remain constant and equation (8) does the latter. A Lowe index holds quantities constant in period *b*, and is given by:

$$(9) \frac{\sum_i p_i^b q_i^b \frac{p_i^o}{p_i^b} \frac{p_i^t}{p_i^0}}{\sum_i p_i^b q_i^b \frac{p_i^0}{p_i^b}} = \frac{\sum_i p_i^t q_i^b}{\sum_i p_i^0 q_i^b} = \sum_i w_i^{b,0} \frac{p_i^{t-1}}{p_i^0} \frac{p_i^t}{p_i^{t-10}}$$

Where $w_i^{b,0}$ are price updated weights given by:

$$(10) w_i^{b,0} = \frac{p_i^b q_i^b \frac{p_i^o}{p_i^b}}{\sum_i p_i^b q_i^b \frac{p_i^0}{p_i^b}}$$

Since the user is responsible for entering the weights they can enter price updated or weights or weights without price updating as in (9) and (8) respectively.

---

[7] As discussed in the *CPI Manual* (2004), "…any set of quantities could serve as the basket. The basket does not have to be restricted to the quantities purchased in one or other of the two periods compared, or indeed any actual period of time. … For practical reasons, the basket of quantities used for CPI purposes usually has to be based on a survey of household consumption expenditures conducted in an earlier period than either of the two periods whose prices are compared."

The Geometric Laspeyres or Young Indices

In the geometric version of the modified Laspeyres index, a weighted geometric average is taken of the price relatives using the expenditure shares of period 0 as weights. It is defined as:

$$(11) \quad I_{0\to t} = \prod_{i=0}^{N} \left[\left(\frac{p_i^{t-1}}{p_i^0} \right) \times \left(\frac{p_i^t}{p_i^{t-1}} \right) \right]^{w_i^0}$$

Similarly, the geometric version of the *Young* Index if period $b \neq 0$, that is, the expenditure shares are different from price reference period 0:

$$(12) \quad I_{0\to t} = \prod_{i=0}^{N} \left[\left(\frac{p_i^{t-1}}{p_i^0} \right) \times \left(\frac{p_i^t}{p_i^{t-1}} \right) \right]^{w_i^b}$$

Users should note that period *b* weights should not be price-updated to period 0 if using the Geometric formula, i.e. a geometric Young may be used but not Geometric Lowe.

Further, whether the index is a Geometric Laspeyres or Geometric Young depends on whether the user enters weights for period 0 or a preceding period *b*, and whether the index is an arithmetic Laspeyres or Young or Lowe depends on whether the user enters weights for period 0 or a preceding period *b*, or price updated weights from *b* to 0. The software uses the terminology “Laspeyres” of “Laspeyres-type” leaving it to the user to define the exact nature of the formula by virtue of the weights used. Laspeyres-type formulas are generally used by countries at the higher level and geometric means at the lower level. In spite of this, the geometric Laspeyres-type index has some advantages. Geometric means are (i) not as sensitive as arithmetic means to the extreme values, (ii) are circular, i.e., fulfill a multi-period transitivity property that the product of the price index change going from a period 1 to a period 2 times the price index change going from period 2 to a period 3 should equal the price index going directly from period 1 to 3; and (iii) are more likely to lie between the Laspeyres and Paasche bounds, a desirable property.

Matched Price Observations

An average price is calculated in each of the geographical areas covered and for each variety comprising the CPI basket. The system allows a different number of

areas and variety structure. The calculation of average prices would be simple if a set of price quotations were available for the current and previous month. In reality, this does not always happen. Quite often, some of the respondents are unable to quote a price for a particular variety because it is out of stock. Whenever a particular price observation is missing from either the previous month or the current month, the corresponding price observations are eliminated from the other period. This is equivalent to imputing the price of variety 1 in period *t* by the short-run price change of the other varieties in the product group. This ensures that the price averages are calculated on the basis of "matched observations", i.e., a consistent sample of price quotations in each period.

In the following example we consider that item's prices are collected for four representative varieties *1*,*2*,*3*, and *4*. In the current month *variety 1*'s price cannot be collected (is missing).

Variety *v* of item *I*	**Month *t-1***	**Month *t***
Variety 1	1.50	-
Variety 2	1.25	1.25
Variety 3	1.25	1.50
Variety 4	1.50	1.50
Average	1.37	1.41
Average for matched observations (Variety 2, Variety 3, Variety 4)	**1.33**	**1.41**
Short-term relative for item *i*	**1.0627 (=1.41/1.33)**	

For the CPI calculation of month *t*, the geometric average price for month *t-1* should be recalculated based on matched observations as $(1.25*1.25*1.50)^{(1/3)}$ = ***1.33*** and not $(1.50*1.25*1.25*1.50)^{(1/4)}$ = ***1.37***.

The month's short-term price relative for item *i* is then 1.0627 (=1.41/1.33), and not 1.0309 (=1.41/1.37).

Impute Missing Indices & Prices

The CPI System program will impute missing STPRs, LTPRs, and updated weights based on the information available from other price quotations of varieties in that commodity group. Imputed indices and prices will be stored database with a flag. These imputed indices and prices are shown in the tabulation in red color.

Missing price index is estimated using its parent index as the proxy, i.e., if a specific variety's index is missing due to the missing prices, the index of the product or item it belongs to will be taken to be representative. The system always uses the next available level index data in the same group or item for the missing index, e.g., if level 7 index is missing, level 6 index will be used; if level 6 index is missing, level 5 index will be used so on so forth. Holding missing prices for a variety constant by carrying the last observation forward (i.e., making the short-term price relative for that variety equal to 1.0) during a period of high inflation would cause short-term distortion in the index, because it would understate inflation while the variety was unavailable and then show a large increase in the index when the variety became available. *The system therefore does not use carry-forward prices.* We do not use the price for the same variety in another outlet to represent the missing price in this outlet. Thus if Coca-Cola price is missing in one particular outlet, the system will not take the other outlet's Coca-Cola price as a proxy, instead it will take soft drink index in the same outlet to represent that of Coca-Cola.

Since parent group price changes are always calculated as geometric mean changes, imputations are based on geometric means.

Missing prices of one or some varieties then are estimated by multiplying the previous period's price by the current period's short-term price relative of that variety, which in turn was estimated using the index of the item/group. If the previous period's price is not available, the missing price will be estimated by multiplying the reference/base period price by the LTPRs. If both previous period price *and* base period price are not available, missing price will/can not be imputed.

If no price is collected for any variety covered by a product (the prices for whole product is missing), Its price relatives will be imputed using average price relatives from the item group of the missing price.

Detection of Outliers

What is an outlier

An observation that is unusually large or small relative to the other values in a price relatives data set is called an outlier. Outliers are the observations that appear to be inconsistent with the remainder of the collected data.

There are several possible sources for outliers:

1. The price quotation of a transaction or variety is observed, recorded, or entered into the computer incorrectly.

2. The price quotation come from a different population, or quality of that transaction/variety has been changed.

3. The price quotation entered is correct, but represents a rare event or novel phenomenon.

Outliers occur when the relative frequency distribution of the data set is extremely skewed. Such distributions have a tendency to include extremely large or small observations.

The PPI/CPI software implemented two procedures to identify the possible errors and outliers. The first one is non-statistical procedure, which is to find whether specific price observation falls outside some pre-specified acceptance interval. In the “Input Price in Time Series Format” screen, the user can specify a threshold value (default is 20%), for any price change is greater than the threshold, the increased percentage will be shown in red and bold which indicates the possibility of errors or outliers. The second one is the statistical procedure, in which three methods are implemented. In both cases outlier detection should *not* result in automatic deletion. Often price changes are undertaken after some time and the “pent-up” prices changes are unusually large. To delete them would bias the index downwards. The outlier detection is to alert the compiler about a possible error that needs further investigation.

Method of using z-score:

In a z-score test, the mean and standard deviation of the entire data set are used to obtain a z-score for each data point, according to following formula:

$$Z_i = \frac{(x_i - \bar{x})}{s}$$

Where

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

If the observations have a bell shaped distribution (standard normal distribution), the interval from $\bar{x} - s$ to $\bar{x} + s$ will contain approximately 68% of the measurements; the interval from $\bar{x} - 2s$ to $\bar{x} + 2s$ will contain approximately 95% of the measurements, and the interval from $\bar{x} - 3s$ to $\bar{x} + 3s$ will contain approximately all of the measurements.

In the case of price relatives, we do not know the underlying distribution of data set. Many studies of price change show that price relatives are not normally distributed. Thus we appeal to Chebyshev's theorem which applies to all possible distributions. According to Chebyshev's theorem, for any set of measurements and any number $k \geq 1$, the interval from $\bar{x} - s$ to $\bar{x} + s$ will contain at least (1 - 1/k2)*100 percent of the measurements.

Thus, at least 88.8 percent of all the observations in a data set will have z-score less than 3 in absolute value i.e. fall into the interval $(\bar{x} - 3s, \bar{x} + 3s)$, and at least 75 percent will fall within 2 standard deviations, where $\bar{x}$ is the mean and *s* is the standard deviation of the sample. Therefore, the observations with z-score greater than 3 will be potential outliers.

Example

The short-term price relatives measured by dividing current period prices by previous period prices are recorded in following table.

100	98	103	**121**	104	102
101	102	101	102	105	105
103	110	100	104	106	102
85	102	101	103	110	101
101	104	105	106	107	100

Sample Data for Short-Term Price Relatives

For this data set, $\bar{x}$ =103.1333, s= 5.4818, 3s=16.4454, z-score of the observation
of 121 is (121-103.1333)/5.4818=3.2593, z-score of 85 is (85-103.1333)/5.4818 = -3.3079.
Since the absolute values of z-score of 121 and 85 are more than 3, the price relatives 121 and the 85 are outliers in the data set.

The Z-score method is biased by the problem that both the mean and standard deviation are affected by the outliers.

Box Plots Method

Another procedure for detecting outliers is to construct box plots of the price relatives data. They make no distributional assumptions and, since they rely on the median and quartiles as parameters, the method of detection is not influenced by the outliers themselves. Below are the steps implemented in constructing the box plots for the software.

- The median M, lower and upper quartiles, QL and QU, and the inter-quartile range, IQR= QU - QL are calculated for the data set.

- Two sets of limits on the box plot are constructed: inner fences are located a distance of below QL and above QU; outer fences are located a distance of below QL and above QU.

Observations that fall between the inner and outer fences are called suspect outliers.

Locate the suspect outliers on the box plot using asterisks (*).Observations that fall outside the outer fences are called highly suspect outliers.

How the Quartile is Calculated?

Quartile calculation depends on the percentiles definition. The First quartile is the 25^{th} percentile (noted Q1), the Median value is the 50^{th} percentile (noted Median), and the Third quartile is the 75^{th} percentile (noted Q3)
The method to calculate the quartiles in CPI/PPI application is same as that used in Excel. It uses n-1 instead of n. the p-th percentile is defined by:
y = (1-g)* x(j+1) + g*x(j+2) where (n-1)*p=j+g (and x(0) is taken to be x(1)).
Let n be the number of observations in a data set (here n=4), and X(1)…X(n) the ordered values of a data set. Let p be the p-th percentile we want to calculate (e.g. p=0.25, 0.5, or 0.75). We'll calculate the product n*p. the product n*p can be split up between j and g, where j is the integer part of n*p and g is the decimal part of n*p.

Example

To better understand this method, we'll apply them on an simple example.
The data set studied is:

Variable	X1	X2	X3	X4

Value	2	1	4	3

Once ordered it becomes:

Variable	X1	X2	X3	X4
Value	1	2	3	4

In this example, for Q1, p=0.25, n=4, (n-1)*p=3*0.25=0.75*X(2) = 0.25*1+0.75*2 = 1.75, thus the 25th percentile is 1.75 with this method.

Log-normal Method

Another procedure that has been implemented into the software to identify the possible errors and outliers is to use 2s from the log-normal distribution, excluding price relatives of 100 (no change of prices from previous to current period).[8] It takes the natural logarithms of price relatives data, which is assumed log-normally distributed. The standard deviation and mean of the logged of all price relatives in the sample are calculated. Those price relatives that fall outside of 2 standard deviations (with 95% confidence level) are considered as possible outliers.

Calculating Adjusted Weights

If sampling outlets so that some are selected with certainty, say as a cut-off sample, and some are selected to be representative of the remaining outlets the weights of each outlet in the latter need to be adjusted. Further, outlets may disappear from the active sample and it may be necessary to redistribute the weight across the active sample. These two effects are picked up in an adjustment routine for the weights. The outlet adjusted weight is calculated based on the sample group to which the outlet is classified. For the sample group selected with certainty, an outlet adjusted weight is equal to the outlet assigned weight. They only represent themselves. However, if outlet(s) disappear from the sample, though still sell goods and services in reality, the weights of the remaining establishments can be adjusted so that those still active get allocated a *prorata* share of those that are “inactive” within its sample segment. If there were 5 selects initially and they each had a value weight of 20 and one disappears and 4 remain, each would get a weight of 25. An outlet adjusted value weight is equal to an outlet assigned value weight (20) divided by the total

[8] The reason to exclude those price relative with 100 value is that, presumably, there are a lot of “no change” price relatives. By including them we will observe a bimodal distribution. Since it is no longer normally distributed, we cannot apply a 2-sigma limit with 95% confidence to detect outliers.

value weight of outlets that are active in the product group (80), then times the total value weight of outlets (both active and inactive outlets) of the product group (100) i.e.

$$W_i^{adj} = \frac{W_i^{Assigned}}{\sum_i^n W_i^{active}} \times \sum_i W_i^{Total}$$

in which W_i^{adj} is adjusted weight of an outlet *i*; $W_i^{Assigned}$ is assigned weight of an outlet *i*, $\sum_{i=1}^{n} W_i^{active}$ is the total weight of active outlets and $\sum_{i=1}^{n} W_i^{Total}$ is the total weights of all the outlets. This is to say that weights of those inactive outlets will be taken and redistributed to the active outlet based on its share in total active weights.

In the probability selected group, the adjusted weight for an outlet being selected with probability to represent others is:

$\frac{\bar{W}^{Active}}{\sum_i W_i^{Active}} \times \sum_i W_i^{Total} = \frac{1}{n} \sum_i W_i^{Total}$ and the adjusted weight to represent outlets no longer active is:

$$W_i^{adj} = \frac{\frac{1}{n}\sum_i W_i^{Total}}{\sum_i W_i^{active}} \times \sum_i W_i^{Total}$$

Hierarchical Structure of Product Classification

The hierarchical levels of product classification used in CPI System are as follows, moving from the general to the more detailed level. The structure is that used for the 2003-2006 ICP round. The basis of this structure and its relationship to COICOP is outlined in the ICP Handbook, Chapter 5.[9] Examples of the six digit "basic headings" code are *Fresh milk* 11.01.14.1; *Shoes and other footwear* 11.03.21.1, *Bread* 11.01.11.3, and *Rice* 11.01.11.1. This is the level at which

[9] Available at: http://web.worldbank.org/WBSITE/EXTERNAL/DATASTATISTICS/ICPEXT/0,,contentMDK:20962711~menuPK:2666036~pagePK:60002244~piPK:62002388~theSitePK:270065,00.html. See also www.unece.org/stats/documents/ece/ces/ge.22/2006/mtg2/12.e.pdf

weights are generally available. An example of the structure is given below. Note that level 7 are “product clusters” and may be more detailed than level 6, for example *Bread* at level 6 may be defined as *White Bread* and *Bread Other Than White* at level 7. However, as in the example below for Rice, levels 6 and 7 may be the same. The main distinction between level 6 and 7 in such cases is that the latter contains the structure product descriptions (SPDs) – a listing of possible characteristics that can be used to precisely define the items. Level 8 is the result of including the precise characteristics of the individual products for which prices are to be collected in the SPDs. There may be more than one defined product specification at level 8. This assortment can be adjusted by the compiler to reflect changes in product supply and consumer behavior.

Weights should be entered at level 6 or a more detailed level where possible and be derived from actual data.

Level	COICOP-ICP Code	Description
0	0	All Products
1	1	GROSS DOMESTIC PRODUCT
2	11	INDIVIDUAL CONSUMPTION EXPENDITURE BY HOUSEHOLDS
3	11.01	FOOD AND NON-ALCOHOLIC BEVERAGES
4	11.01.1	FOOD
5	11.01.11	Bread and cereals
6	11.01.11.1	Rice
7	11.01.11.1_01	Rice/RICE
8	11.01.11.1_01a	Imported Uncle Ben’s long-grain rice 1lb boxed in supermarket
8	11.01.11.1_01c	Domestic brown medium-grain rice 1lb sold loose in open market

Chapter

4

Installation of the Software

The Basic Information

The CPI System was developed using Microsoft Visual Basic ® 6.0. Both cross-section data and time-series data are stored in Microsoft Access format. The system uses Microsoft ActiveX Data Objects (ADO 2.8) for the data access and database management that is included in Microsoft Data Access Components (MDAC). The system uses VideoSoft ActiveX Controls® VSFlexGrid® 7.0 and Formula One® for the tabulation, Wise InstallBuilder® 8.03 for software packaging and RoboHelp® for generating standard help file.

Package Contents

The installation package on the CPI distribution CD-Rom includes all necessary files and ActiveX components for the system to operate. The MDAC 2.8 setup file is also included in the installation package and will be installed into the user machine.

The package contains standard COICOP-ICP and COICOP product classification together with an SPD list.

System Requirements

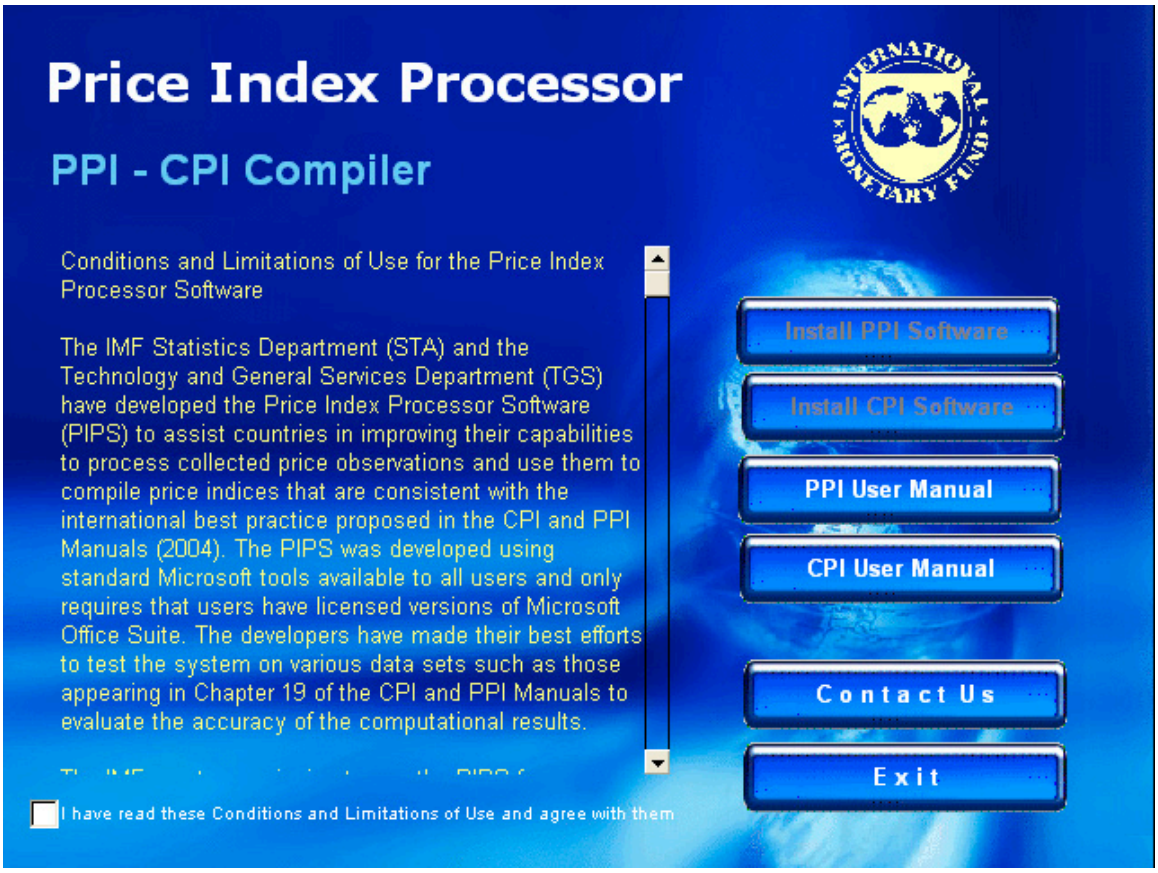
The system requirements are as follows:

a) Operating Systems: Windows 98, Windows NT 4.0, Windows 2000, Windows ME, or Windows XP.
b) Microsoft Office 97 or above.
c) CPU: Pentium-600 MHz or higher
d) 512 MB RAM or more
e) 60 MB free disk space
f) VGA–True Color video mode, displays at 800 x 600 or 1024 x 768 pixels.

The installation package on the CPI distribution CD-Rom includes all necessary files, including ActiveX components for the system to operate and MDAC 2.8 setup file.

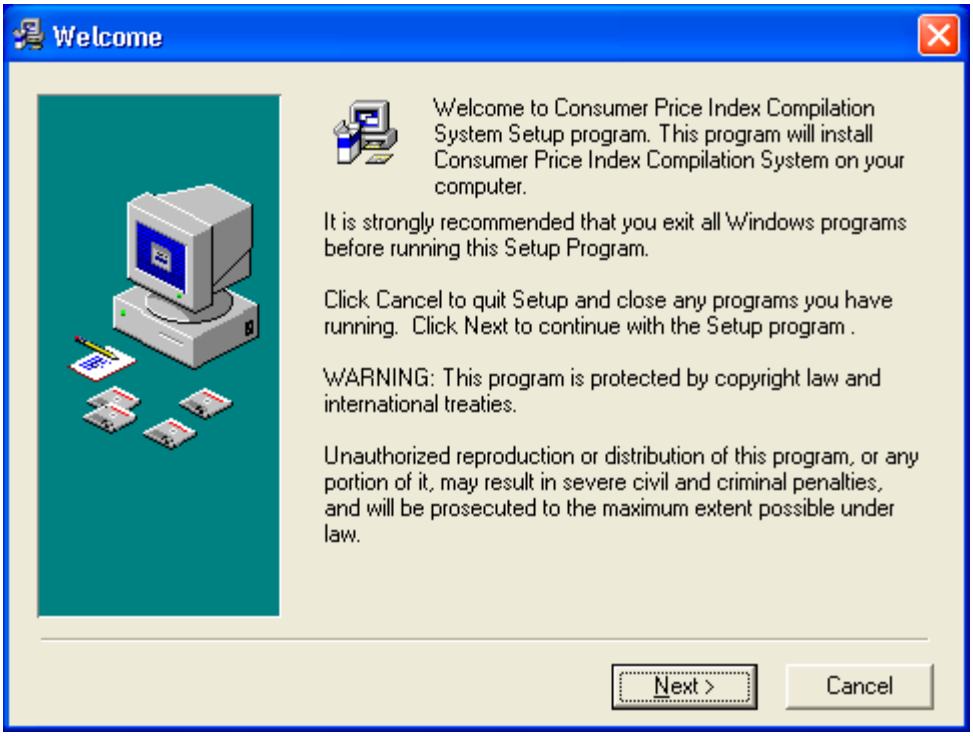
Installation Procedures

Insert the Price Index Processor CD-ROM in the CD drive, The user will see following screen:

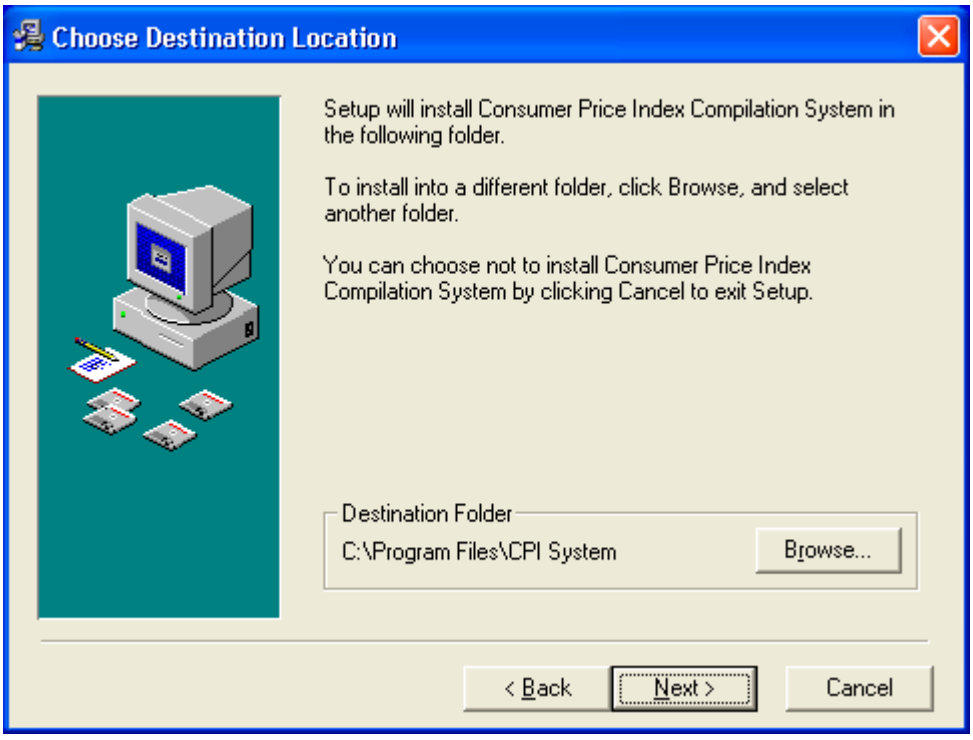


If the user agrees with disclaimer agreement, he should click **I have read these Conditions and Limitations of Use and agree with them** checkbox to activate **Install PPI Software** and **Install CPI Software** buttons.

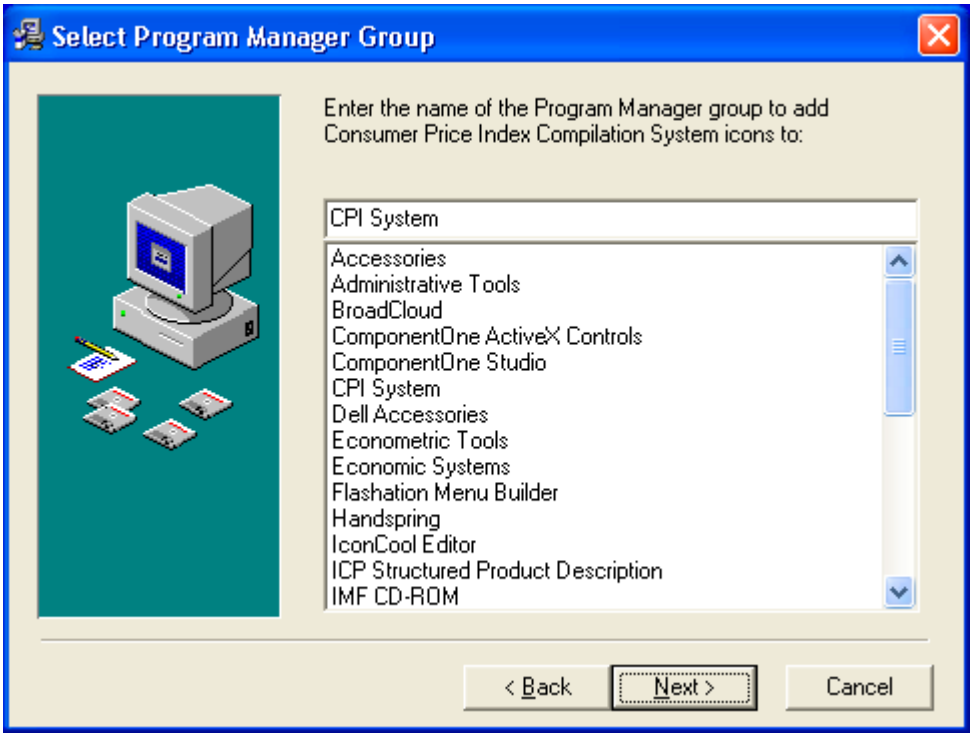
Click either **Install CPI Software** or **Install PPI Software** button to trigger installation process. Click **CPI User Manual** or **PPI User Manual** to read the manual in PDF format.



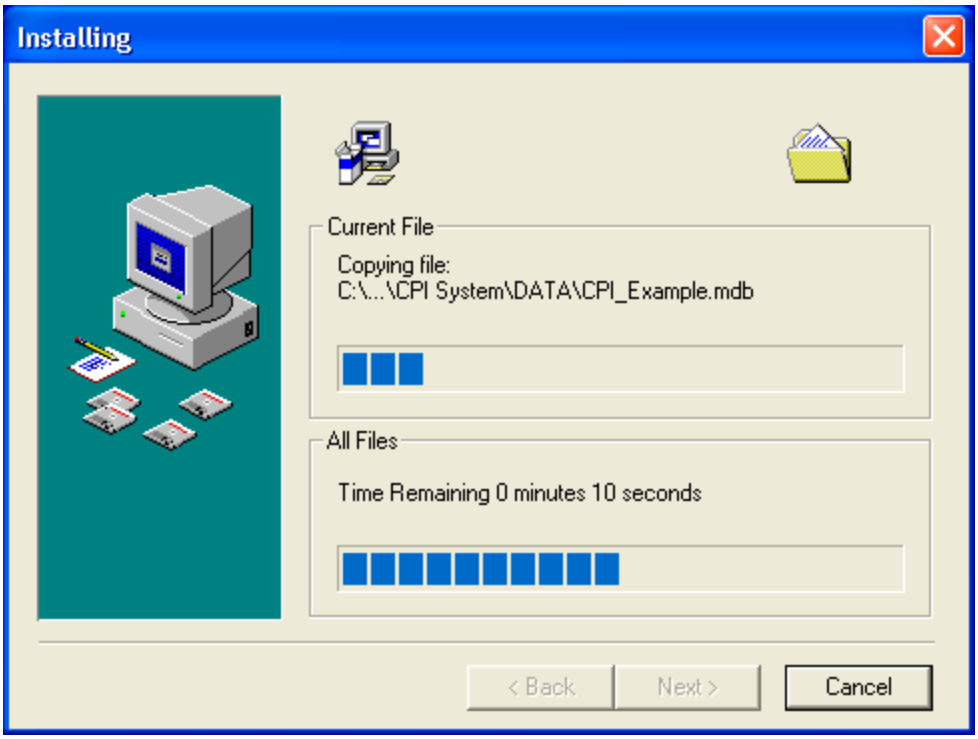
Click **Next**



Click **Next.**


Select Program Manager Group
Enter the name of the Program Manager group to add Consumer Price Index Compilation System icons to:
CPI System
Accessories
Administrative Tools
BroadCloud
ComponentOne ActiveX Controls
ComponentOne Studio
CPI System
Dell Accessories
Econometric Tools
Economic Systems
Flashation Menu Builder
Handspring
IconCool Editor
ICP Structured Product Description
IMF CD-ROM
< Back
Next >
Cancel

Click **Next.**


Installing
Current File
Copying file:
C:\...\CPI System\DATA\CPI_Example.mdb
All Files
Time Remaining 0 minutes 10 seconds
< Back
Next >
Cancel

Wait until installation is complete.

Click Cancel to exit.

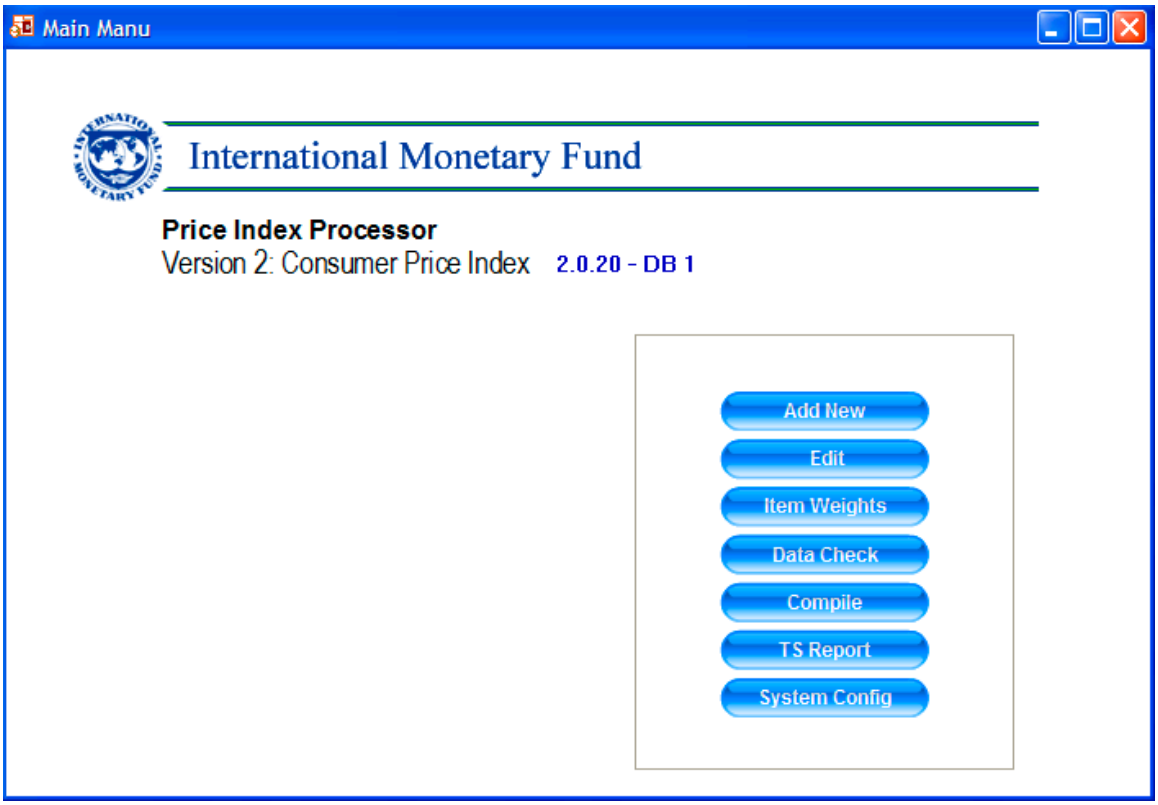
Click the **Finish** button to complete the installation process.
The Installation procedure will create CPI System shortcut at Desktop and create icon at Program Manager Group.

Chapter

5

The Basics

Click **Start**, **All Programs**, **CPI System,** then click  icon or click CPI System icon from the desktop to start CPI program.



"2.0.20" is the application version number and "DB1" is the database version.

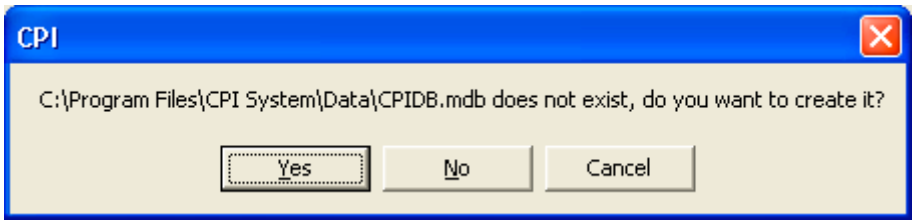
System Configuration

The user can configure the CPI System work environment by clicking 

button in the main screen or clicking Options, Configure in toolbar. The user can define following settings:

1) select the default database location;

2) select default database name (the user can disconnect to a CPI database and re-connect to another CPI database. If the database the user selected does not exist, the user will be prompted to following message:



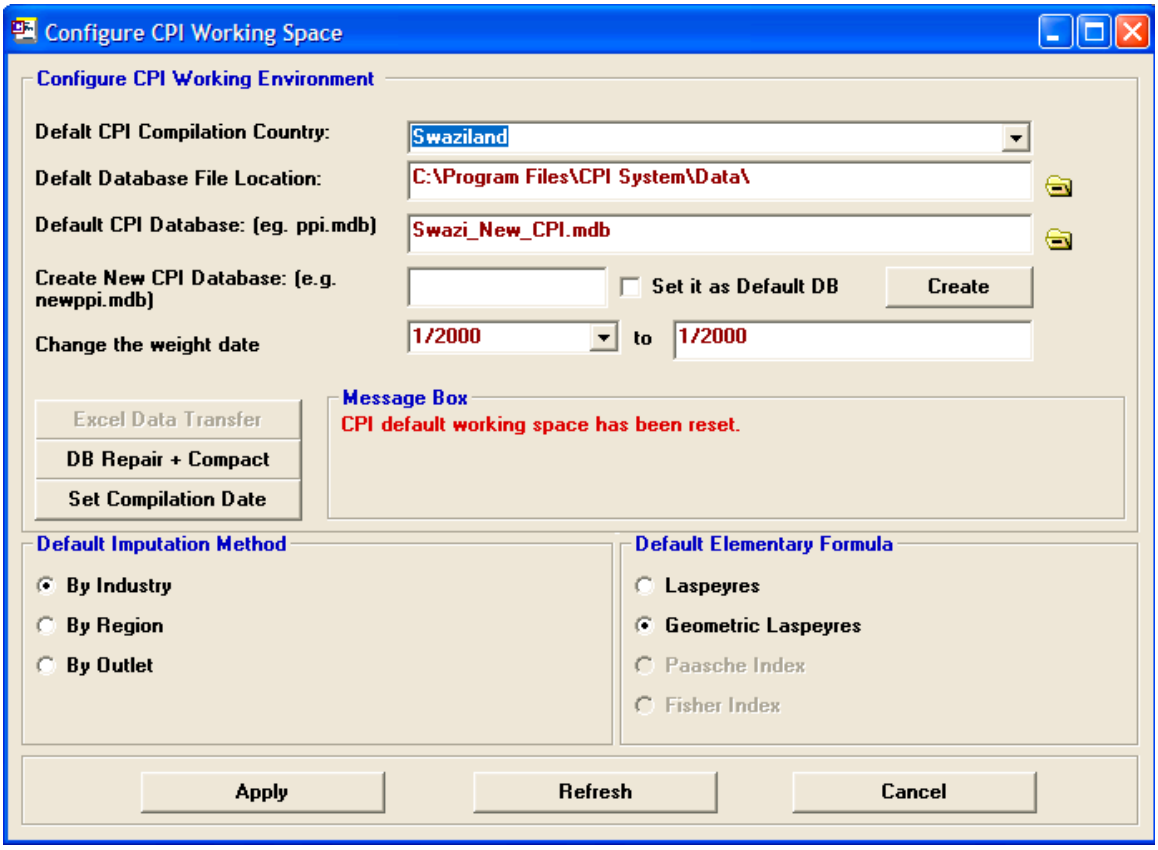
If the user clicks Yes, the system will create a new database.

3) create a new CPI database;

4) define default imputation method.

5) define default elementary formula

Click **Apply** after changes have been made. Click Refresh button to view recent changes that have been made.



Click **DB Repair + Compact** button to repair possible database error and compact Access database. Click **Set Compilation Date** to specify the compilation date.

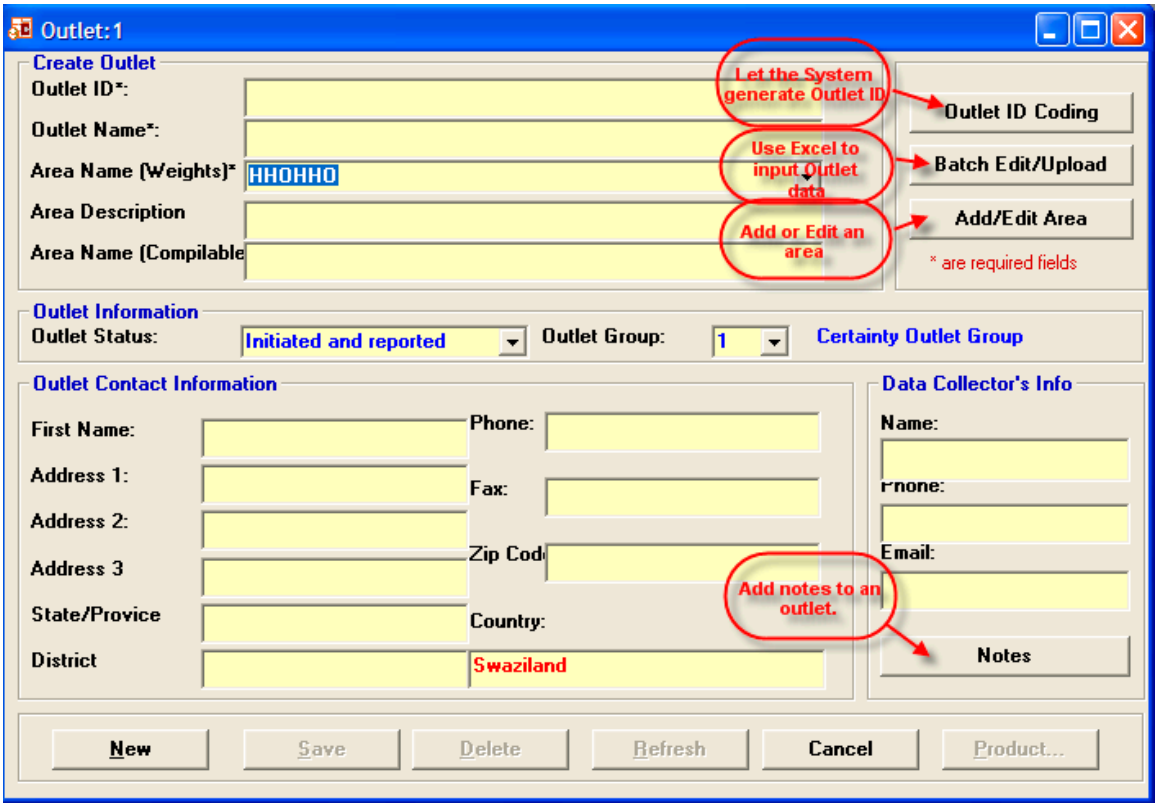
Add New Outlet

Click the 

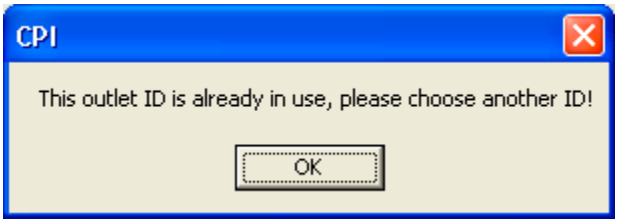
button to add a new outlet. Enter the outlet information in each field. Field with * indicates required filed. Click 

to select an area. If no area has been defined, 

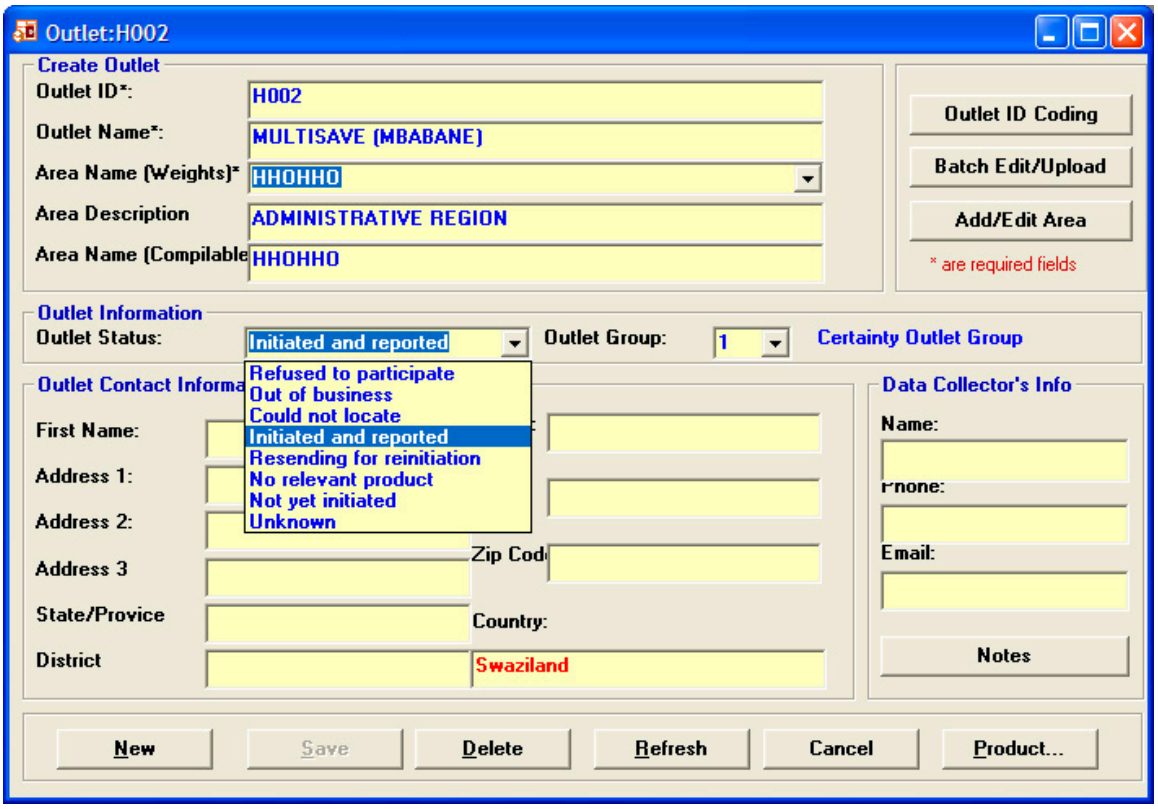
will be displayed here.



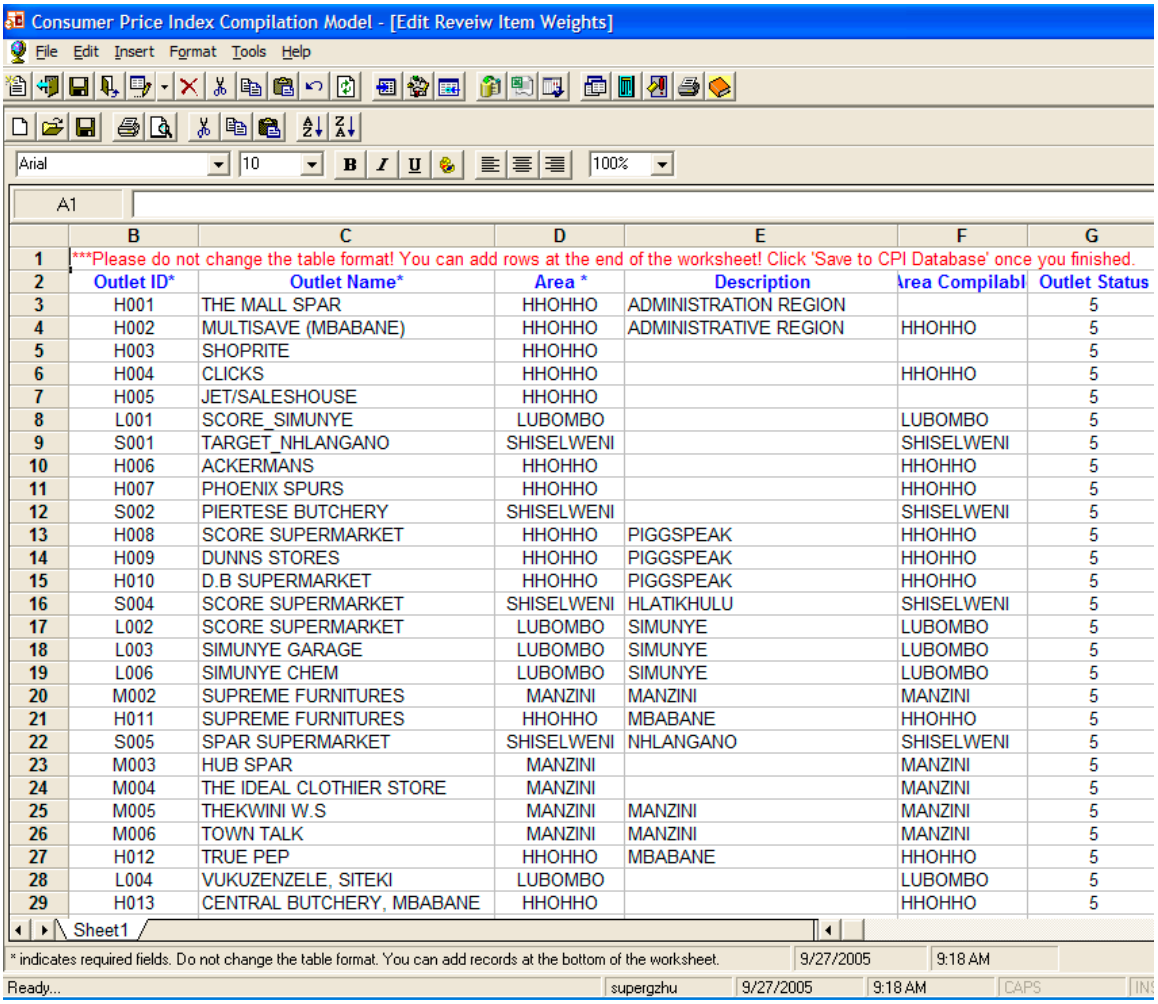
Outlet ID is unique field. If the outlet ID is already taken, the following message will be displayed:

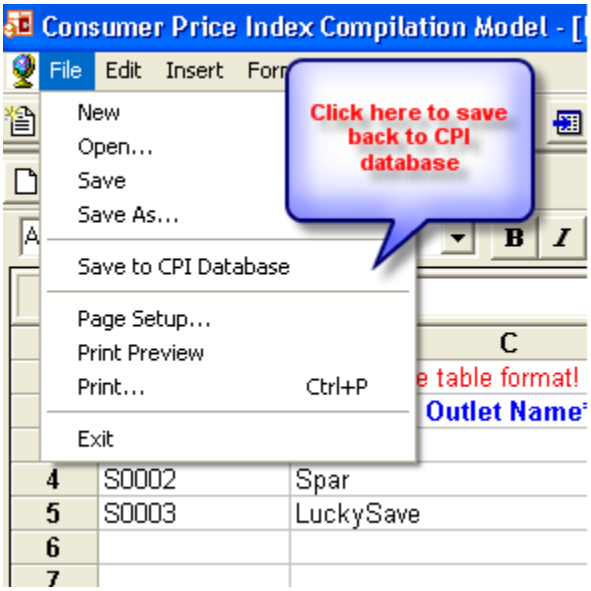


Click **Outlet Status** dropdown list to define the outlet status (default value is "Initiated and Reported". Click Outlet Group to define the group that outlet belongs to (default value is "Group I - Certainty Group". Outlet status will affect the distribution of expenditure weights and calculation of indices. An outlet that is defined as "Out of Business" or "Refuse to Participate" will be suppressed from calculation even though they are physically exist in the database.

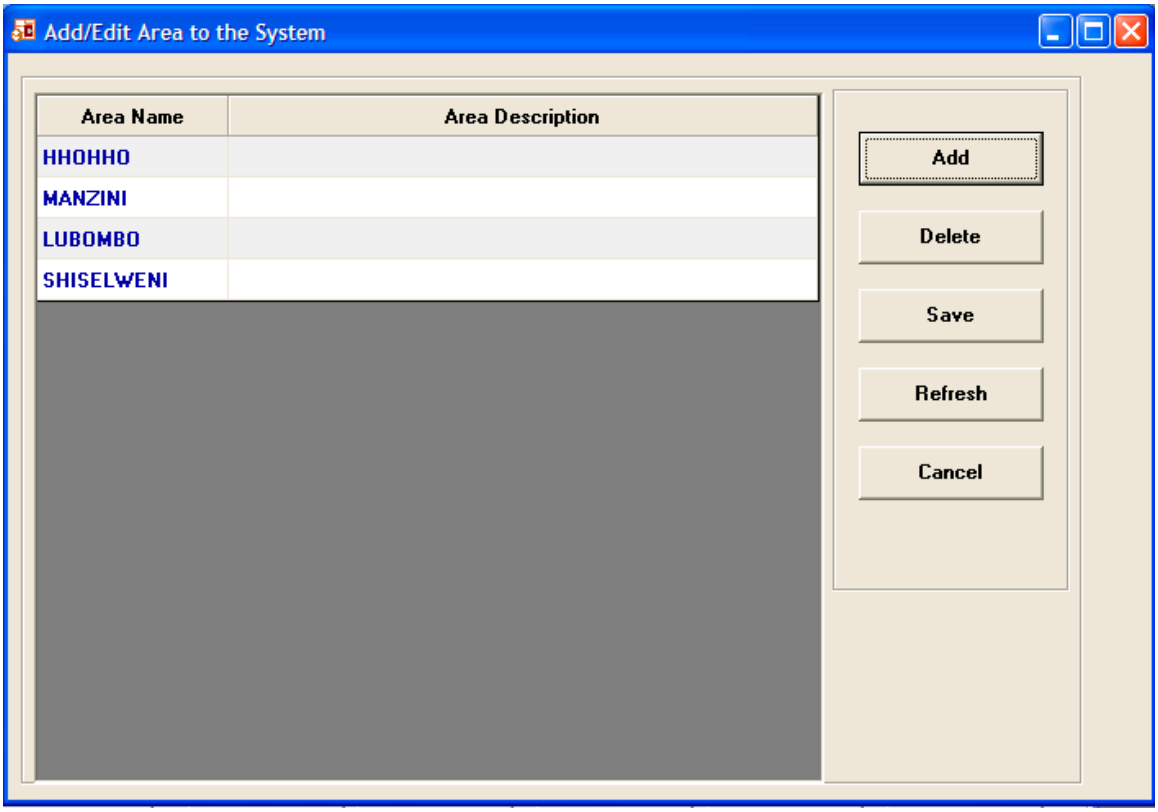


Click  button to input or edit multiple outlets at a time.


Consumer Price Index Compilation Model - [Edit Reveiw Item Weights]
File Edit Insert Format Tools Help
Arial 10 100%
A1
***Please do not change the table format! You can add rows at the end of the worksheet! Click 'Save to CPI Database' once you finished.
Outlet ID* Outlet Name* Area * Description Area Compilabl Outlet Status
H001 THE MALL SPAR HHOHHO ADMINISTRATION REGION 5
H002 MULTISAVE (MBABANE) HHOHHO ADMINISTRATIVE REGION HHOHHO 5
H003 SHOPRITE HHOHHO 5
H004 CLICKS HHOHHO HHOHHO 5
H005 JET/SALESHOUSE HHOHHO 5
L001 SCORE_SIMUNYE LUBOMBO LUBOMBO 5
S001 TARGET_NHLANGANO SHISELWENI SHISELWENI 5
H006 ACKERMANS HHOHHO HHOHHO 5
H007 PHOENIX SPURS HHOHHO HHOHHO 5
S002 PIERTESE BUTCHERY SHISELWENI SHISELWENI 5
H008 SCORE SUPERMARKET HHOHHO PIGGSPEAK HHOHHO 5
H009 DUNNS STORES HHOHHO PIGGSPEAK HHOHHO 5
H010 D.B SUPERMARKET HHOHHO PIGGSPEAK HHOHHO 5
S004 SCORE SUPERMARKET SHISELWENI HLATIKHULU SHISELWENI 5
L002 SCORE SUPERMARKET LUBOMBO SIMUNYE LUBOMBO 5
L003 SIMUNYE GARAGE LUBOMBO SIMUNYE LUBOMBO 5
L006 SIMUNYE CHEM LUBOMBO SIMUNYE LUBOMBO 5
M002 SUPREME FURNITURES MANZINI MANZINI MANZINI 5
H011 SUPREME FURNITURES HHOHHO MBABANE HHOHHO 5
S005 SPAR SUPERMARKET SHISELWENI NHLANGANO SHISELWENI 5
M003 HUB SPAR MANZINI MANZINI 5
M004 THE IDEAL CLOTHIER STORE MANZINI MANZINI 5
M005 THEKWINI W.S MANZINI MANZINI MANZINI 5
M006 TOWN TALK MANZINI MANZINI MANZINI 5
H012 TRUE PEP HHOHHO MBABANE HHOHHO 5
L004 VUKUZENZELE, SITEKI LUBOMBO LUBOMBO 5
H013 CENTRAL BUTCHERY, MBABANE HHOHHO HHOHHO 5
Sheet1
* indicates required fields. Do not change the table format. You can add records at the bottom of the worksheet.
9/27/2005 9:18 AM
Ready... supergzhu 9/27/2005 9:18 AM CAPS


Consumer Price Index Compilation Model - [
File Edit Insert Form
New
Open...
Save
Save As...
Save to CPI Database
Page Setup...
Print Preview
Print... Ctrl+P
Exit
Click here to save back to CPI database
C
e table format!
Outlet Name
4 S0002 Spar
5 S0003 LuckySave
6
7

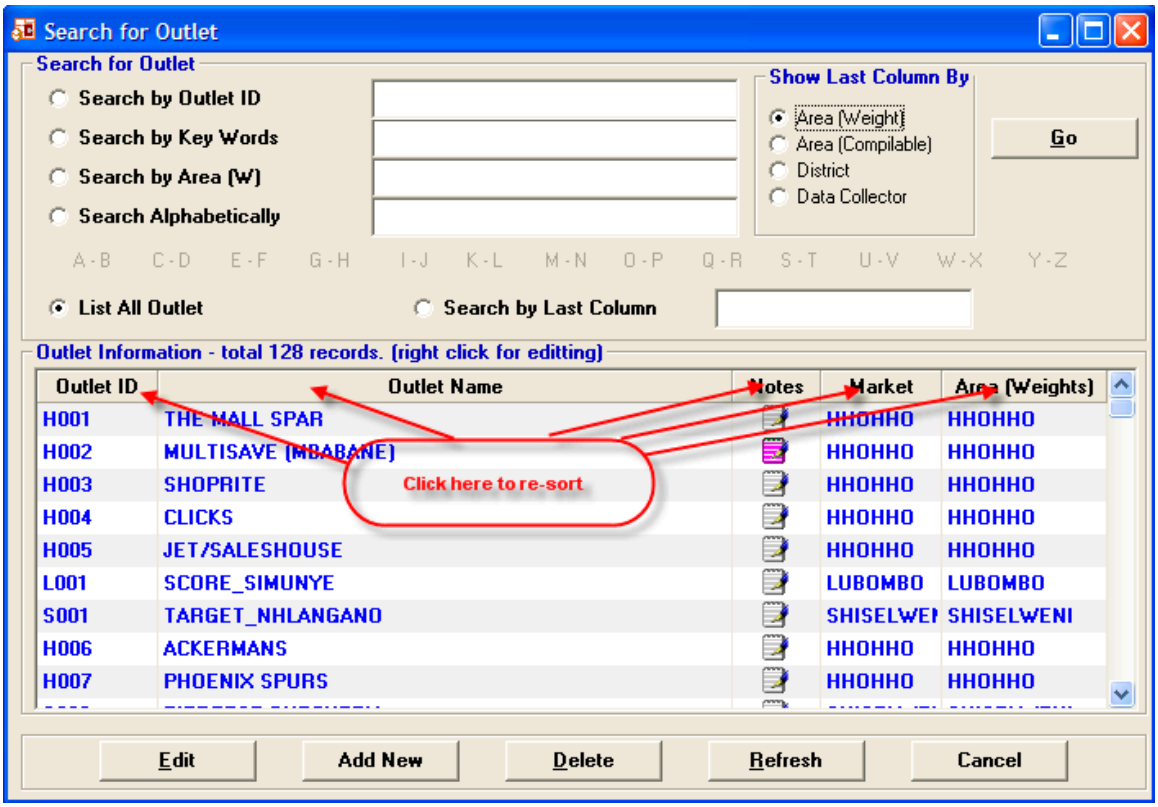
Click  button to add or edit area.



The user can also manage the area list here. It is important to note that the area defined here corresponds to the geographical area where household expenditure survey is conducted and where weights information is available. *If regional weights are not available, national weights by product group will be distributed equally across the regions that have been defined and by variety.*

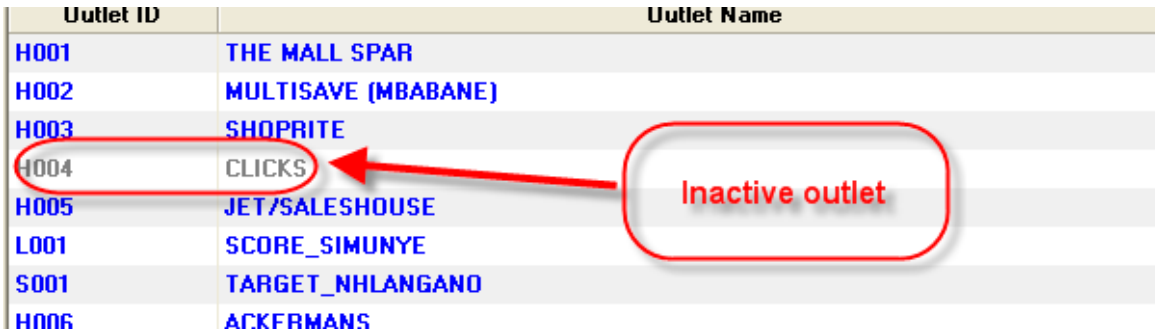
Edit Outlet, Product or Variety

Click  button to edit outlet information.

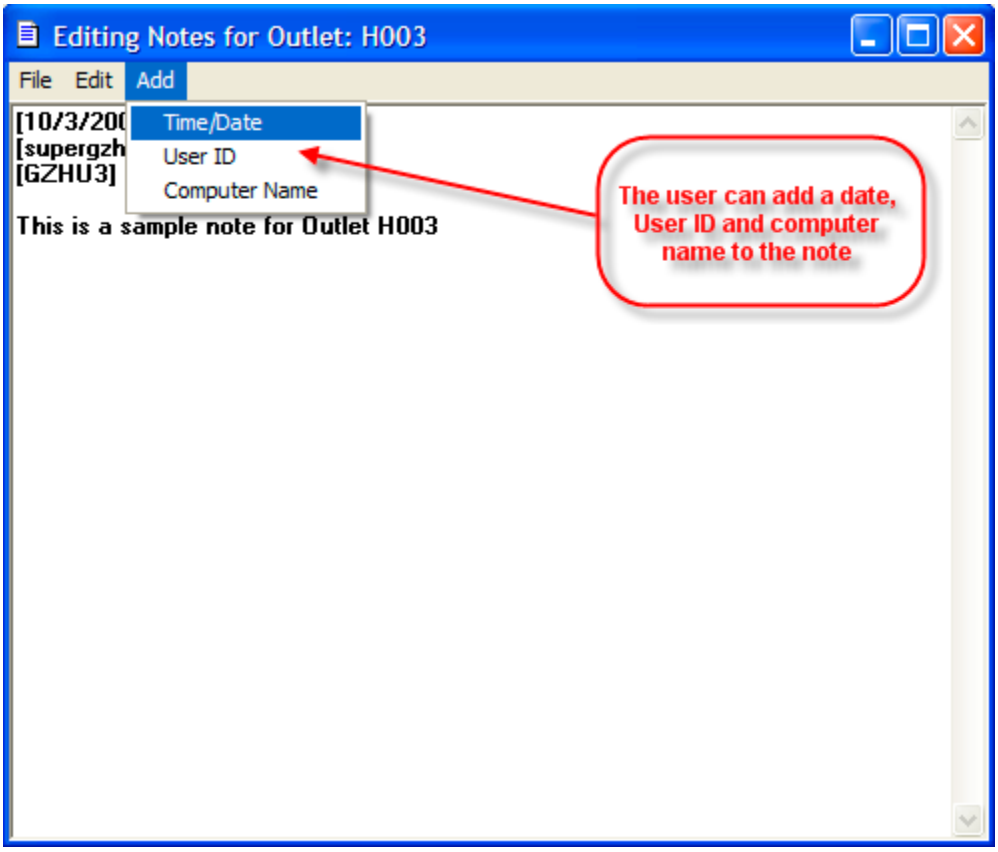


Click the first row of the respective column to re-sort the information in ascending or descending alphabetic order. Select a specific outlet (by clicking once on the particular row), right click (once) to invoke an edit Manu. It is important to note that you can change an outlet ID provided that ID is not taken. It the user wants to create a new outlet using an D that already exist, she can either delete that ID first, or hit the save button three times. In that way, the existing outlet information in the database will be overwritten by the information newly inputted. The user can edit outlet, product or price information by clicking the appropriate button.

Those inactive outlets (including those of out of business, not yet initiated or refuse to participate) will be showed in grey.

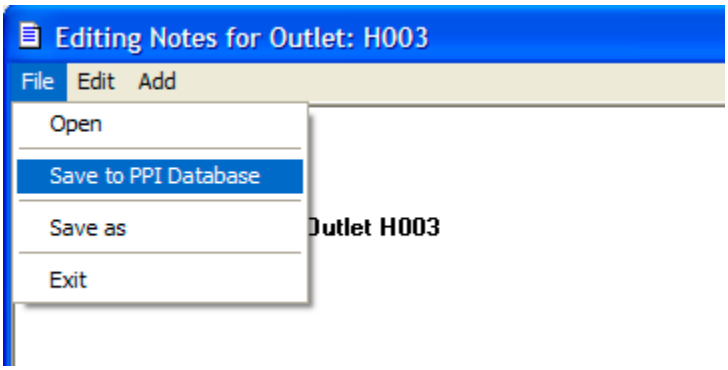


The user can add notes to an outlet by double clicking icon in Notes column. An icon indicates that there is a note associated with this outlet.

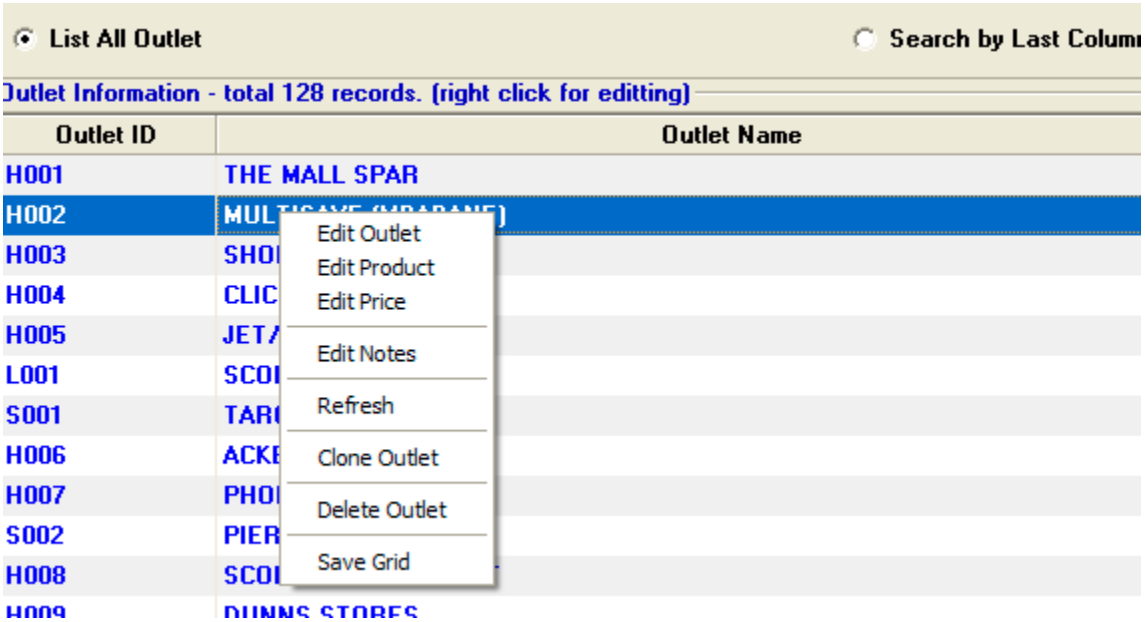


It is important to note that meta data text length should not exceed 65353 characters (a memo datatype in the database)

After adding a note, the user can save it to the CPI database or save it externally for the future reference.



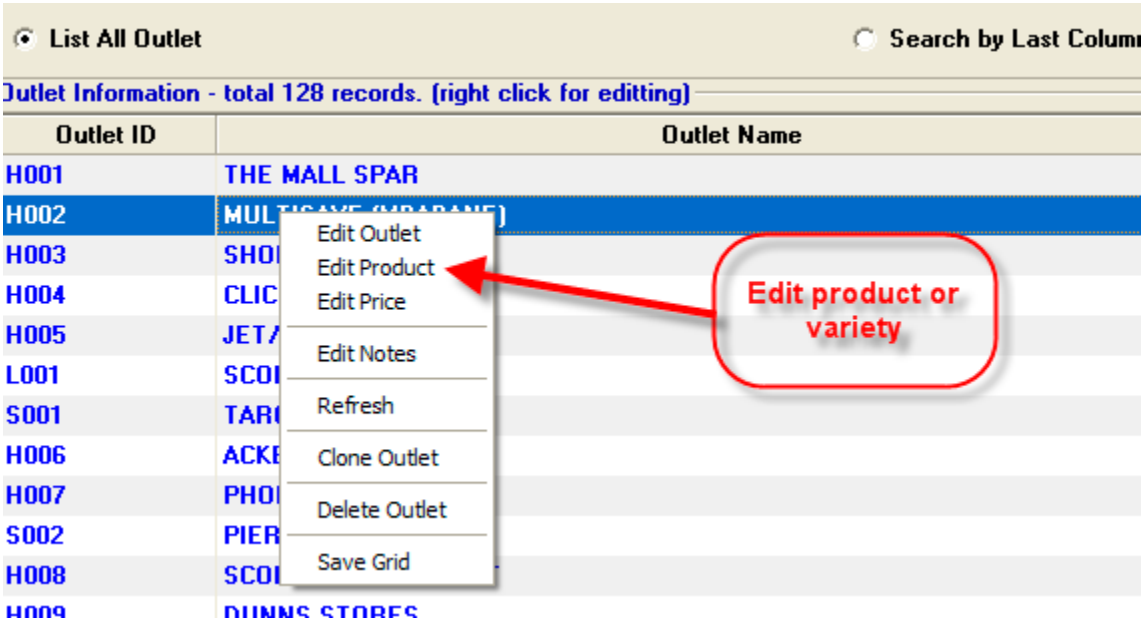
The user has the choice of editing outlet, editing a product, editing a price, editing notes, cloning an outlet, deleting an outlet or saving the contents of the grid to a spreadsheet.



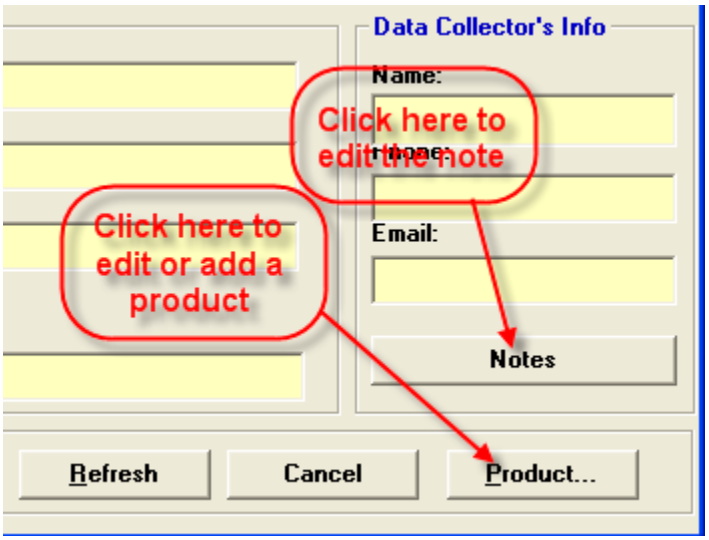
Cloning an outlet is to duplicate the complete outlet information, including the product structure. This is to provide the user with a quick way to create an outlet record without typing every piece of information all over again. This is also useful in the sample replacement exercise, where newly selected outlets replace the existing one.

Save Grid feature allows the user to save grid info into a spreadsheet.

Add New Product and Variety



To add product or variety information, the user can select an outlet (single click), Click **Edit Product** button, or click **Product** button on the outlet edit screen.

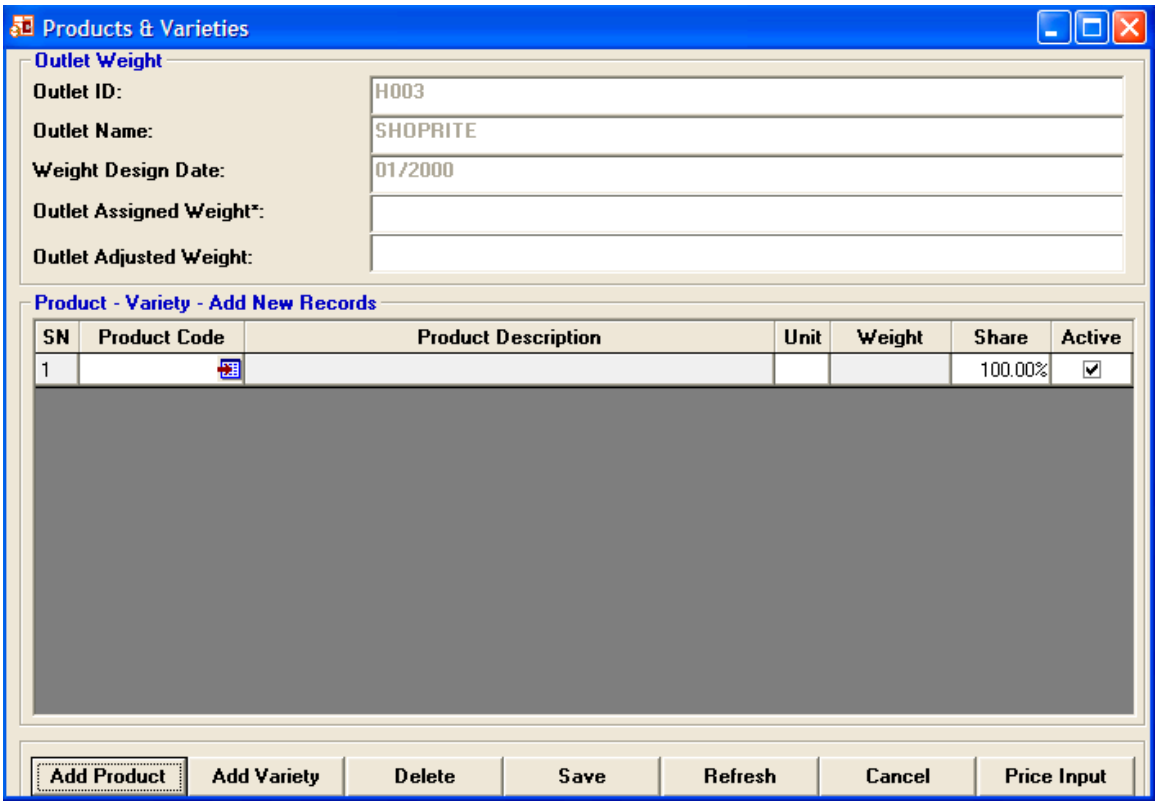


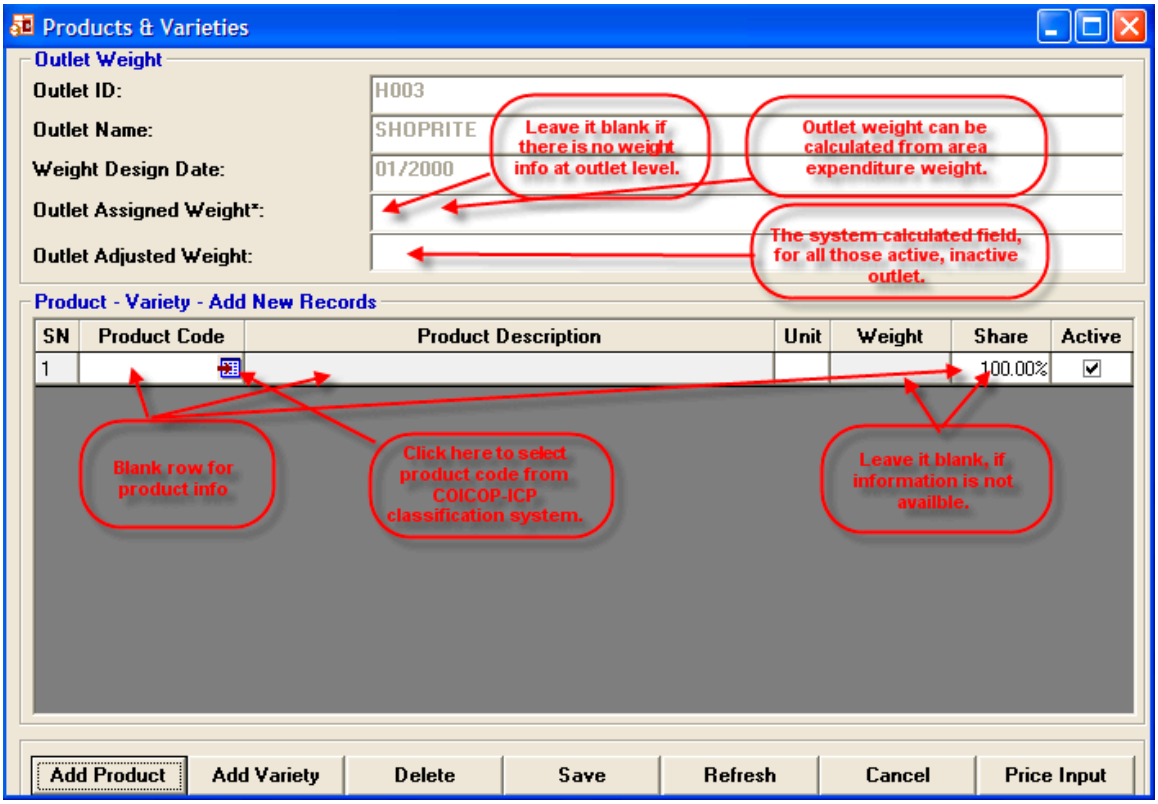
Click **Add Product** button to add a product. A blank row will be added for the user to input product information. If the **Outlet Assigned Weight** is not available for that outlet, the user can leave it blank. It will be calculated once area

expenditure weights for the product become available. Outlet Adjusted Weight is an imputed value based on the Outlet Assigned Weight, Outlet Status and Sample Group it belongs to.

Weight and Share fields for a product will be calculated once the Outlet Adjusted Weight becomes available.

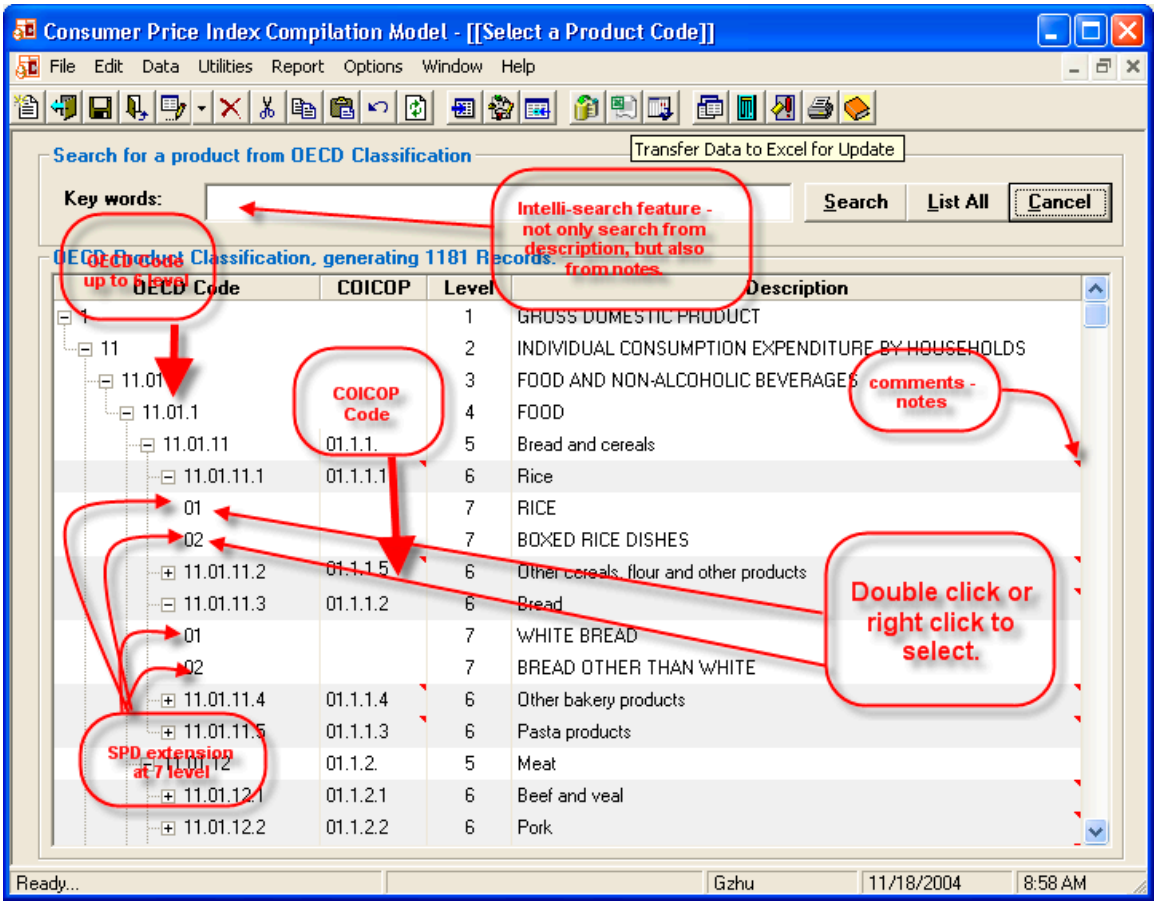
The following screen will allow the user to input product and variety information: Leave the unknown field blank, define the product structure by selecting from COICOP-ICP classification system.



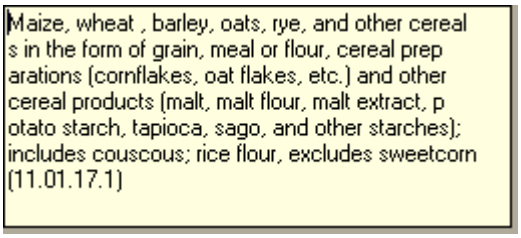


Click button to select a product from a pre-defined product classification system. In CPI System the default product classification is COICOP-ICP system with SPD extension.

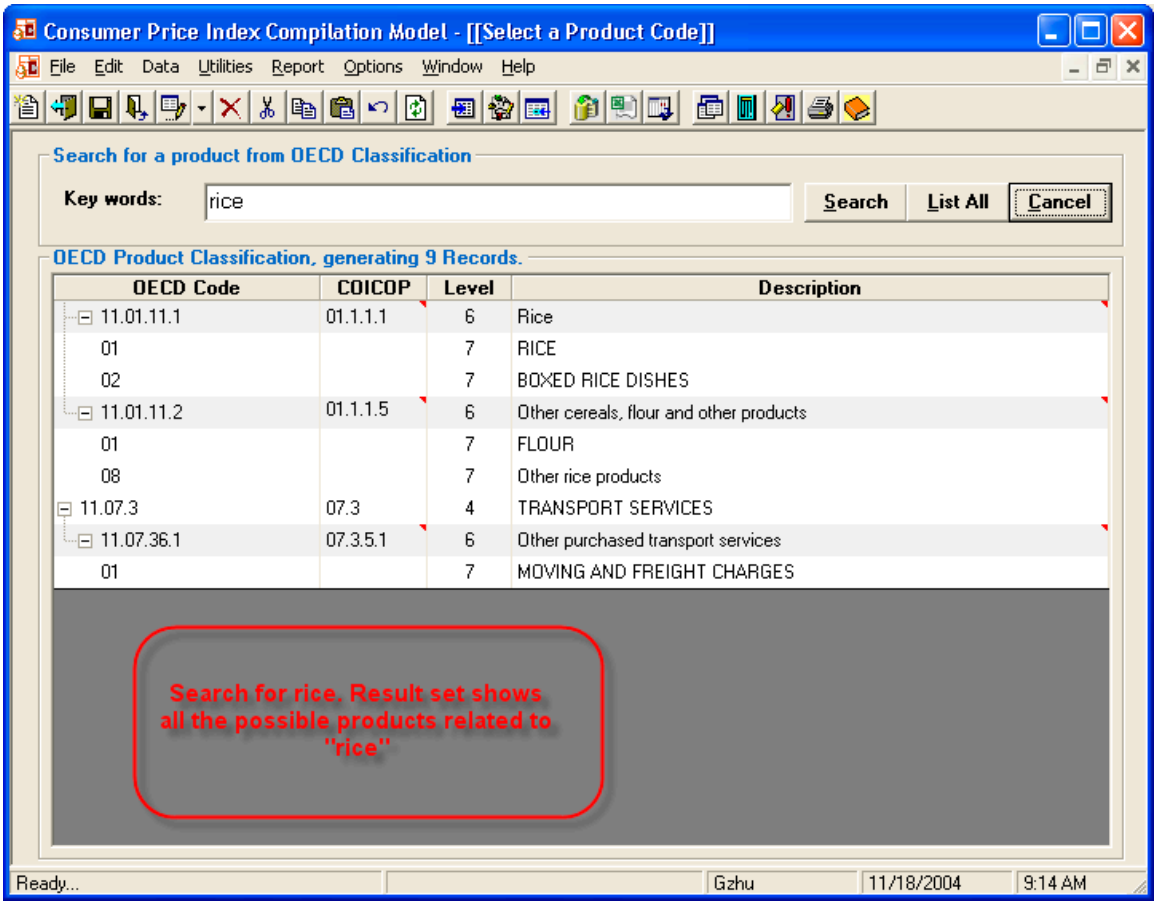
The user should double click to select a product from the most detailed level (in this case, at 7th level, which is also SPD level) from the product list. A search feature is provided for the user to quickly locate the product she wants to find. The user can type the complete word or part of that word; the search will be performed not only on the description but also on notes field.



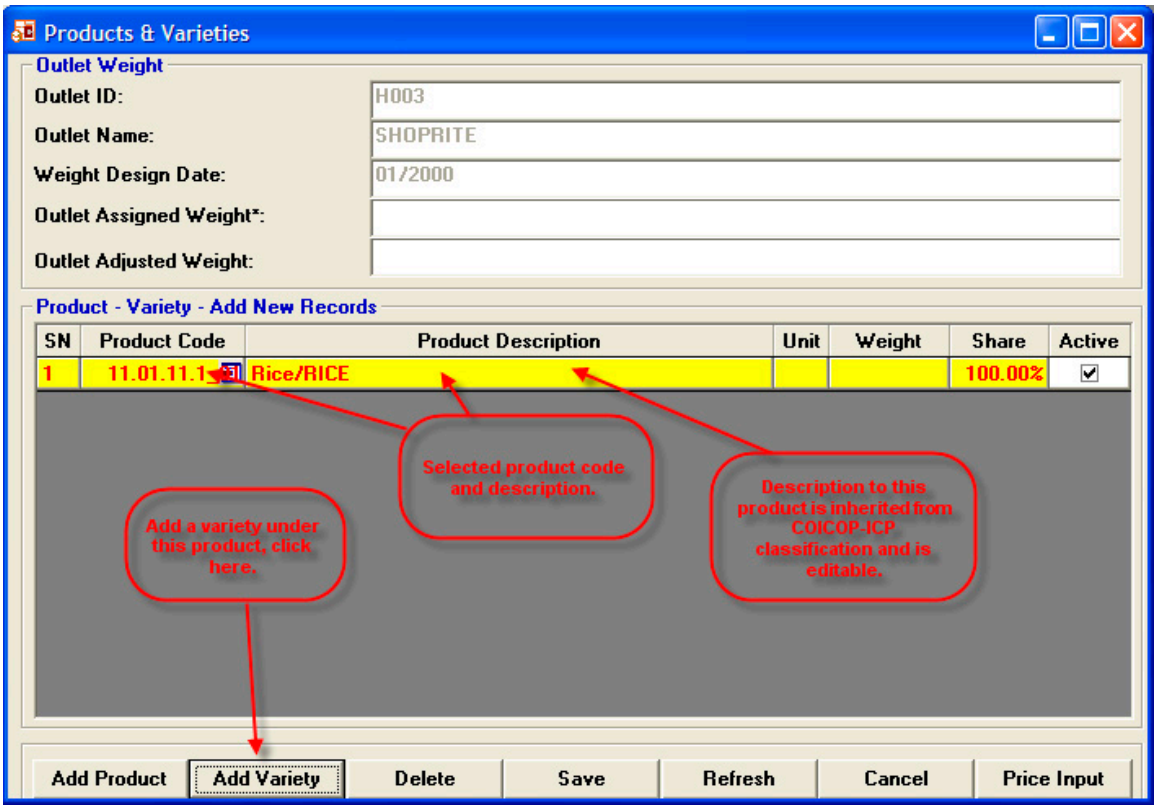
Additional information is provided in the comment field which was indicated by a red triangle at up-right of the cell. Following comment is associated with product “11.01.11.2, Other cereals flours and other product”.



The CPI System provides the user with an intelligent search facility. Let us assume the user wants to search for rice-related products, she entered “rice” in search textbox, click the **Search** button, and the following screen will be shown.

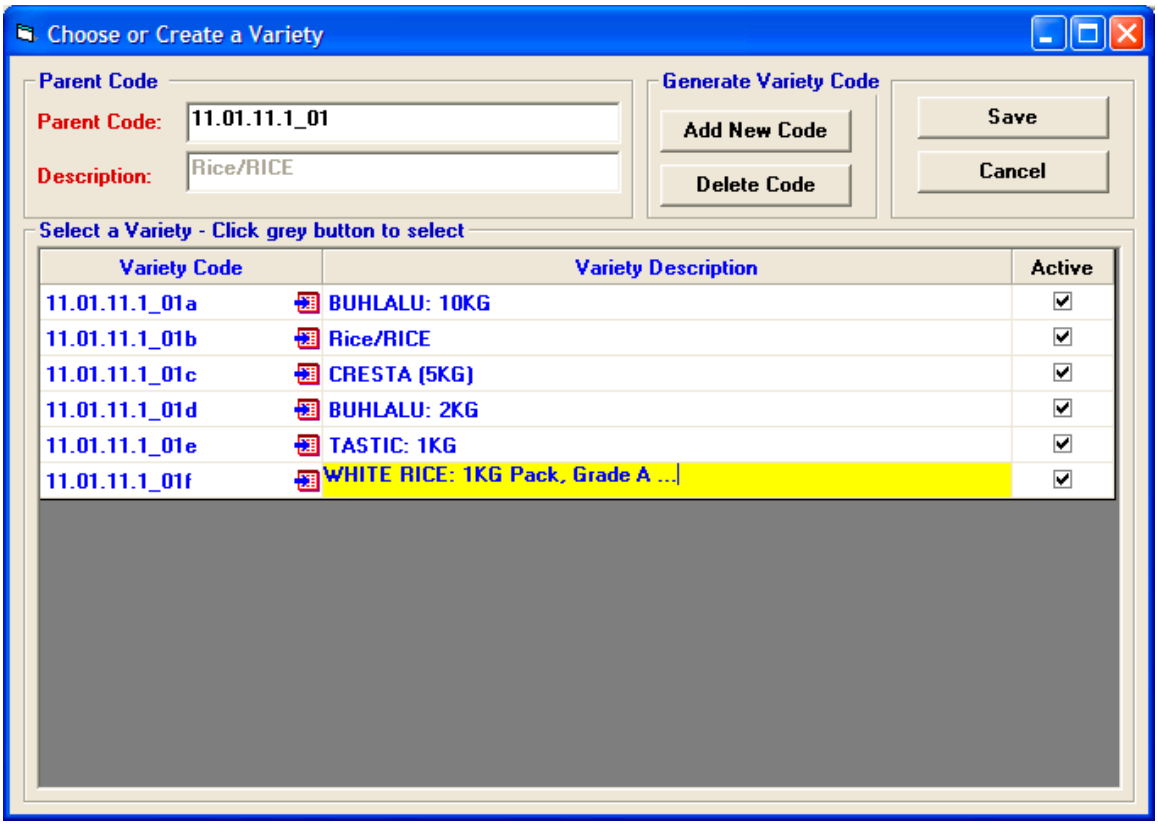


By double clicking the product the user wants to select, the product code with description will be extracted and displayed on the product selection screen.



The user then can add a variety to this product by clicking **Add Variety**  button.

The variety list is a user-defined list which extends the international standard product classification defined by COICOP-ICP. The Variety list comprises many variety items. Each variety item can have an user-definable, unique code and description with a SPD checklist type of details in the Description field, such as package type, serving, number of units in package, size of unit, unit of measure, origin, seasonal availability, product characteristics etc. The variety code is typically concatenated by its parent product code and additional one or two alphabetical letter. The system will generate such code when the user press the **Add New Code** button and the user can change this code if he wants to. The variety list allows each country to maintain additional dimension of product/variety information which reflects that its own specific consumption pattern. Each variety list has a parent, which is listed in the standard COICOP-ICP/SPD product classification. **In this way, the system maintains a international standard while provides each country with flexibility of creating additional dimension of variety with its unique characteristics.**

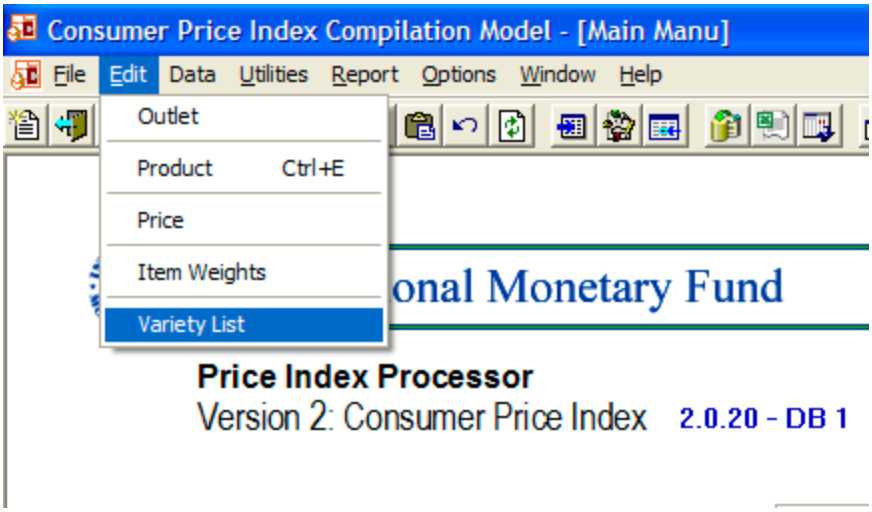


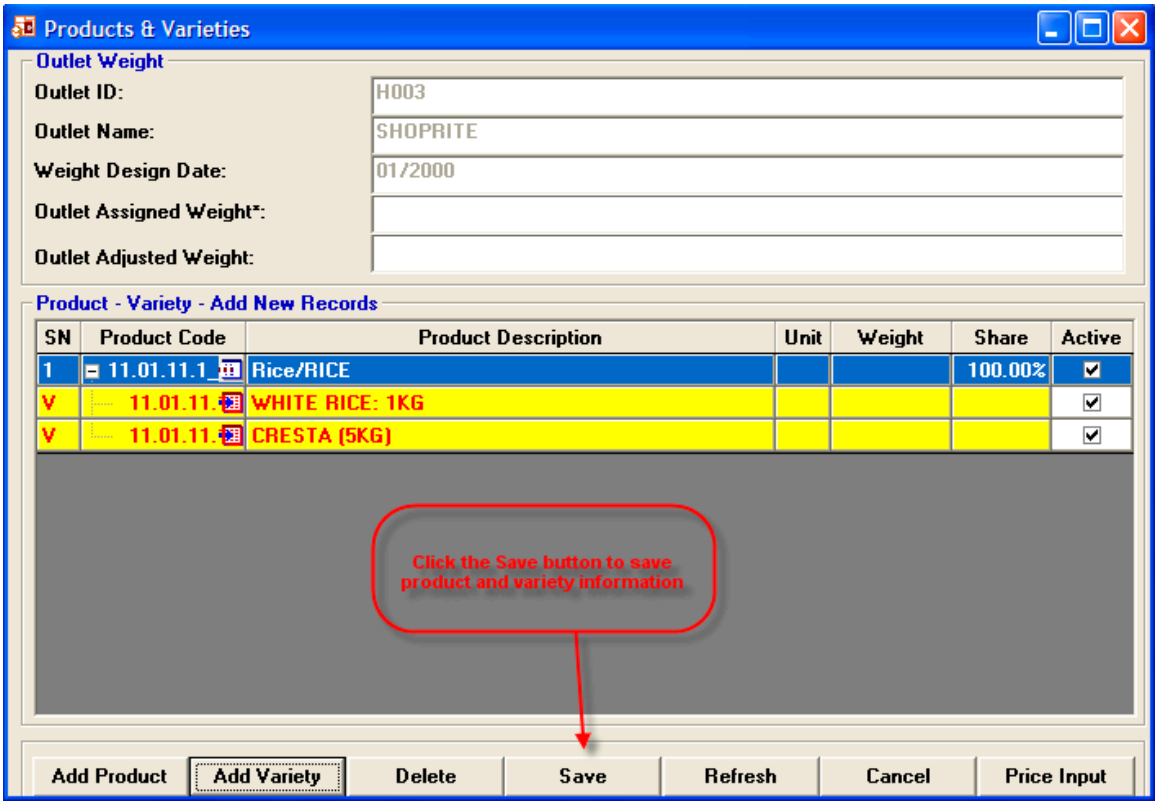
It is important to note that the variety list is shared by multiple outlets, if one particular variety is already used by another outlet, it can not be deleted from database.

The user should look at the variety list carefully to decide whether an existing variety code should be used or a new variety should be added and defined. Click the small grey button at right side of the first column to select a variety. Once a variety is selected, the “Choose or Add Variety” screen will disappear and variety information will show in the product screen. The user has to click **Save** button to store the product and variety information.

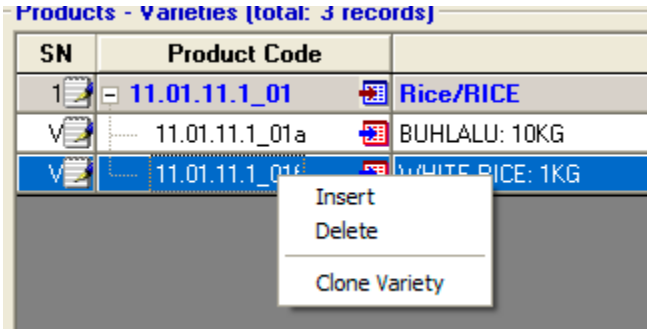
Please note, if there are multiple products in one outlet, the user has to select a product to which she wants add variety. The user can not add a variety directly under another variety.

Variety lists can be managed by clicking **Edit**, **Variety List** button as showed in following screen.





Select a Product Code (move the mouse to Product Code column, then single click) right click. The user can Insert, Delete or Clone Variety. Click icon, the user can add a note to the product or variety. Again, indicates that there is already a note for this product or variety.

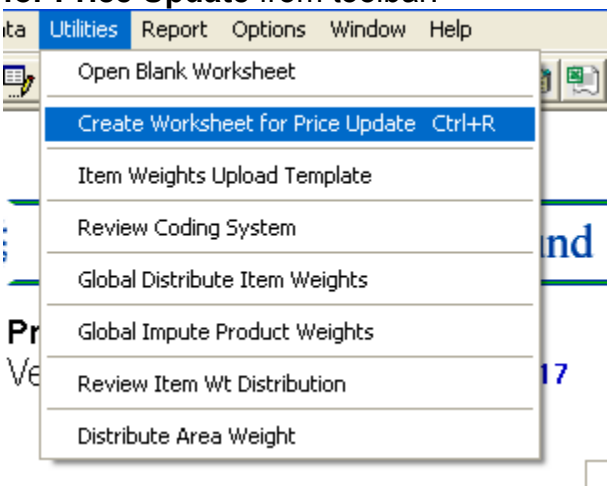


Input Price Quotation

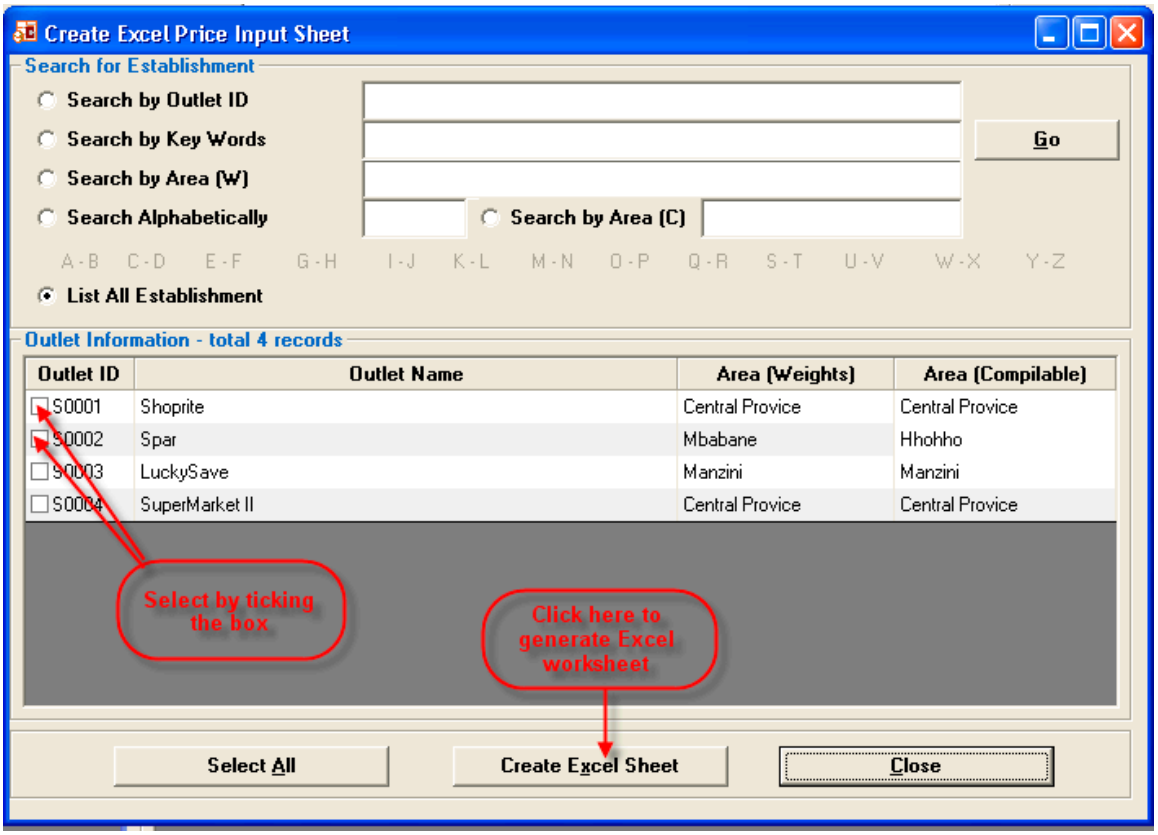
After defining products and varieties structure, the user is ready to input price quotations. The user has several methods to input price:

Method 1, Batch Price Data Input

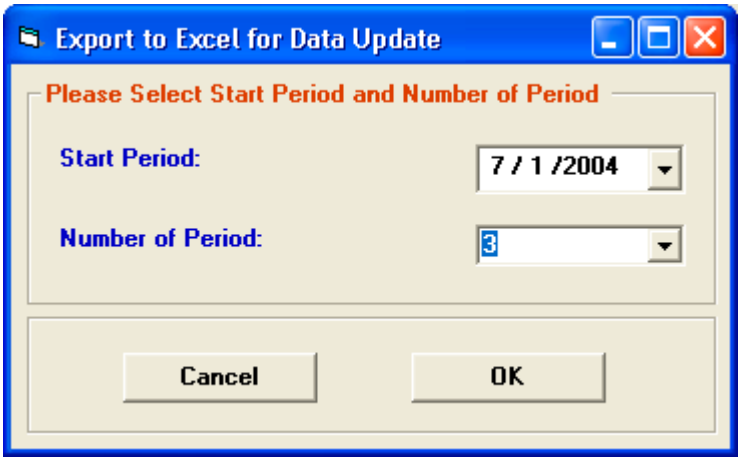
Click **Input Sheet** button at product screen or click **Utilities**, **Create Worksheet for Price Update** from toolbar.



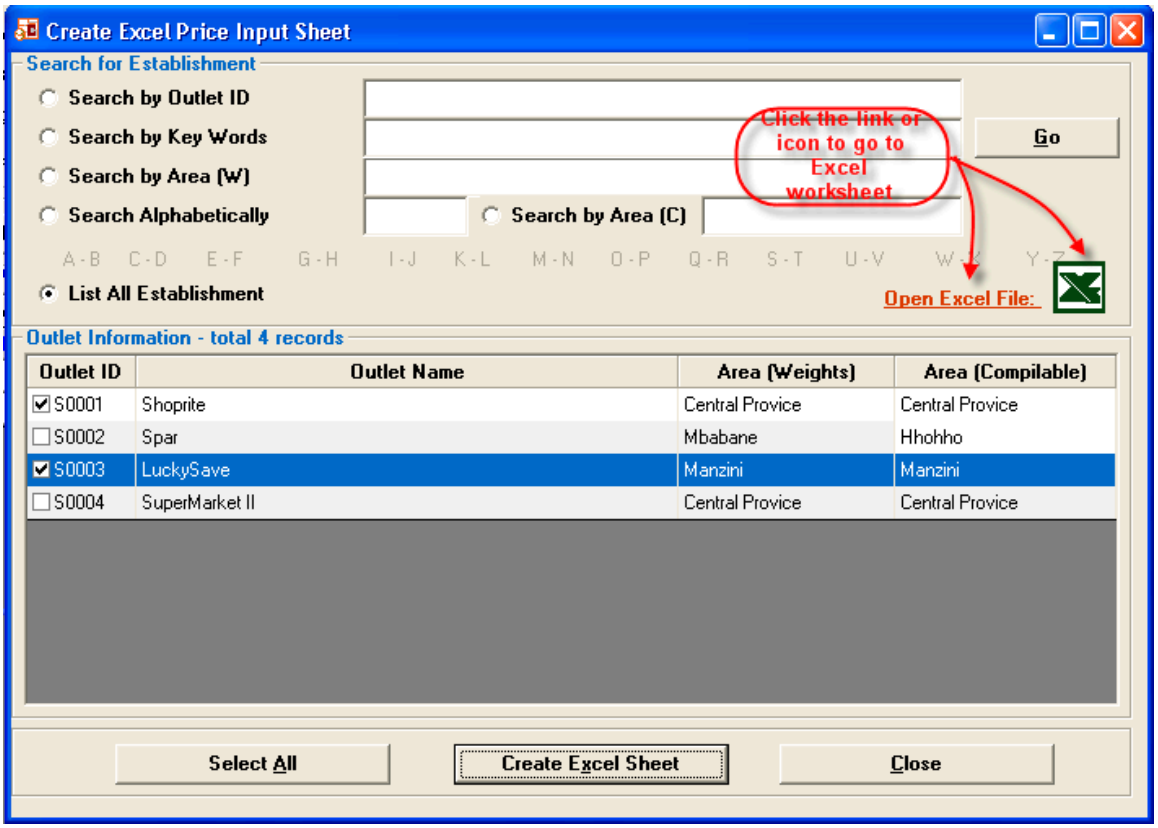
Select outlets by ticking the small square boxes in the first column. Then click **Create Excel Sheet** button.



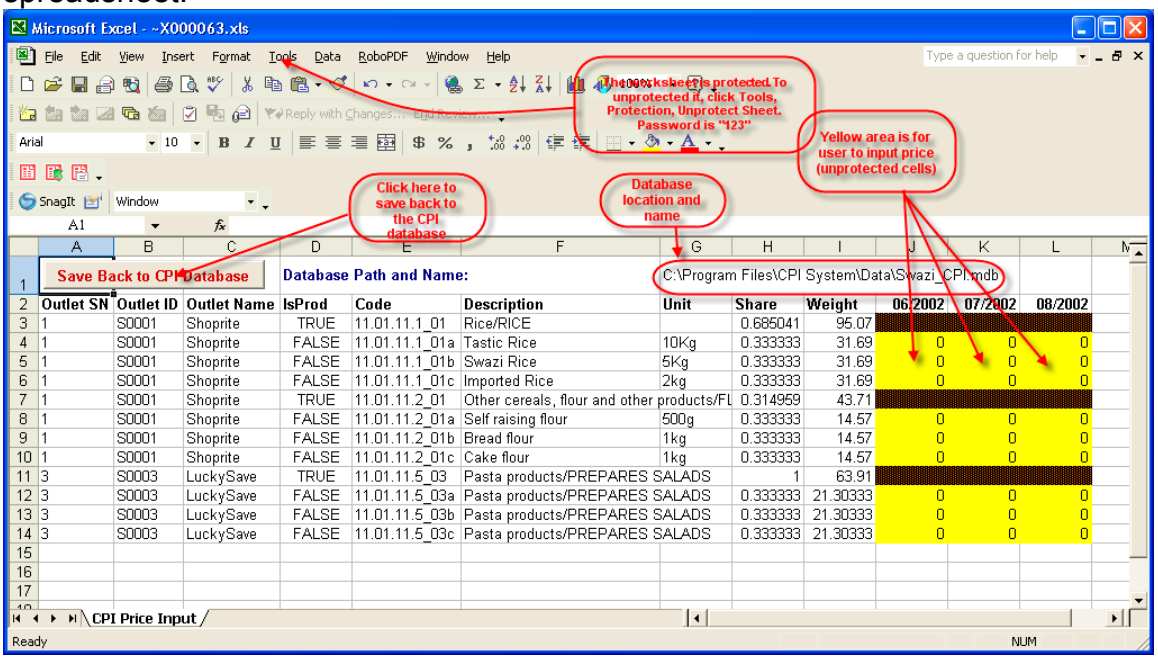
The user will see the following screen: which allow the user to specify the start period and the number of period.



Click **OK** button. The user will see the previous screen with the Excel Icon on the right hand side of the form.



Click the **Open Excel File** link or Excel icon, the user will see following Excel spreadsheet:



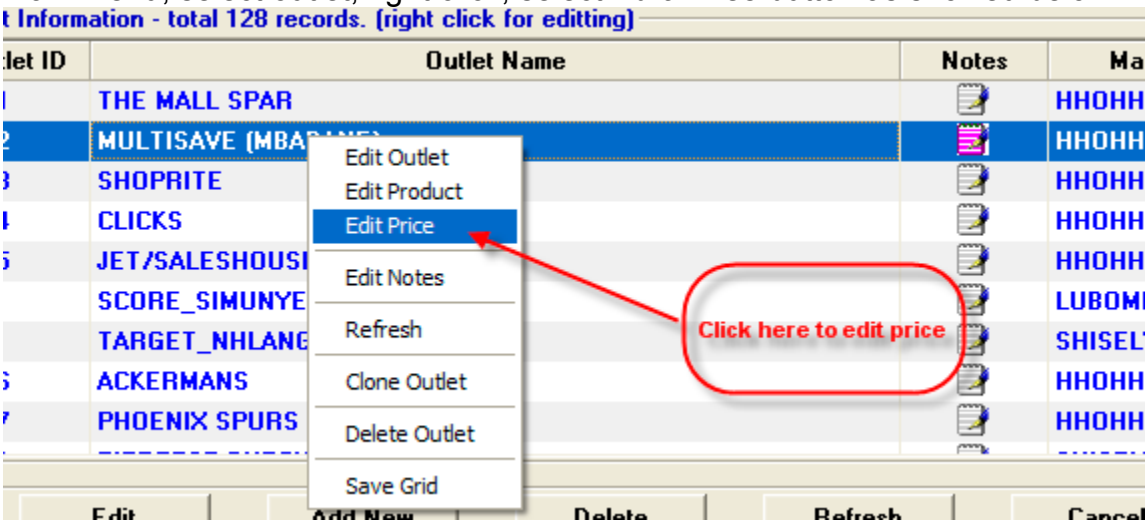
Outlet SN	Outlet ID	Outlet Name	IsProd	Code	Description	Unit	Share	Weight	06/2002	07/2002	08/2002
1	S0001	Shoprite	TRUE	11.01.11.1_01	Rice/RICE		0.685041	95.07			
1	S0001	Shoprite	FALSE	11.01.11.1_01a	Tastic Rice	10Kg	0.333333	31.69	0	0	0
1	S0001	Shoprite	FALSE	11.01.11.1_01b	Swazi Rice	5Kg	0.333333	31.69	0	0	0
1	S0001	Shoprite	FALSE	11.01.11.1_01c	Imported Rice	2kg	0.333333	31.69	0	0	0
1	S0001	Shoprite	TRUE	11.01.11.2_01	Other cereals, flour and other products/FL		0.314959	43.71			
1	S0001	Shoprite	FALSE	11.01.11.2_01a	Self raising flour	500g	0.333333	14.57	0	0	0
1	S0001	Shoprite	FALSE	11.01.11.2_01b	Bread flour	1kg	0.333333	14.57	0	0	0
1	S0001	Shoprite	FALSE	11.01.11.2_01c	Cake flour	1kg	0.333333	14.57	0	0	0
3	S0003	LuckySave	TRUE	11.01.11.5_03	Pasta products/PREPARES SALADS		1	63.91			
3	S0003	LuckySave	FALSE	11.01.11.5_03a	Pasta products/PREPARES SALADS		0.333333	21.30333	0	0	0
3	S0003	LuckySave	FALSE	11.01.11.5_03b	Pasta products/PREPARES SALADS		0.333333	21.30333	0	0	0
3	S0003	LuckySave	FALSE	11.01.11.5_03c	Pasta products/PREPARES SALADS		0.333333	21.30333	0	0	0

This worksheet is protected except for the yellow area, which is the area where the user can input price information. It is important to note that this is a pre-designed spreadsheet with a specific table format for price update purposes. Some cell information is for technical purposes and the table structure cannot be changed at the user's discretion. Any unsolicited change of table structure will cause errors and data cannot be saved properly. But if it is absolutely necessary, the user can unprotect this worksheet by clicking **Tools, Protection, Protect Sheet** button. The password is "123".

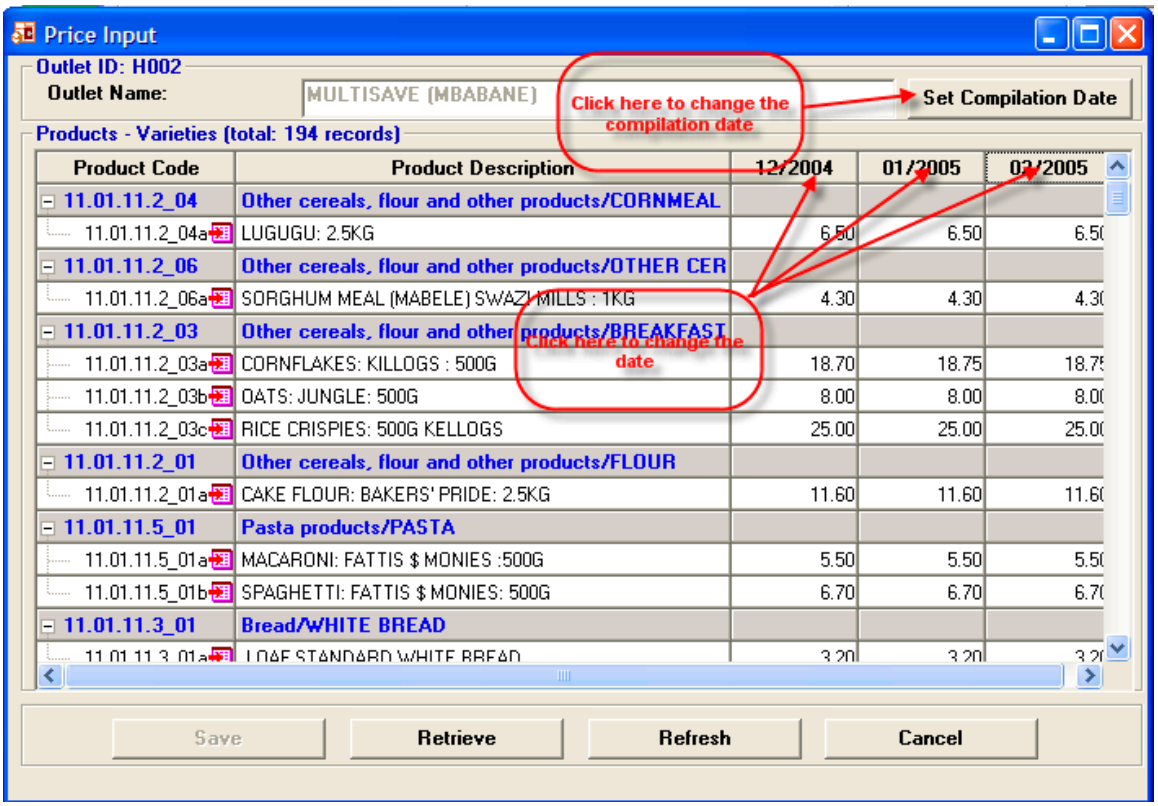
Method 2, Enter/Update Price Through Windows Form

Click **Price Input** 

button in edit product screen, or click **Edit** from main menu, select outlet, right click, select **Edit Price** button as showed below:

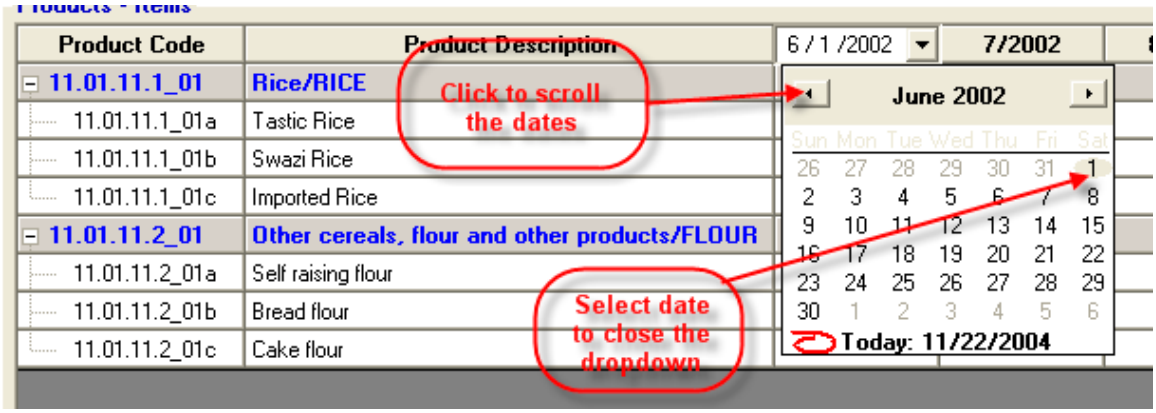


Edit price screen looks like:

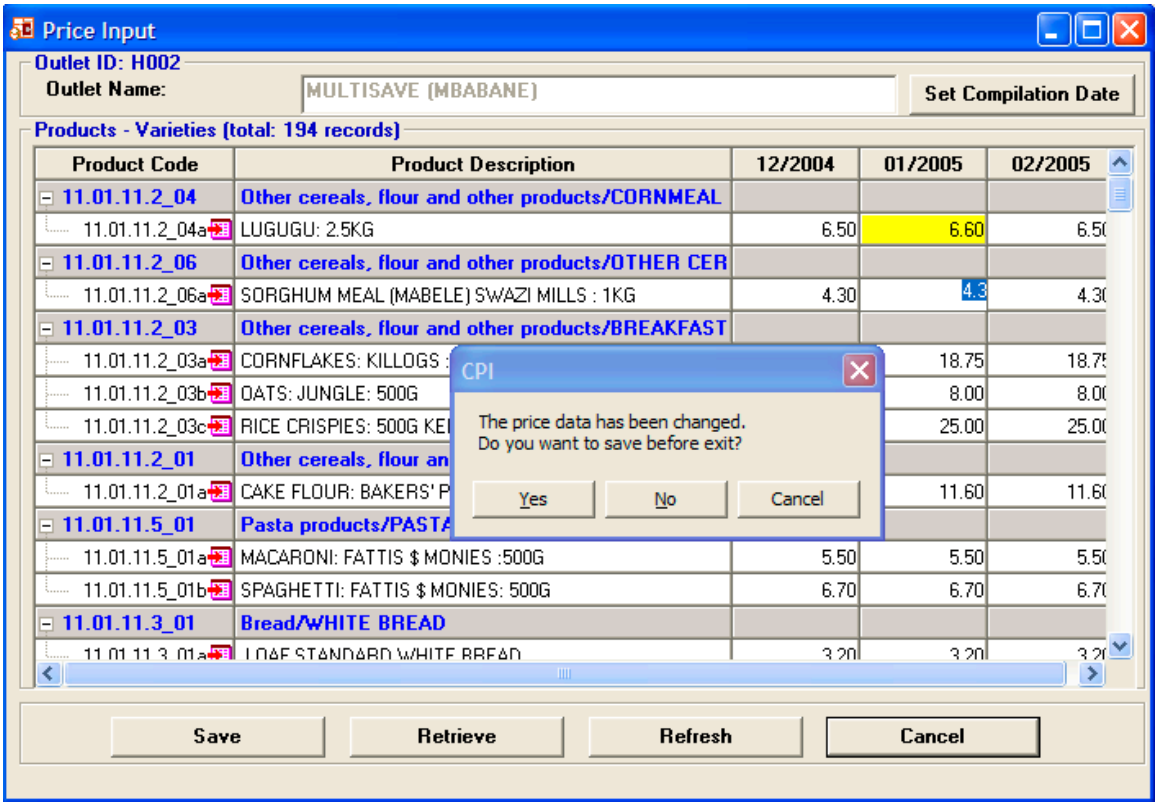


Enter price in relevant cells. The user can enter/edit price for 3 month (for base period, preceding period and current period). Note that once cell in editing mode the background color will turn yellow.

The user can retrieve historical price data by clicking **Retrieve** button. To change the date, the user can click date cell.



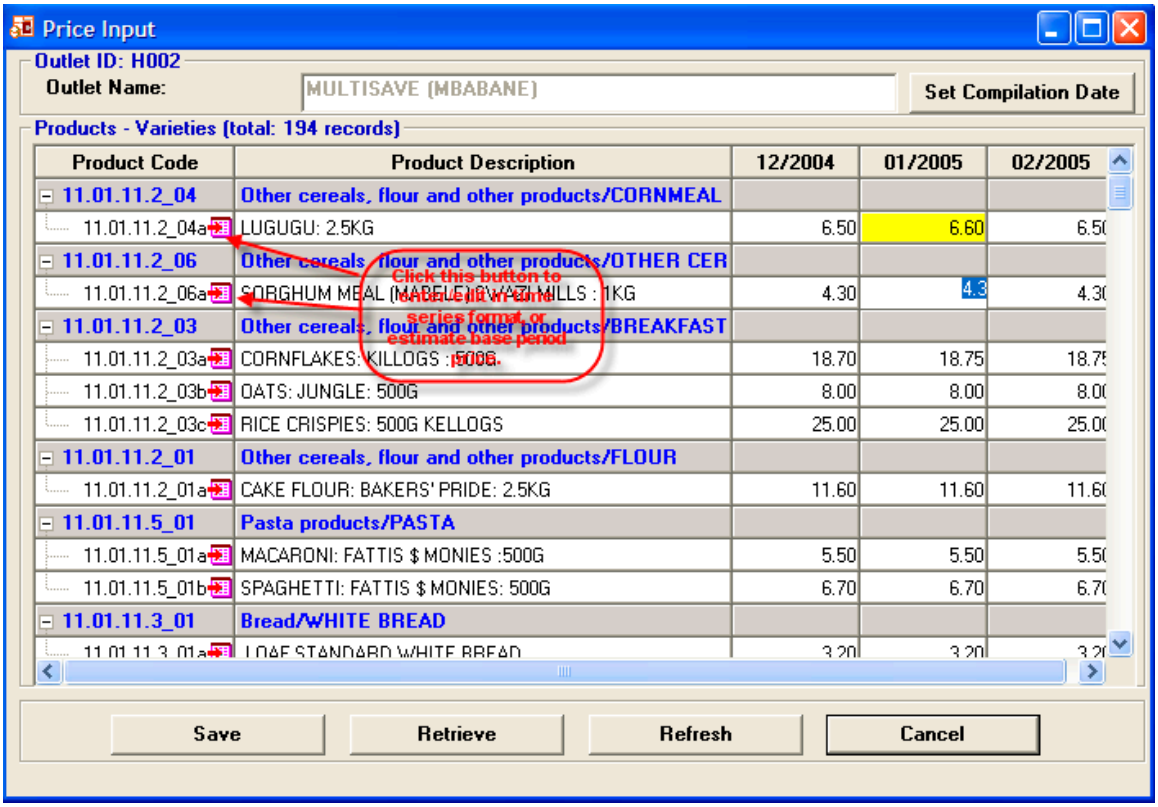
If the user exit the price edit form without saving revised data, he will see following message:



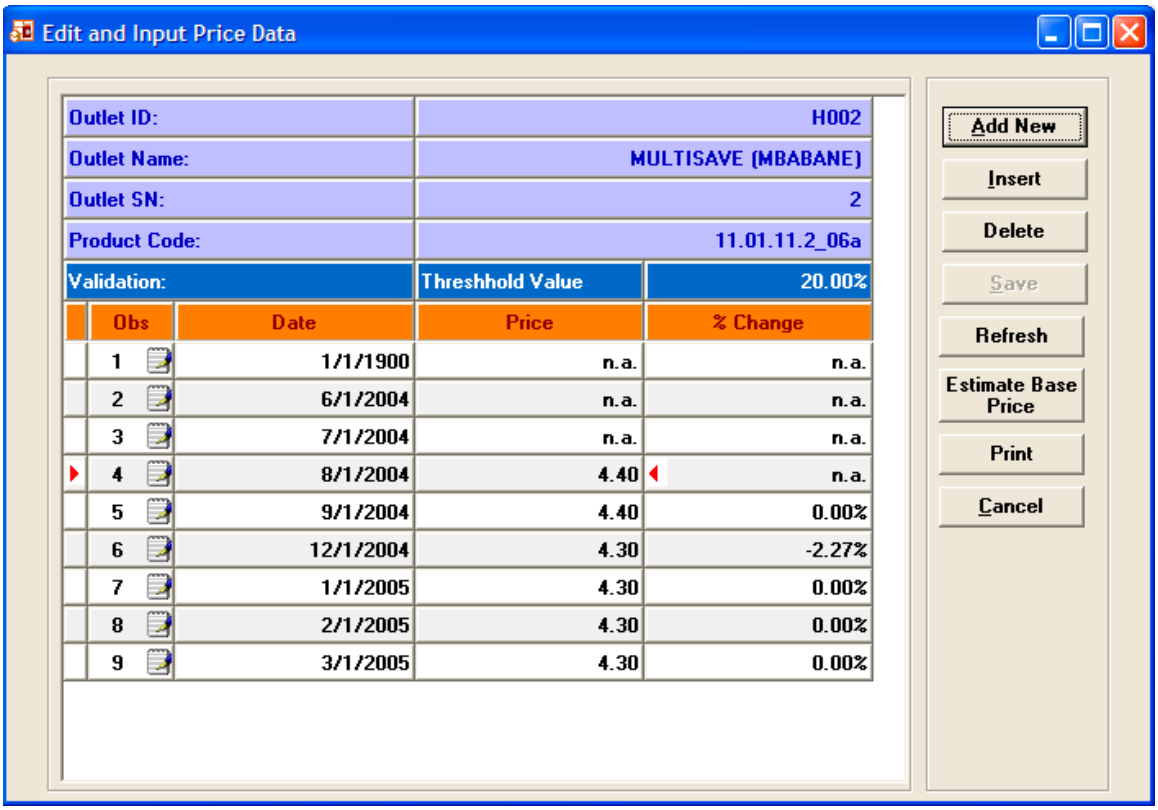
Click **Yes** to save the information, click **No** to ignore the changes, click **Cancel** to cancel the exit and continue to edit the price.

Method 3, Enter/Update Price in Time Series Format

The user can enter/edit price data for one specific variety for multiple periods in time series format. Click the right side of cell that contains the variety code.



Click the red button, the user will be give following screen for price entry:

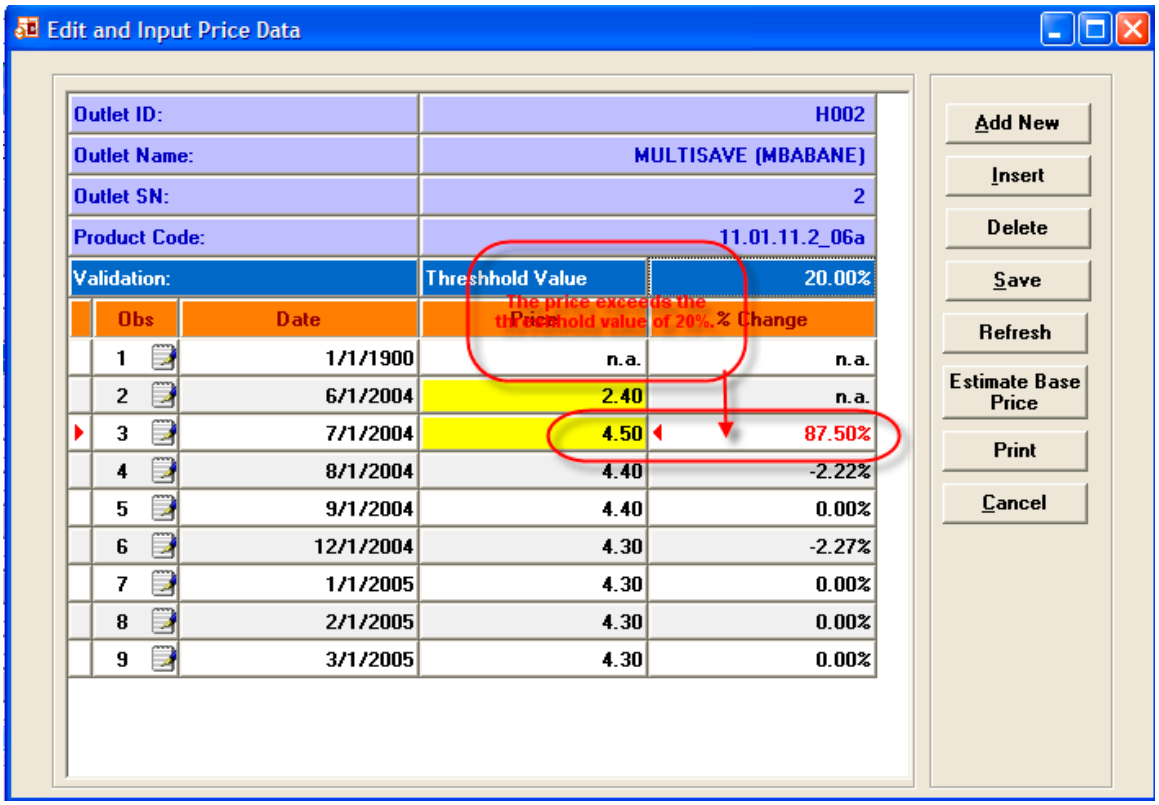


At this screen, the user can (a) add an observation at the end of time series; (b) can also insert one between two observations; (c) remove an observation; (d) estimate base period price using another series as a proxy. At this screen, the user can also define a threshold value, in percentage term. If the percentage increase exceeds this threshold value, the font color of percentage will be showed in red.

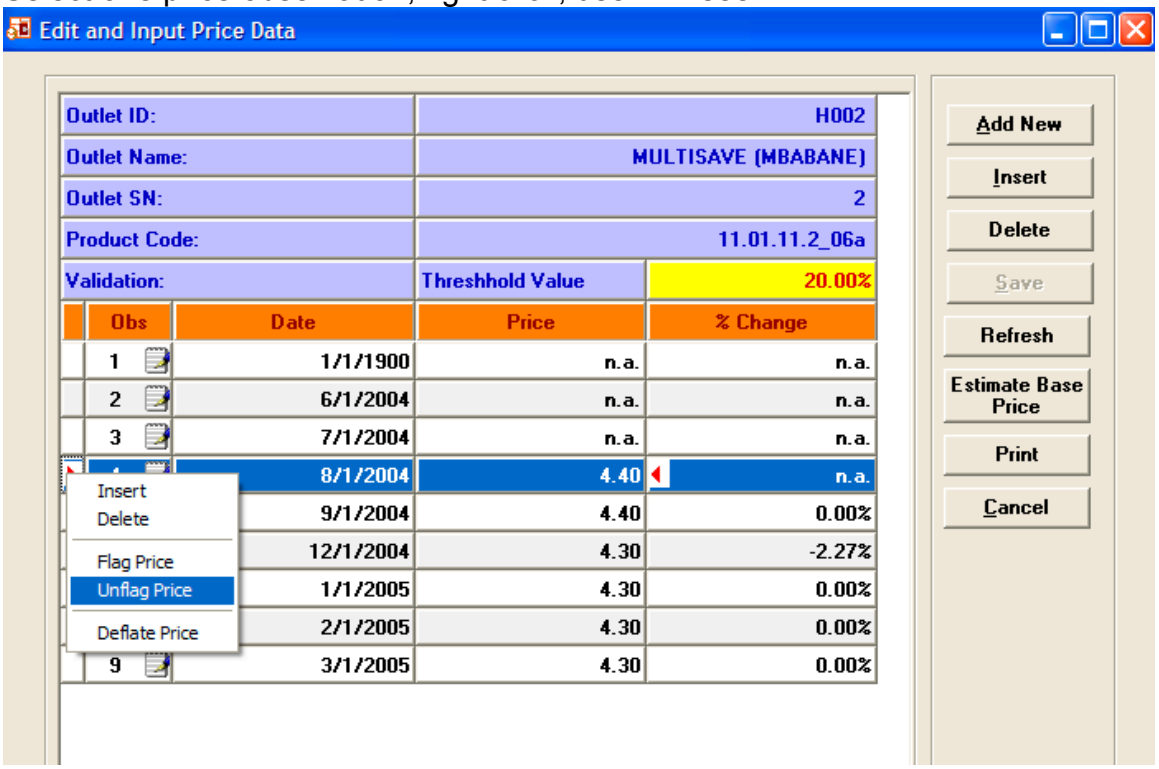
Please note that if the price value is in red, this indicates that this value is estimated value during the imputation.

If more than one price observations have been entered, the system will take the average of these prices in the compilation process. The system is able to handle multiple price collections per month and will impute averages to enter the index calculation.

Click icon to add a note to a price observation, and indicates that there is note for that observation.



Select one price observation, right click, user will see:

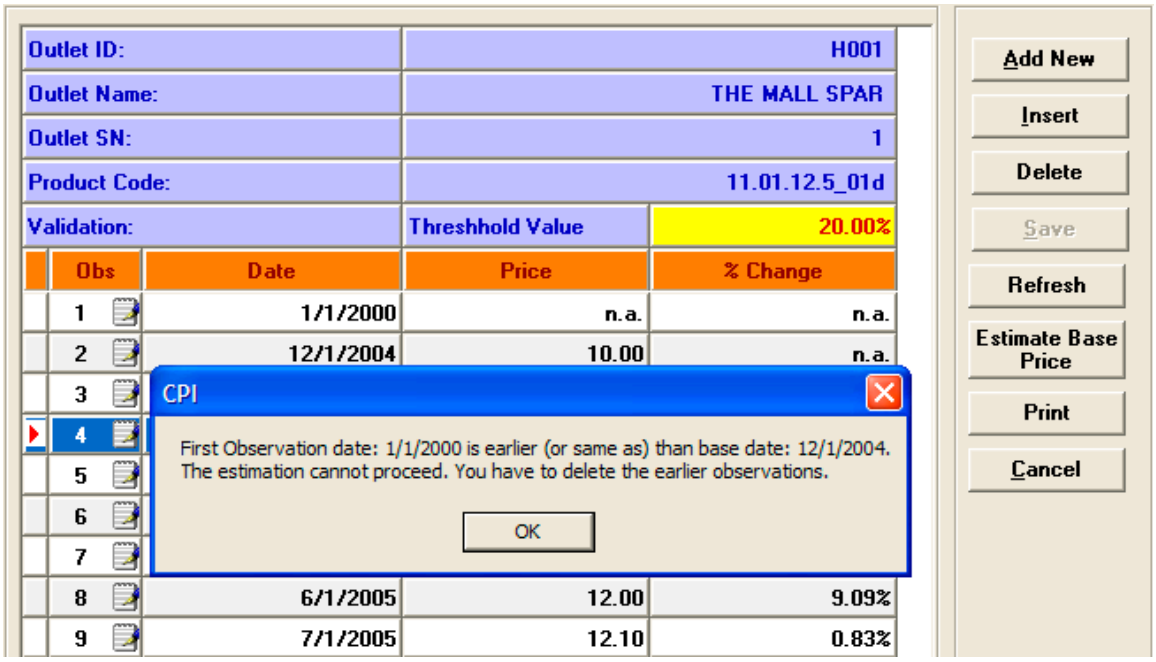


The user should be able to **Insert, Delete, Flag Price, Unflag Price or Deflate Price.**

When a new variety is introduced to substitute an old variety due to the quality change or a brand new variety is added to the price index compilation process, a base period price needs to be estimated.

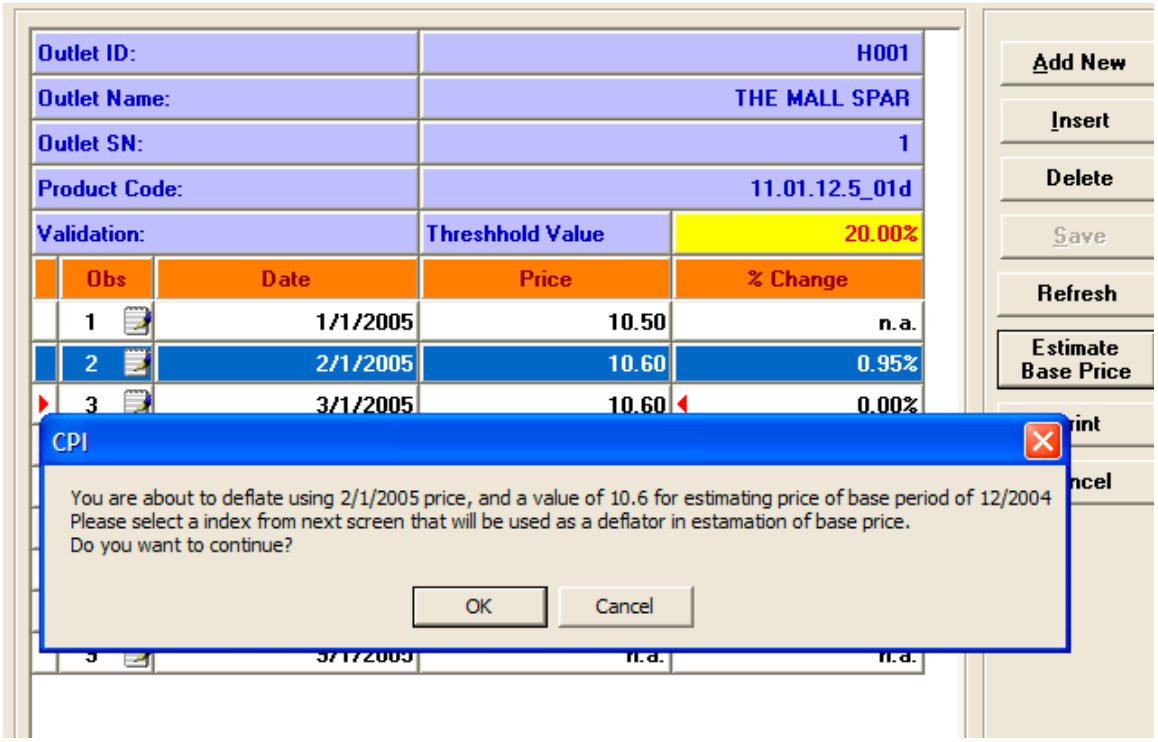
The user can click **Estimate Base Price** or **Deflate Price** button, to use another price index to estimate the base period price for a given price observation.

If an observation exists prior to or at same period of that of base period, the user has to delete them before estimation can be made.

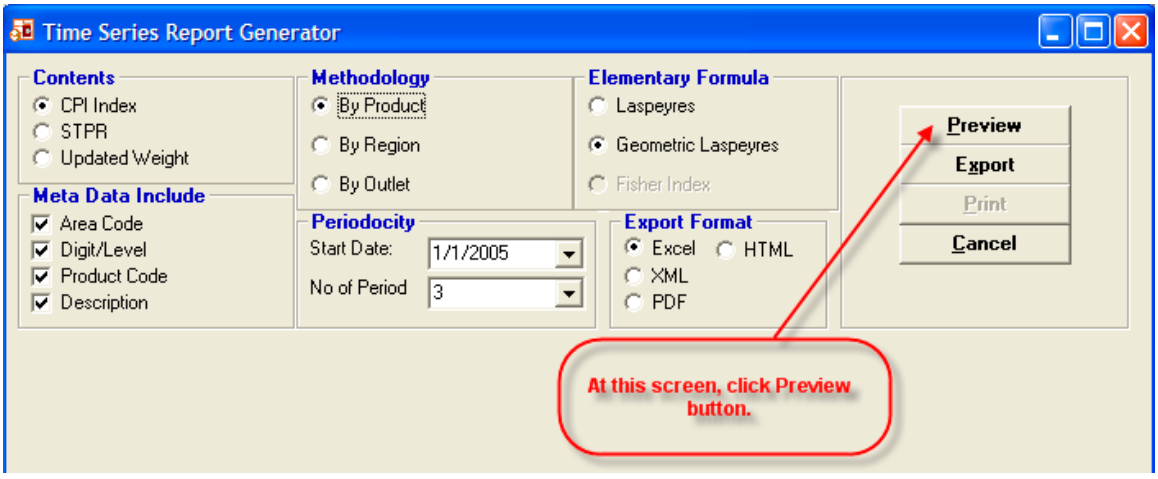


Delete the value by selecting the first column for the observation, right click and click **delete** button as showed in the diagram below:

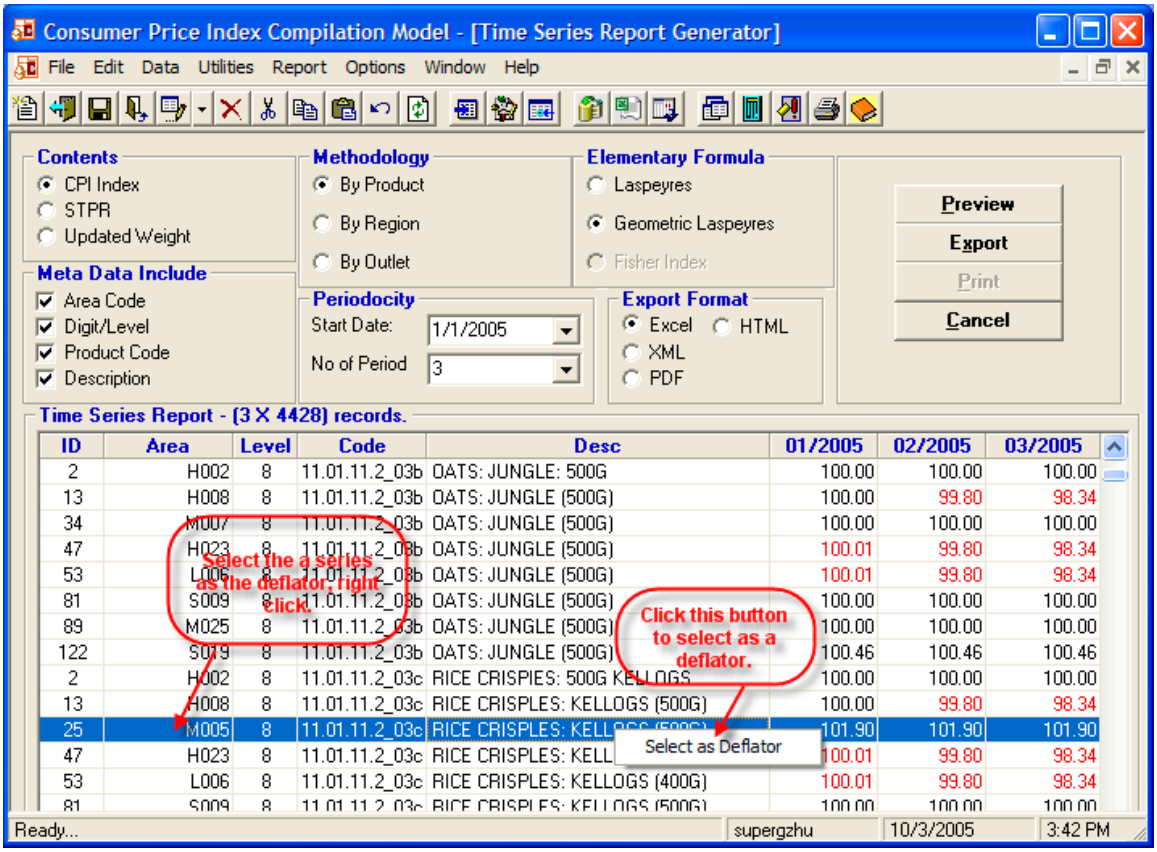
Then click the **Estimate Base Price** again.



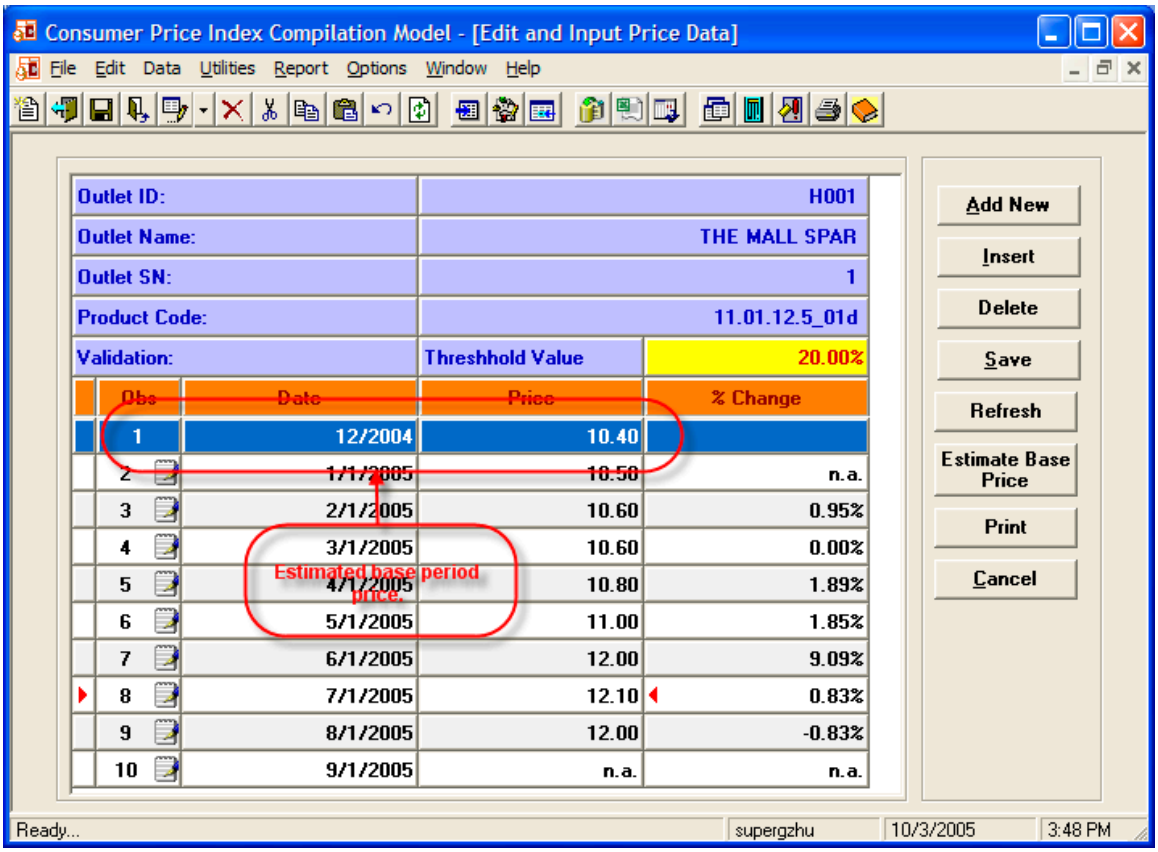
Click OK button to continue. At next screen click **Preview** button.



Then select a time series as the **deflator**:



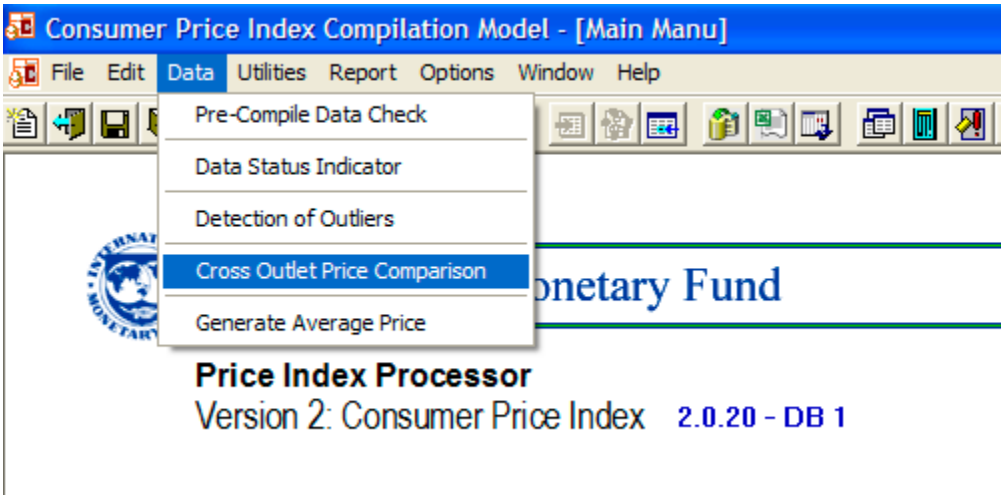
The base period price will be estimated as showed below:



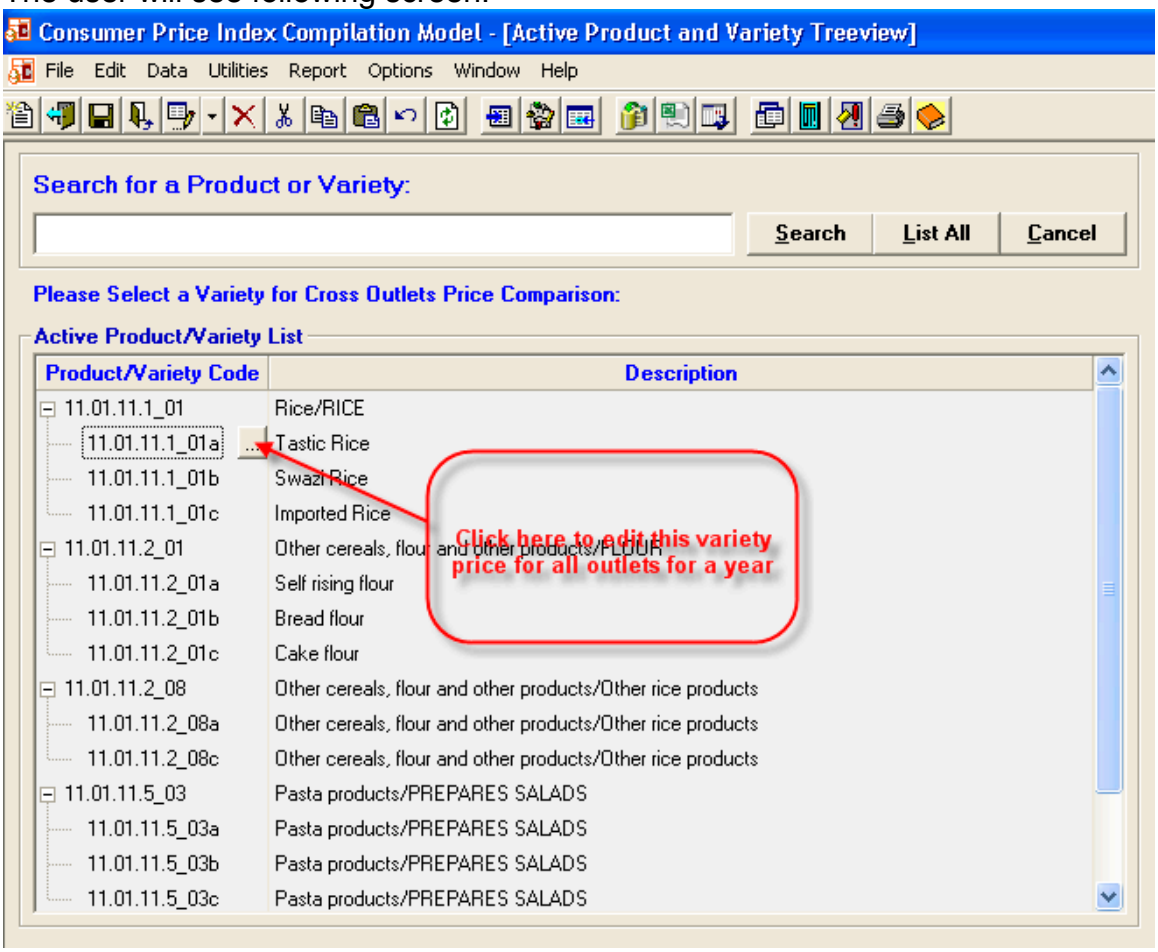
Click the **Save** button to save price information.

Method 4, Enter/Update Price for Multiple Outlets and for Single Variety

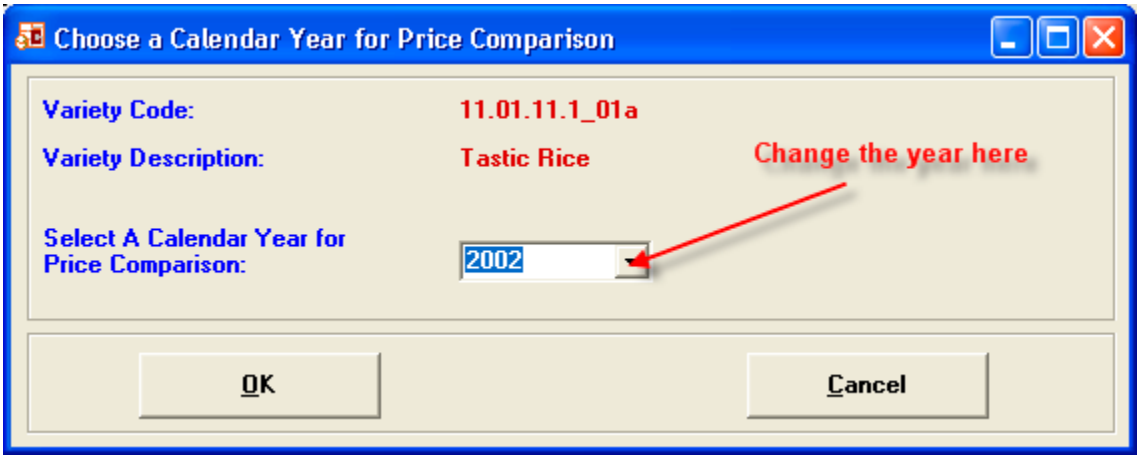
The user can also enter or edit price for multiple outlets for a single variety. Click **Data**, **Cross Outlet Price Comparison** button as showed below:



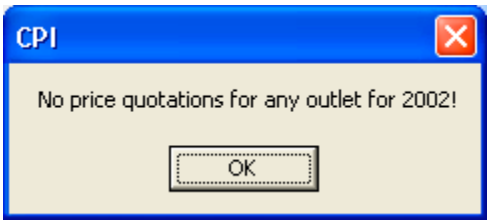
The user will see following screen:



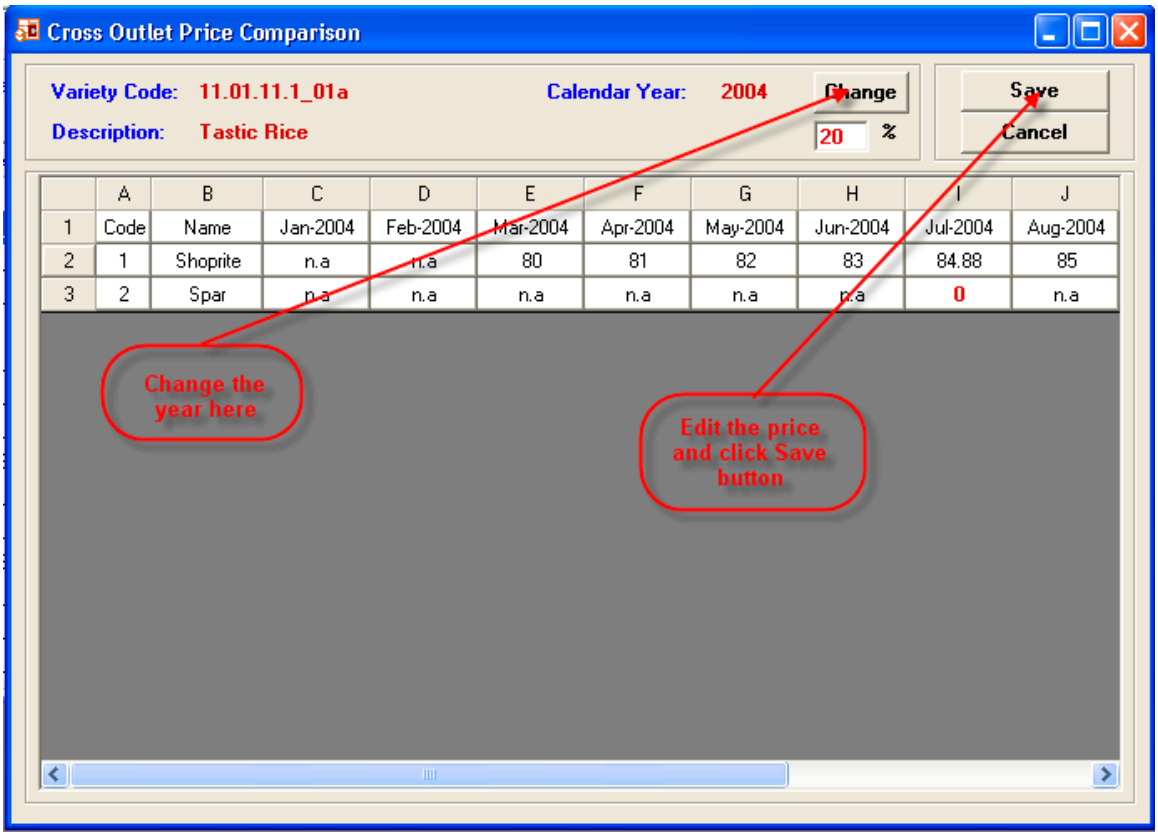
The user can click the grey button to edit price for variety 11.01.11.1_01a. The user will be given for the following screen to select a specific year:



Select year and then hit **OK** button. If no data is available, the user will see this message box:



Otherwise, the user will see:

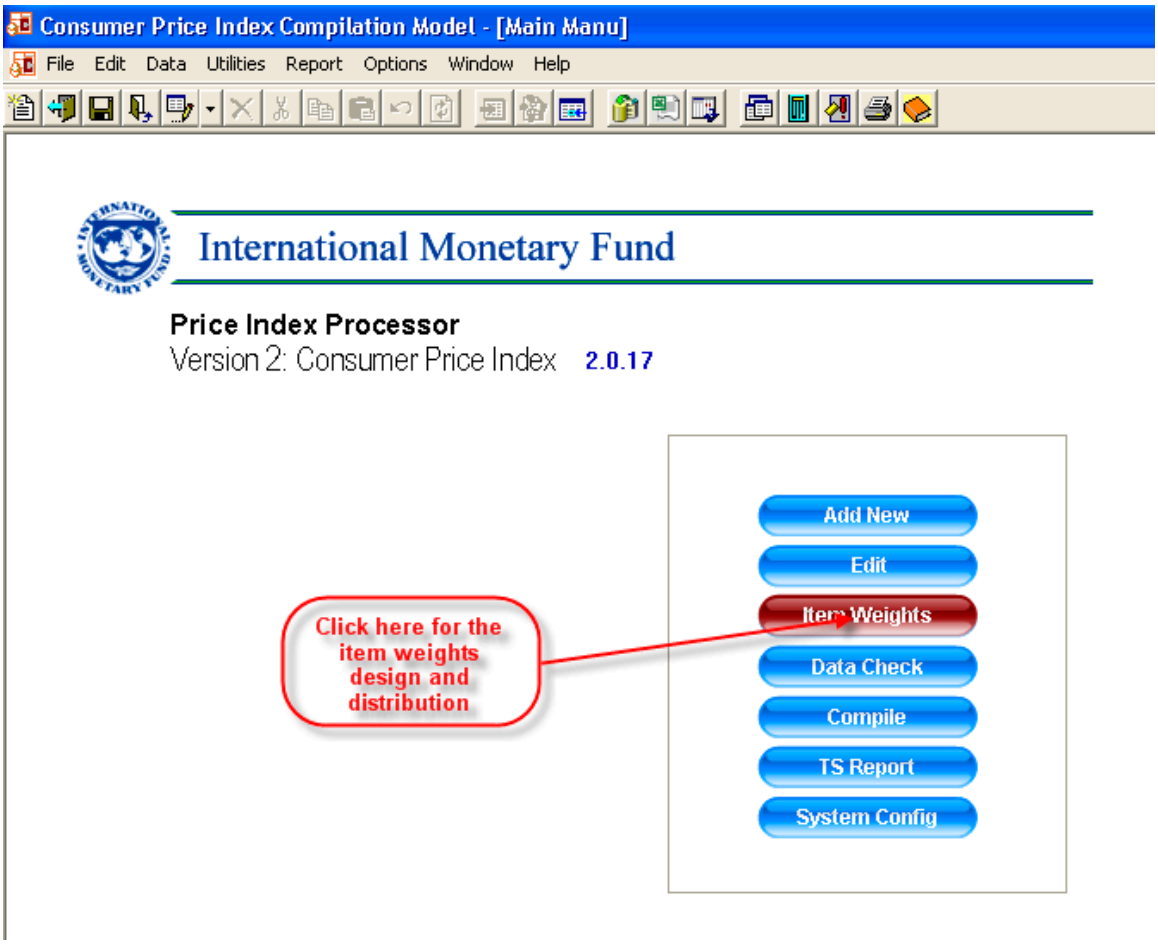


	A	B	C	D	E	F	G	H	I	J
1	Code	Name	Jan-2004	Feb-2004	Mar-2004	Apr-2004	May-2004	Jun-2004	Jul-2004	Aug-2004
2	1	Shoprite	n.a	n.a	80	81	82	83	84.88	85
3	2	Spar	n.a	n.a	n.a	n.a	n.a	n.a	0	n.a

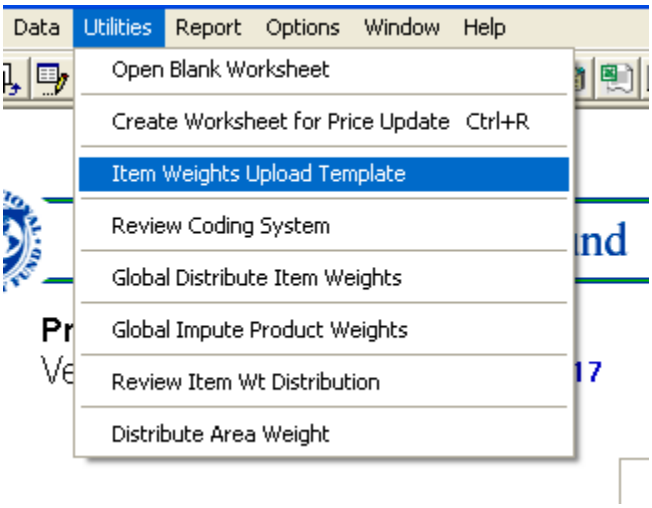
The user can change or input price here and hit the **Save** button.

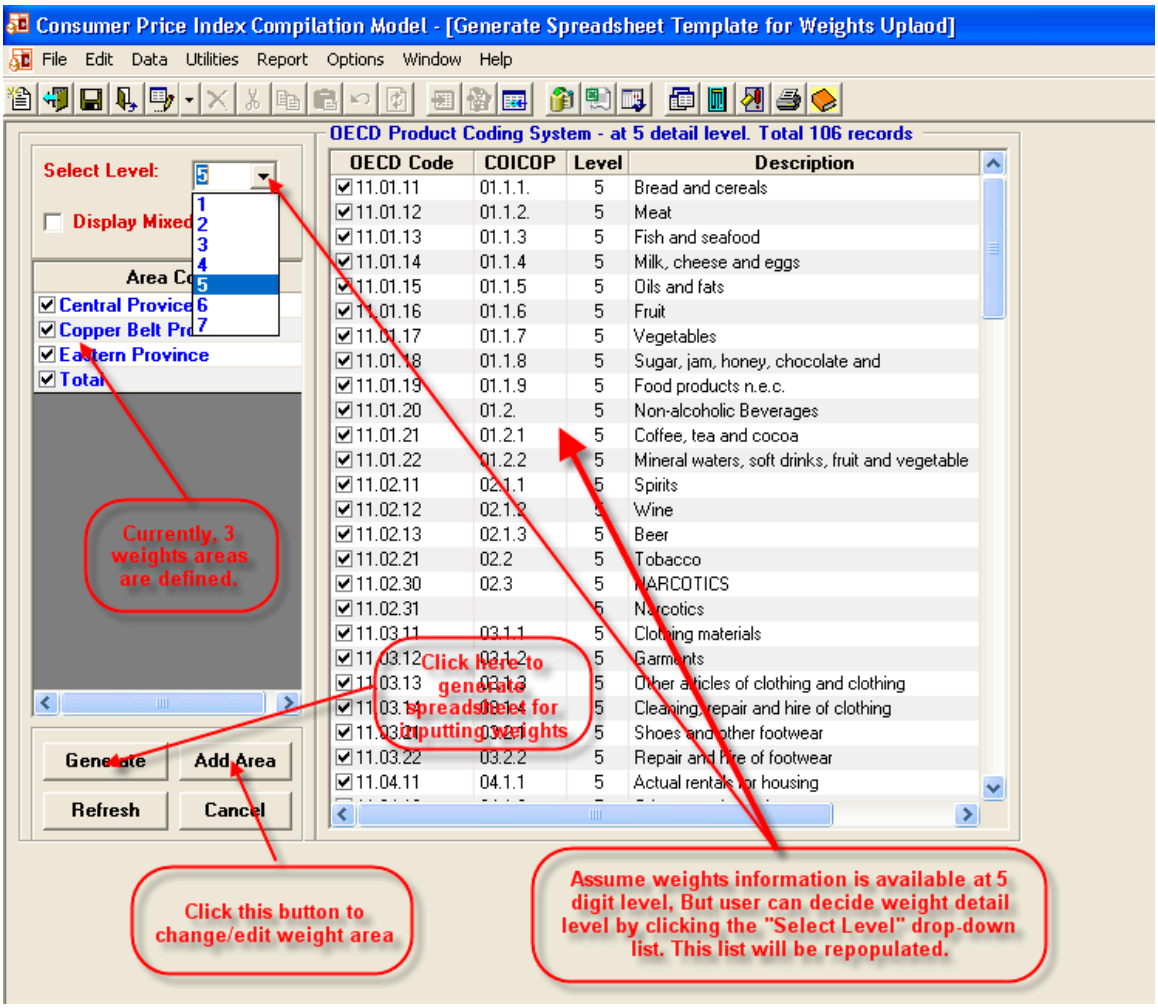
Item Weights Design and Distribution

Item weights are typically obtained from household expenditure survey by area and by each individual item. To start item weights distribution, the user clicks Item Weights at the main menu:

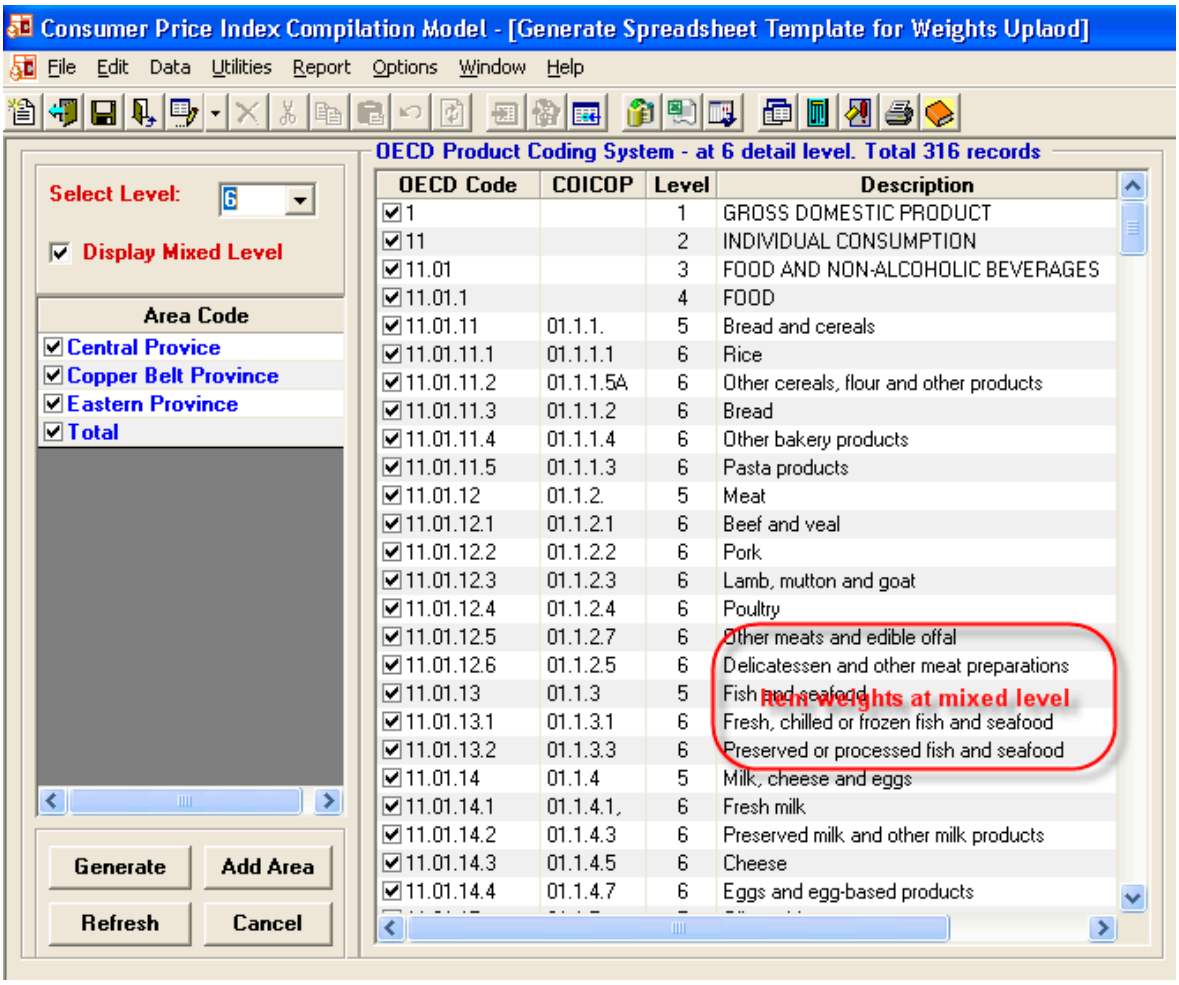


Click **Item Weights** button in the main menu or at click Utilities, Item Weights Upload Template button on toolbar:



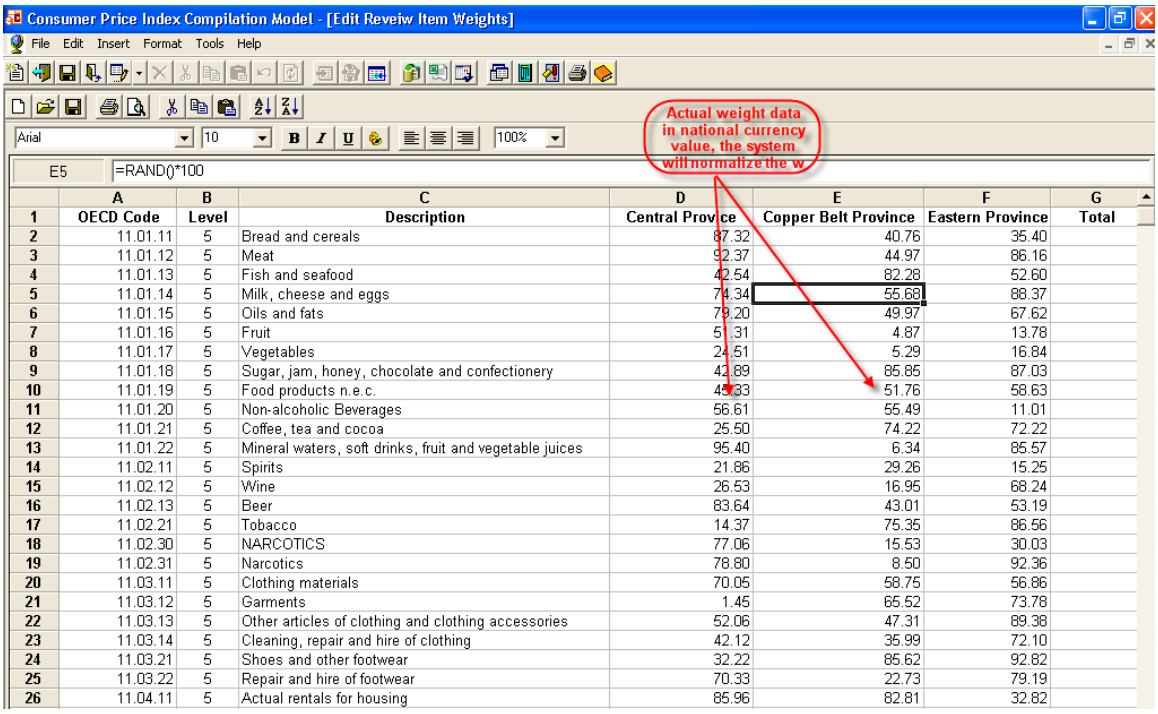


The user should decide the weights level from 1 to 7 according to the specific weight survey data. The user can also select item details at mixed level. The following screen shows that the user weight data is at mixed level:



It is important to note that the user should not select an item from level 5 and 6 at same time. For example, the user should not select “11.01.11 - Bread and Cereals” (level 5 item) and item 11.01.11.3 – Bread (level 6 item) simultaneously. By doing so, only weight data at more detail level will be distribution to the price quotations underneath.

The user clicks the **Generate** button, and inputs the weight data on the following screen:

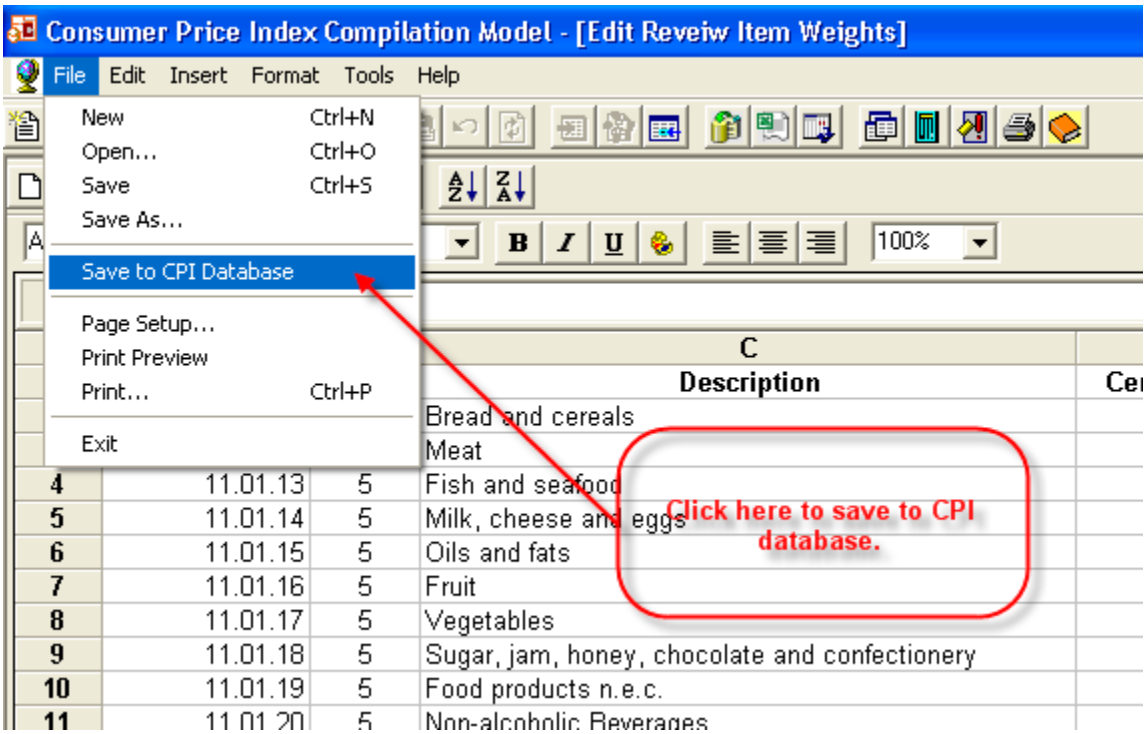


	A	B	C	D	E	F	G
1	OECD Code	Level	Description	Central Province	Copper Belt Province	Eastern Province	Total
2	11.01.11	5	Bread and cereals	87.32	40.76	35.40	
3	11.01.12	5	Meat	92.37	44.97	86.16	
4	11.01.13	5	Fish and seafood	42.54	82.28	52.60	
5	11.01.14	5	Milk, cheese and eggs	74.34	55.68	88.37	
6	11.01.15	5	Oils and fats	79.20	49.97	67.62	
7	11.01.16	5	Fruit	51.31	4.87	13.78	
8	11.01.17	5	Vegetables	24.51	5.29	16.84	
9	11.01.18	5	Sugar, jam, honey, chocolate and confectionery	42.89	85.85	87.03	
10	11.01.19	5	Food products n.e.c.	43.83	51.76	58.63	
11	11.01.20	5	Non-alcoholic Beverages	56.61	55.49	11.01	
12	11.01.21	5	Coffee, tea and cocoa	25.50	74.22	72.22	
13	11.01.22	5	Mineral waters, soft drinks, fruit and vegetable juices	95.40	6.34	85.57	
14	11.02.11	5	Spirits	21.86	29.26	15.25	
15	11.02.12	5	Wine	26.53	16.95	68.24	
16	11.02.13	5	Beer	83.64	43.01	53.19	
17	11.02.21	5	Tobacco	14.37	75.35	86.56	
18	11.02.30	5	NARCOTICS	77.06	15.53	30.03	
19	11.02.31	5	Narcotics	78.80	8.50	92.36	
20	11.03.11	5	Clothing materials	70.05	58.75	56.86	
21	11.03.12	5	Garments	1.45	65.52	73.78	
22	11.03.13	5	Other articles of clothing and clothing accessories	52.06	47.31	89.38	
23	11.03.14	5	Cleaning, repair and hire of clothing	42.12	35.99	72.10	
24	11.03.21	5	Shoes and other footwear	32.22	85.62	92.82	
25	11.03.22	5	Repair and hire of footwear	70.33	22.73	79.19	
26	11.04.11	5	Actual rentals for housing	85.96	82.81	32.82	

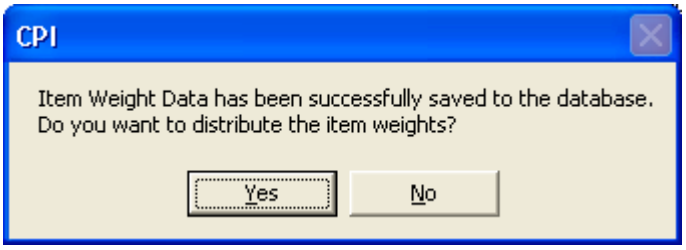
Please note that the user should input the unprocessed weight data in national currency denomination (raw weight survey data), even though correctly normalized data are also acceptable.

To correctly normalize the weight data, sum the columns to get regional totals first; then sum the regional totals to get the national total. Divide each cell by national total to get a fraction of each item in a specific weight area as a percentage of the national total—though the CPI System prefers to have expenditure weight data without normalization.

The user then should save the weight data to the CPI database by clicking **File**, **Save to CPI Database** (weight data will be saved to the default CPI database; use the system configuration to change the default database) as showed in the screen below:



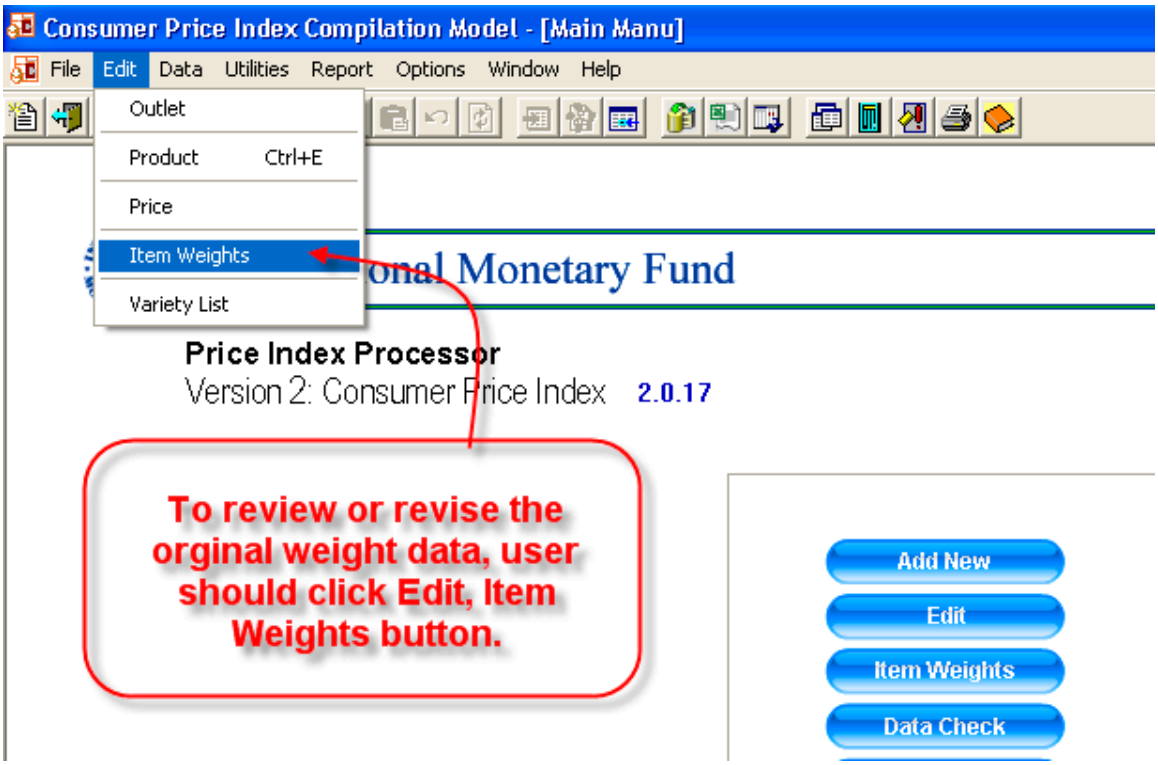
After the weight data are saved, the user will be given following message box:



Click Yes button to have the weight distributed to the price quotation under this item.

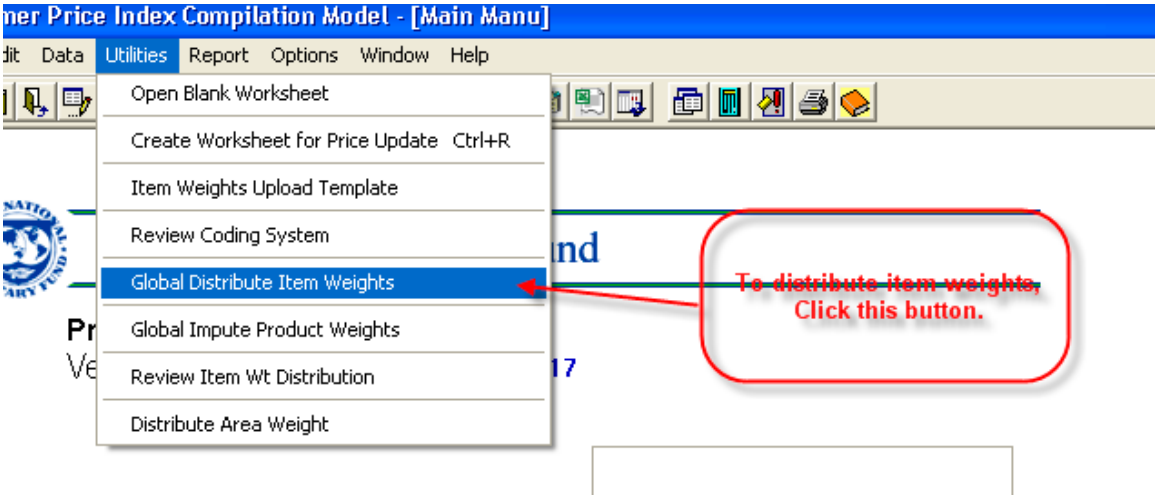
The message box indicates whether the item weight has been distributed successfully.

To review or revise the original item weight data, the user should click:

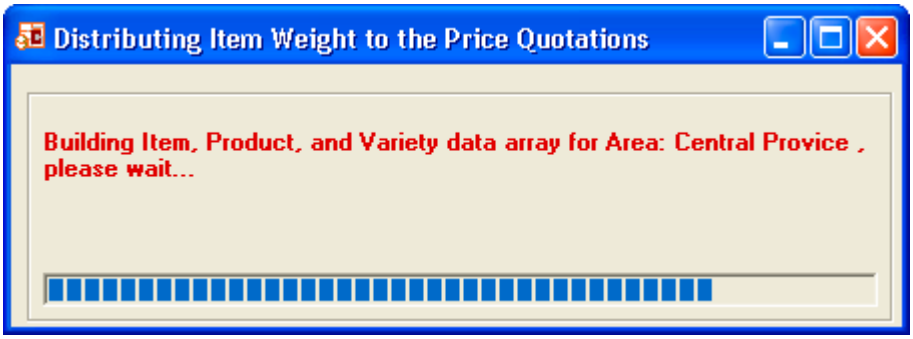


The CPI System can retrieve the weight data from database and display the data in spreadsheet format. The user can revise the item data if necessary, and click Save to CPI Database button.

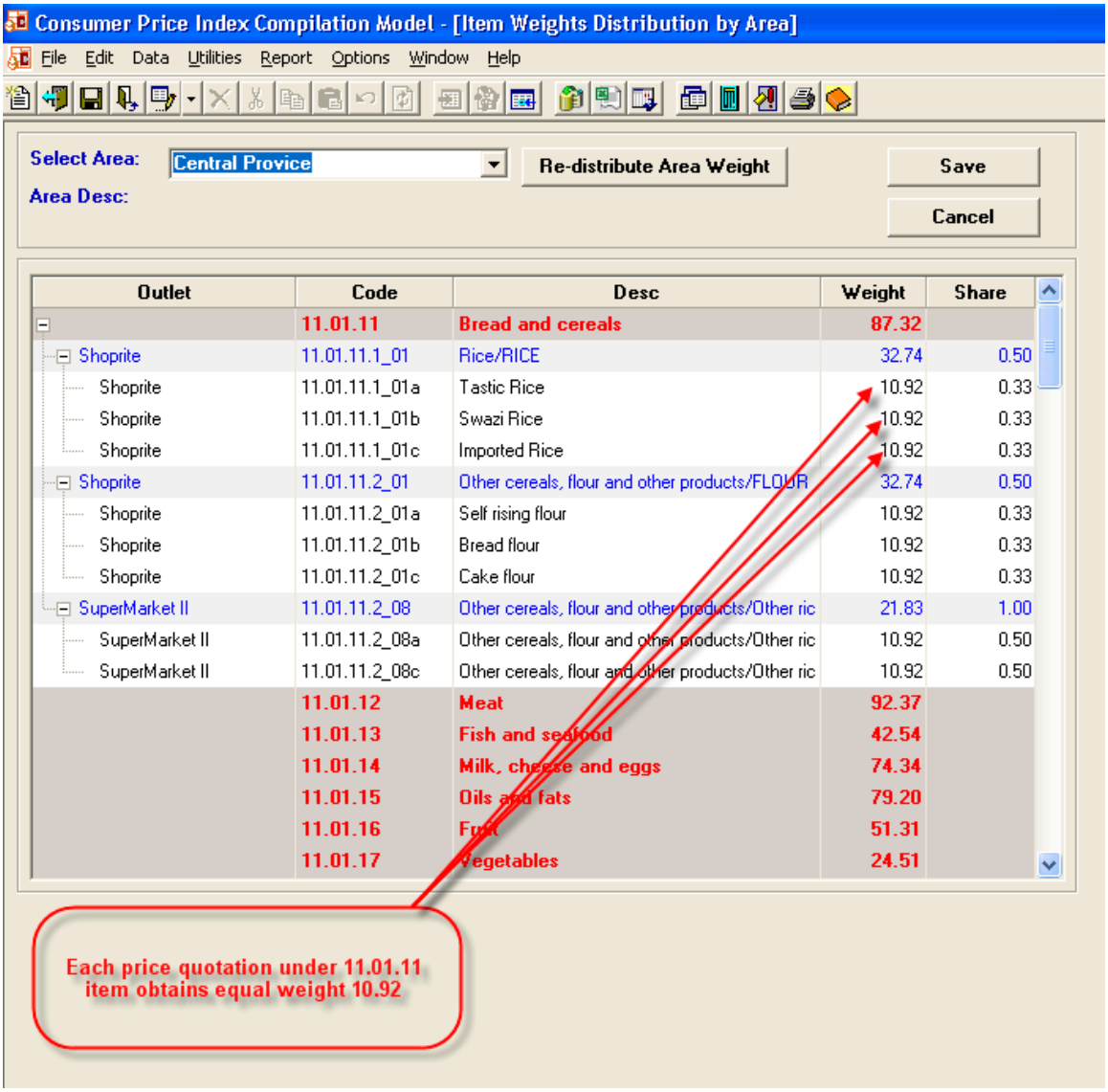
An area item weight will be equally distributed to the every price quotation under this item if the user click **Utilities**, **Global Distribute Item Weights** button.



It is important to note that the user can review the weight distribution by clicking **Review Item Wt Distribution** button showed in above screen.



Wait until the system finishes the process. The user should be given the information of how weights are distributed across the outlets.



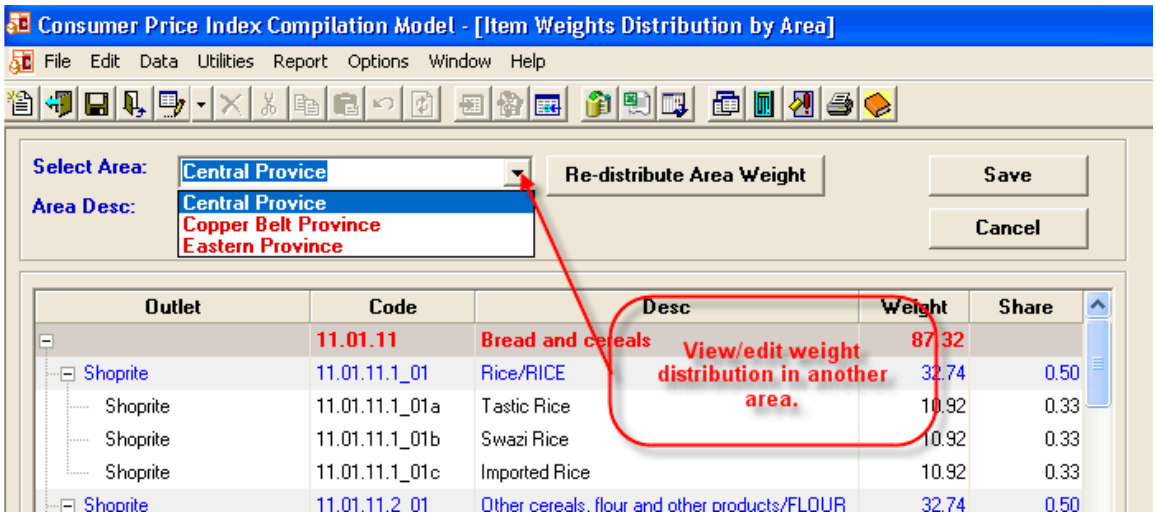
Outlet	Code	Desc	Weight	Share
	11.01.11	Bread and cereals	87.32	
Shoprite	11.01.11.1_01	Rice/RICE	32.74	0.50
Shoprite	11.01.11.1_01a	Tastic Rice	10.92	0.33
Shoprite	11.01.11.1_01b	Swazi Rice	10.92	0.33
Shoprite	11.01.11.1_01c	Imported Rice	10.92	0.33
Shoprite	11.01.11.2_01	Other cereals, flour and other products/FLOUR	32.74	0.50
Shoprite	11.01.11.2_01a	Self rising flour	10.92	0.33
Shoprite	11.01.11.2_01b	Bread flour	10.92	0.33
Shoprite	11.01.11.2_01c	Cake flour	10.92	0.33
SuperMarket II	11.01.11.2_08	Other cereals, flour and other products/Other ric	21.83	1.00
SuperMarket II	11.01.11.2_08a	Other cereals, flour and other products/Other ric	10.92	0.50
SuperMarket II	11.01.11.2_08c	Other cereals, flour and other products/Other ric	10.92	0.50
	11.01.12	Meat	92.37	
	11.01.13	Fish and seafood	42.54	
	11.01.14	Milk, cheese and eggs	74.34	
	11.01.15	Oils and fats	79.20	
	11.01.16	Fruit	51.31	
	11.01.17	Vegetables	24.51	

The item weight for Central Province (an weight area) for item 11.01.11 – Bread and cereals is 87.32 which is equally distributed among the price quotations under it.

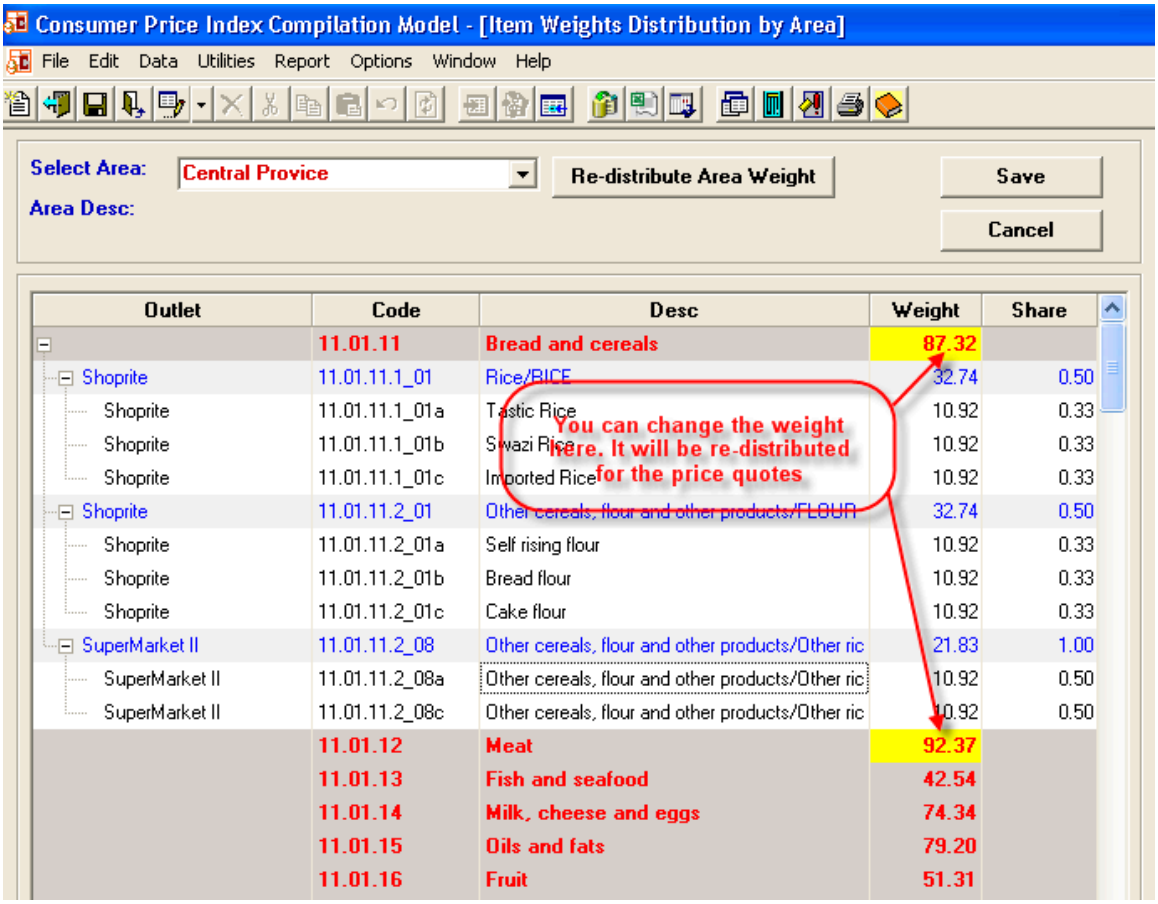
	A	B	C	D	E
	OECD Code	Level	Description	Central Provice	Copper Belt Pro
	11.01.11	5	Bread and cereals	87.32	40.76
	11.01.12	5	Meat	92.37	44.97
	11.01.13	5	Fish and seafood	42.54	82.28
	11.01.14	5	Milk, cheese and eggs	74.34	55.68
	11.01.15	5	Oils and fats	79.2	49.97
	11.01.16	5	Fruit	51.31	4.87
	11.01.17	5	Vegetables	24.51	5.29

This weight is equally distributed.

To view another area's weight distribution, click:



The weight for particular item for a specific area are editable, Suppose that the user wants to change the weight data for item 11.01.11 for Central Province, she can double click 87,32 and input a new number. This new figure will be redistributed based on the existing share.



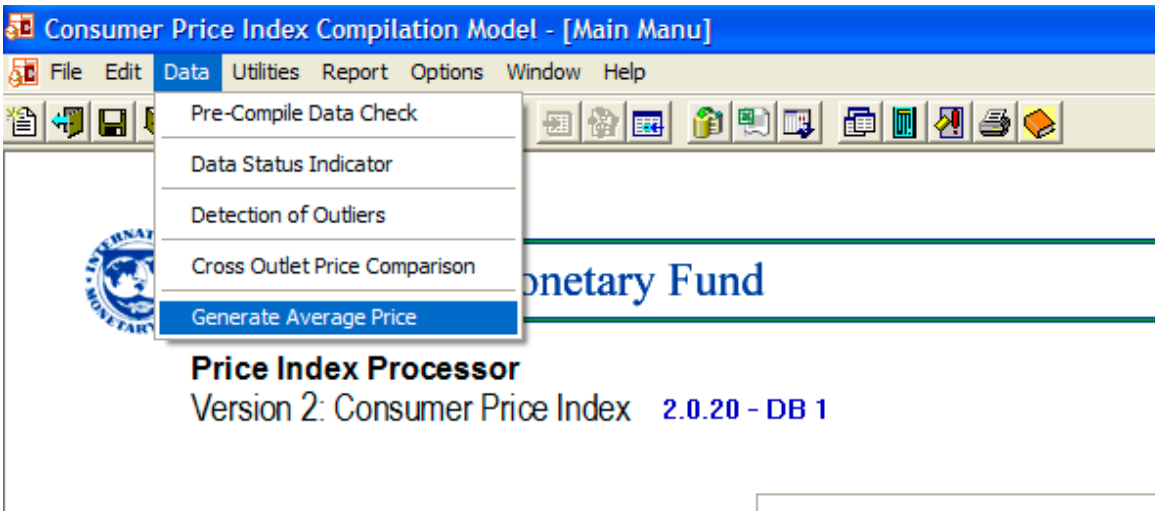
Outlet	Code	Desc	Weight	Share
	11.01.11	Bread and cereals	87.32	
Shoprite	11.01.11.1_01	Rice/RICE	32.74	0.50
Shoprite	11.01.11.1_01a	Tastic Rice	10.92	0.33
Shoprite	11.01.11.1_01b	Swazi Rice	10.92	0.33
Shoprite	11.01.11.1_01c	Imported Rice	10.92	0.33
Shoprite	11.01.11.2_01	Other cereals, flour and other products/FLOUR	32.74	0.50
Shoprite	11.01.11.2_01a	Self rising flour	10.92	0.33
Shoprite	11.01.11.2_01b	Bread flour	10.92	0.33
Shoprite	11.01.11.2_01c	Cake flour	10.92	0.33
SuperMarket II	11.01.11.2_08	Other cereals, flour and other products/Other ric	21.83	1.00
SuperMarket II	11.01.11.2_08a	Other cereals, flour and other products/Other ric	10.92	0.50
SuperMarket II	11.01.11.2_08c	Other cereals, flour and other products/Other ric	10.92	0.50
	11.01.12	Meat	92.37	
	11.01.13	Fish and seafood	42.54	
	11.01.14	Milk, cheese and eggs	74.34	
	11.01.15	Oils and fats	79.20	
	11.01.16	Fruit	51.31	

Rows in red indicate the items whose weights are available. If there are products or varieties below these entries the view can expand via a tree structure.

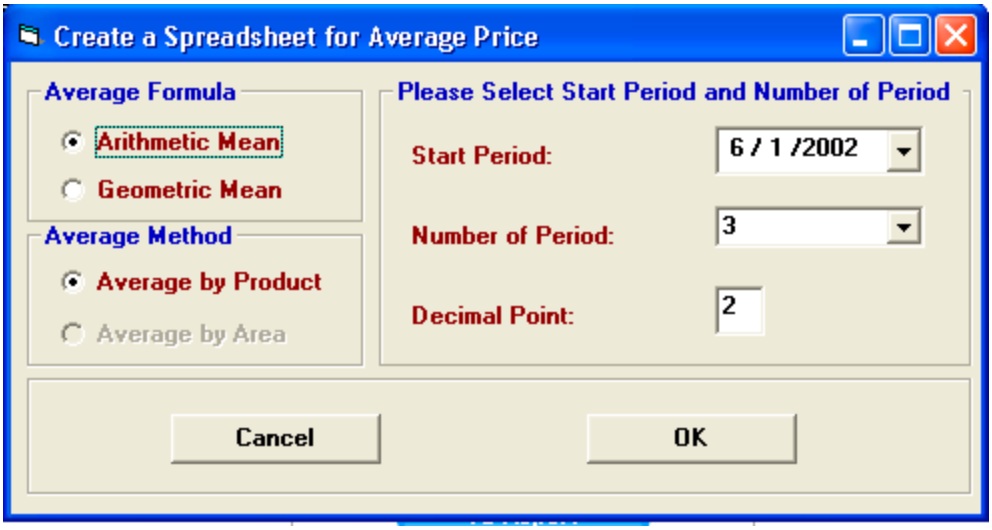
Once each variety obtains its weight from the item it belongs to, product weights can be imputed, as can outlet weights.

Generate Average Price for Varieties

The system can generate a report for average price for all the varieties across the outlets. The user can click **Data**, **Generate Average Price** button as showed below:

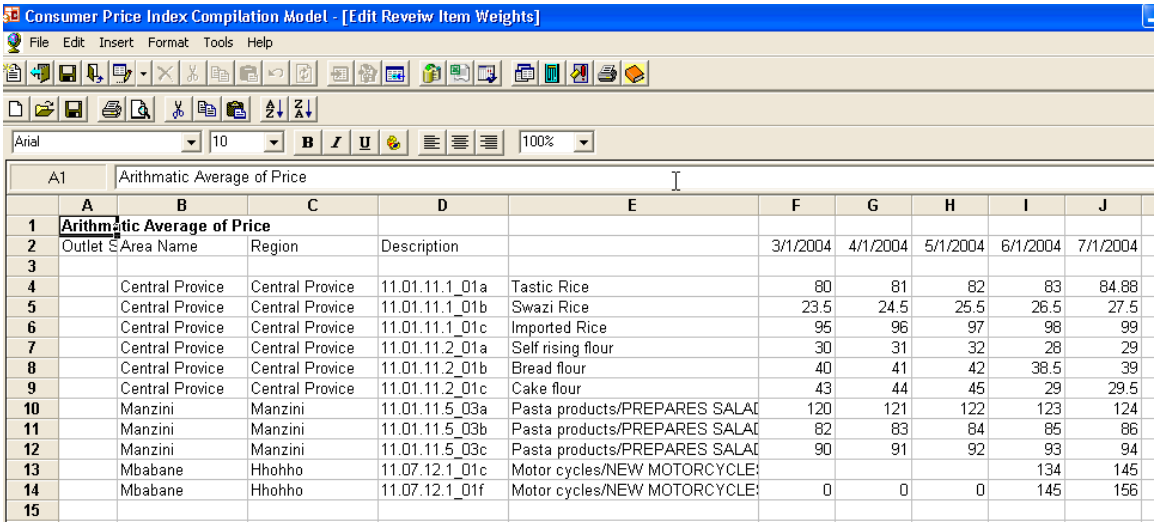


The user will see following screen:



The user can select whether either an arithmetic mean or geometric mean isbe used in calculation. The user should also select **Start Period**, **Number of Period** and **Decimal Point** she would like to have in the report.

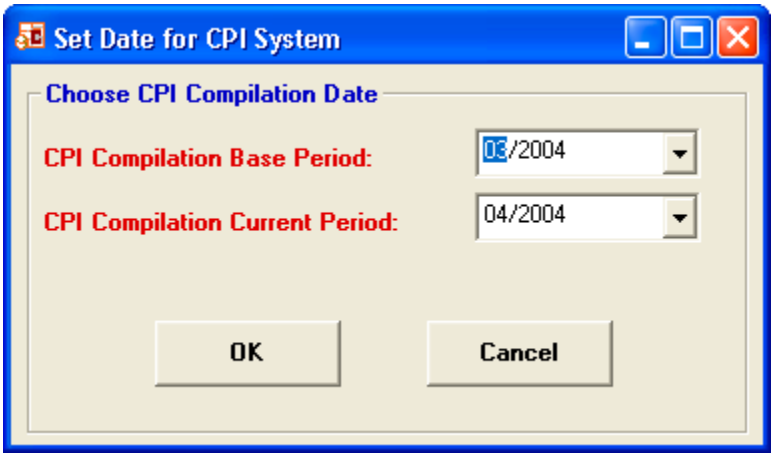
The following report shows the arithmetic average price for different varieties across the outlets in the country for 5 periods.



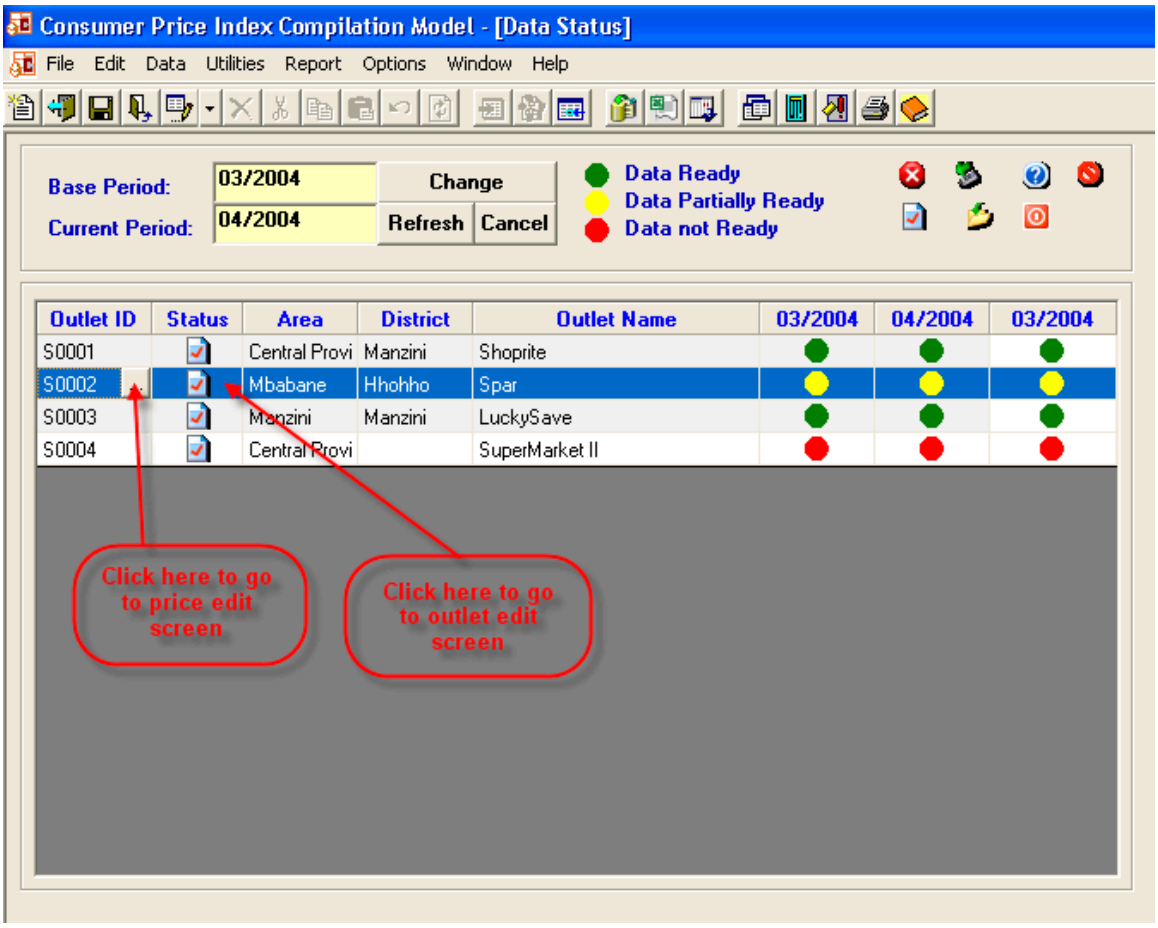
	A	B	C	D	E	F	G	H	I	J
1	Arithmatic Average of Price									
2	Outlet S	Area Name	Region	Description		3/1/2004	4/1/2004	5/1/2004	6/1/2004	7/1/2004
3										
4		Central Provice	Central Provice	11.01.11.1_01a	Tastic Rice	80	81	82	83	84.88
5		Central Provice	Central Provice	11.01.11.1_01b	Swazi Rice	23.5	24.5	25.5	26.5	27.5
6		Central Provice	Central Provice	11.01.11.1_01c	Imported Rice	95	96	97	98	99
7		Central Provice	Central Provice	11.01.11.2_01a	Self rising flour	30	31	32	28	29
8		Central Provice	Central Provice	11.01.11.2_01b	Bread flour	40	41	42	38.5	39
9		Central Provice	Central Provice	11.01.11.2_01c	Cake flour	43	44	45	29	29.5
10		Manzini	Manzini	11.01.11.5_03a	Pasta products/PREPARES SALAI	120	121	122	123	124
11		Manzini	Manzini	11.01.11.5_03b	Pasta products/PREPARES SALAI	82	83	84	85	86
12		Manzini	Manzini	11.01.11.5_03c	Pasta products/PREPARES SALAI	90	91	92	93	94
13		Mbabane	Hhohho	11.07.12.1_01c	Motor cycles/NEW MOTORCYCLE				134	145
14		Mbabane	Hhohho	11.07.12.1_01f	Motor cycles/NEW MOTORCYCLE	0	0	0	145	156
15										

Data Status Indicator and Pre-compile Data Check

Data Status Indicator is a graphical feature provided to help the compiler to examine the data availability and status before undertaking the index compilation. The user has to determine the compilation period by selecting it in the following screen.



Click **OK** button, the user will see following screen:

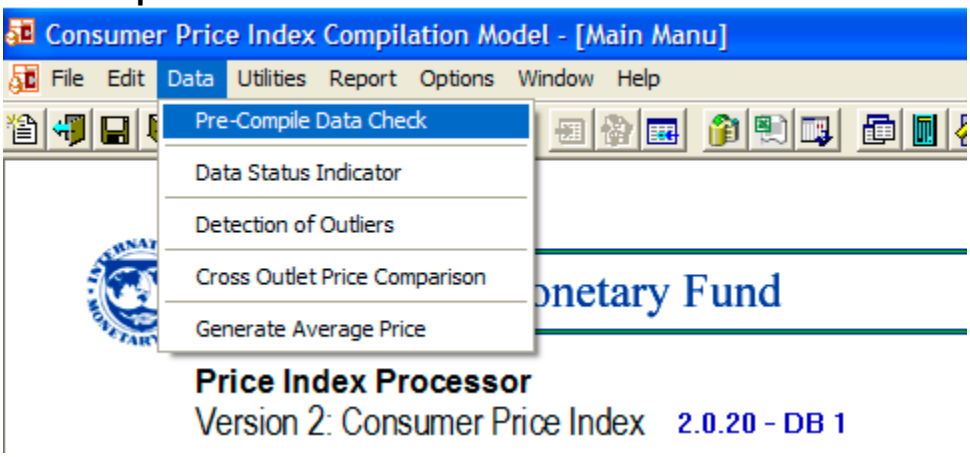


The user will have graphic view of the data status. Green dots indicates that data are available; yellow dots indicate that data are partially available and red dots indicates that data are not available. By clicking the grey button, which will appear when the user single clicks the **Outlet ID** column, the user will enter into price edit screen; from there the user can find the missing price data. If the user click the **Status** column, the user will be able to edit outlet information. This feature provides the user quick access to the data to resolve the data missing problem.

Different icons represent different outlet status. For example, represents **Initiated and reported** ("initiated" means that the outlet has entered the active sample and representative items have been selected for pricing); represents **Refuse to participate**; represents **Out of business**; represents **Could not locate**; represents Not get initiated; represents Resending for initiation and for **No relevant product**.

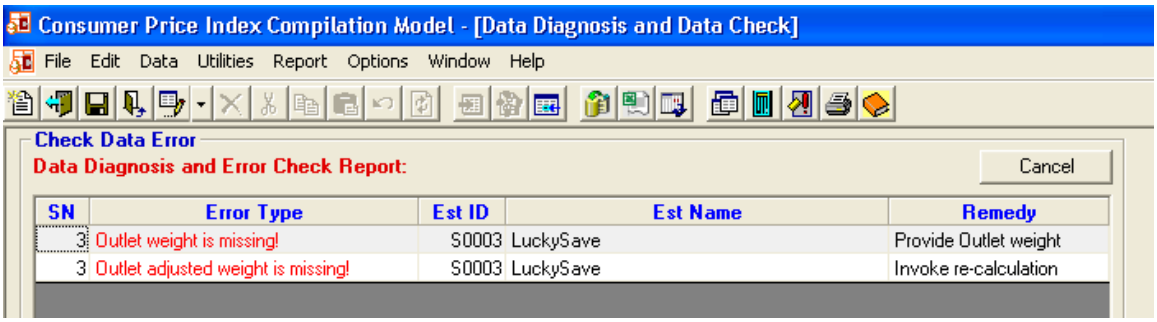
The user can change the base period and current period by clicking the **Change** button.

To examine whether the data is good for compilation, the user can click **Data, Pre-Compile Data Check** button.



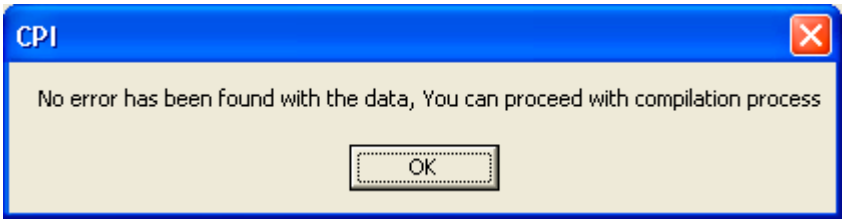
The system is able to identify the following data errors: (1) missing outlet ID; (2) missing weight area; (3) missing product and variety weight information; (4) missing outlet status; (5) missing outlet weights and (6) whether product and variety weights exceed 100%.

For example, the following screen shows that there are two errors for outlet S0003. They are interrelated: the outlet weight is missing and the adjusted weight is missing.

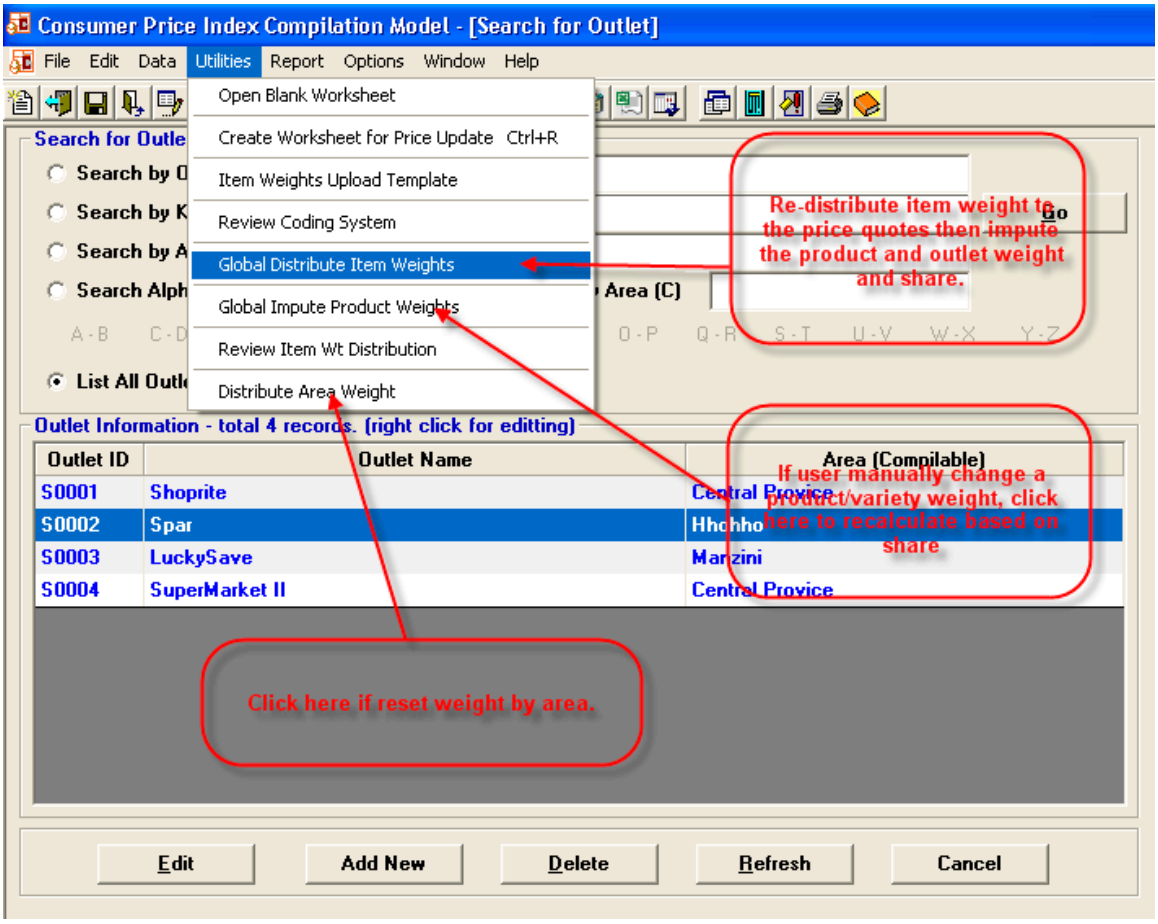


By double-clicking the red colored fonts, the program will bring to the screen where the error can be corrected.

If no error has been found, the user will see following message box:



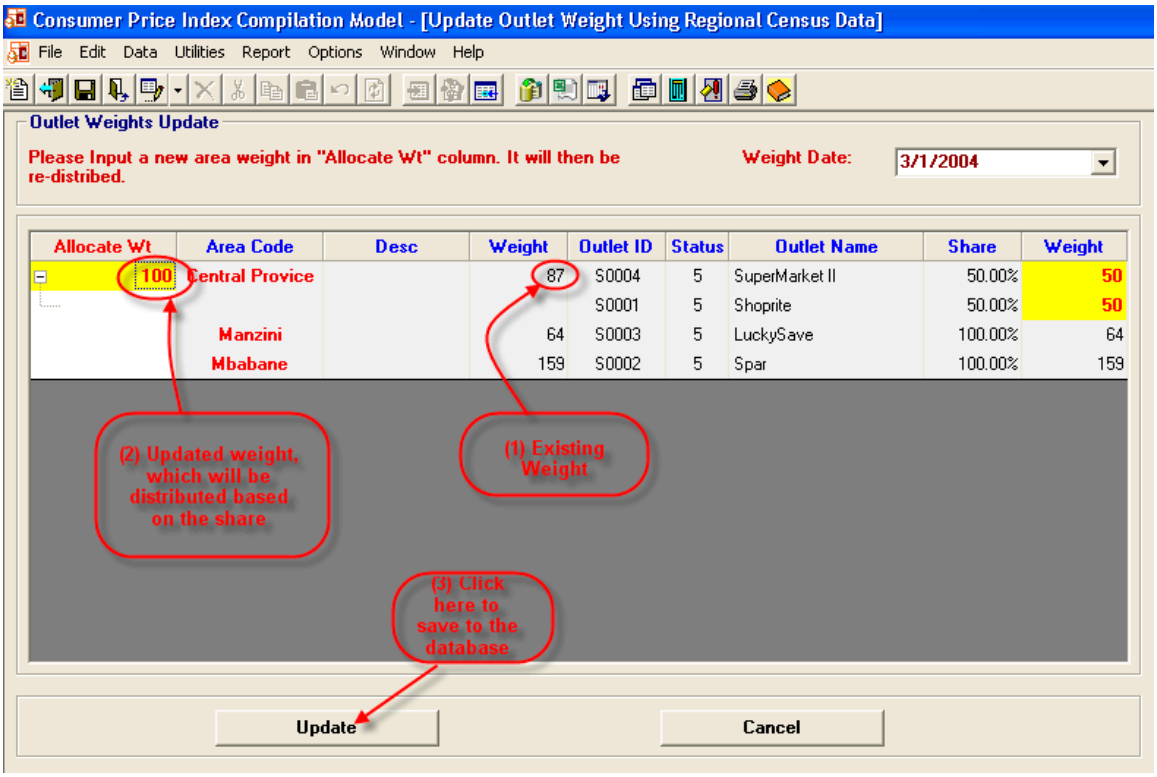
Recalculation of Weights



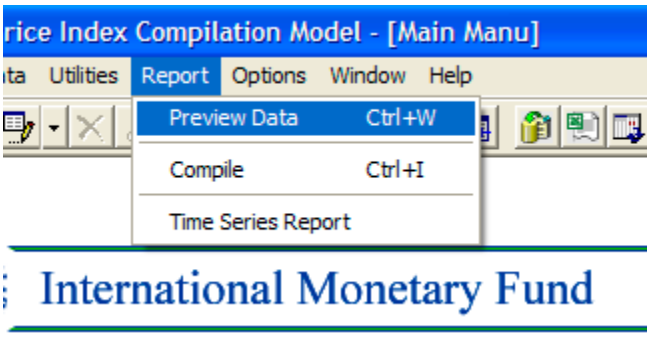
If item weights by area are revised, the user should click **Global Distribute Item Weights** to trigger recalculation process.

It is important that the user can manually define the relative importance of a product or variety by assigning its weight. Clicking **Global Impute Product Weights** will trigger a recalculation based on outlet status and sample group and newly assigned weight. If excess weight is assigned to an outlet, these extra weight will be taken from other outlets. Similarly, if less weight is assigned to a particular outlet, excess weight will be redistributed to the other outlets.

Distribute Area Weights allows the user to undertake weights reset or update. The user can reset an old area weight with a new one, which will be distributed based on the existing share of each outlets. Click **Utilities**, **Global Impute Product Weights** button, newly assigned outlet weights will be redistributed to the product and varieties below.

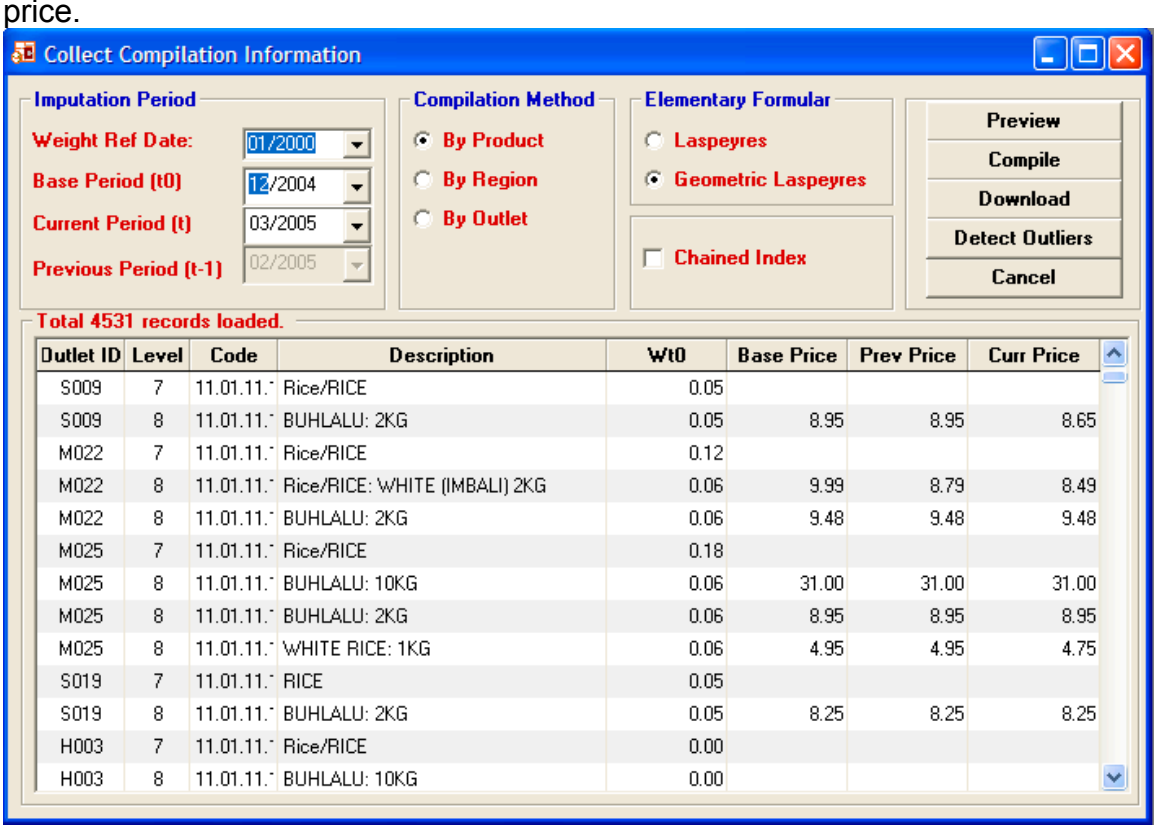


Pre-compilation Preview



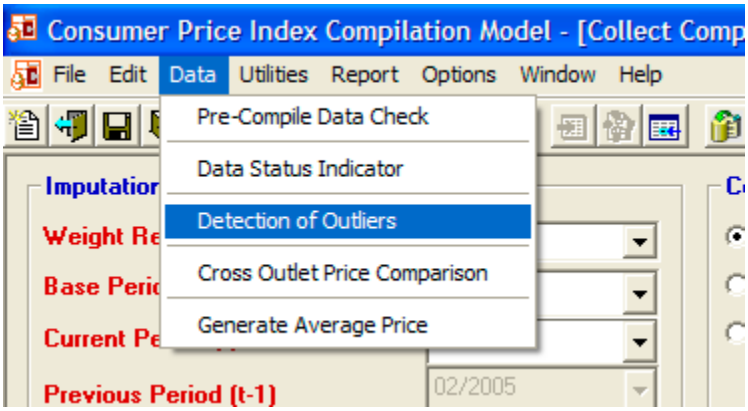
Click **Report**, **Preview Data** button at toolbar to preview the data before compilation.

The user can select Base Period, Current Period. Wt0 column indicates weights data followed by base period price, previous period price and current period price.

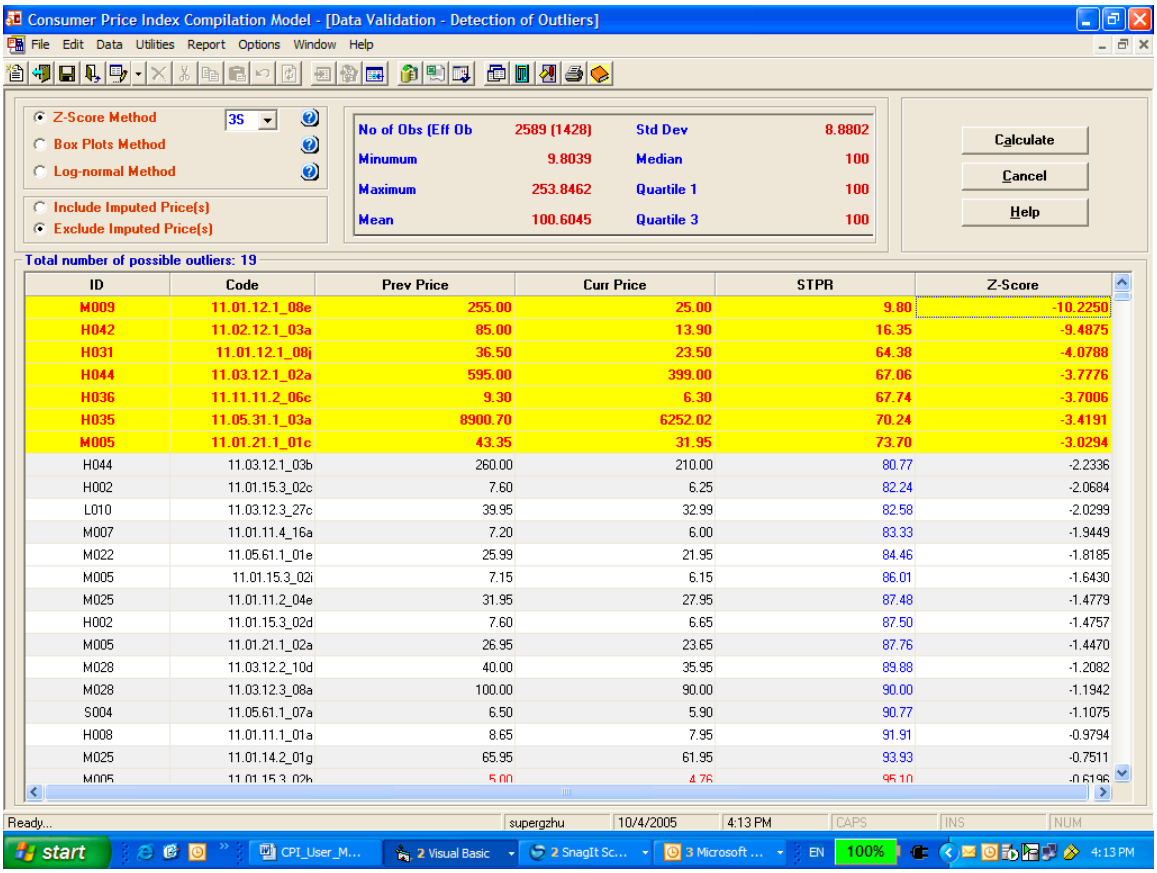


Detecting Outliers

To minimize data entry error, the system has implemented independent Detect Outliers module on checking the outliers of short term price relatives (one month price changes). For the methodology, please refer to Chapter 3, Detection of Outliers. Click **Data, Detection Outliers** button.



Wait for preview data to be loaded into the table. Click **Detect Outliers** button. Click Z-Score fixed row to re-sort information.

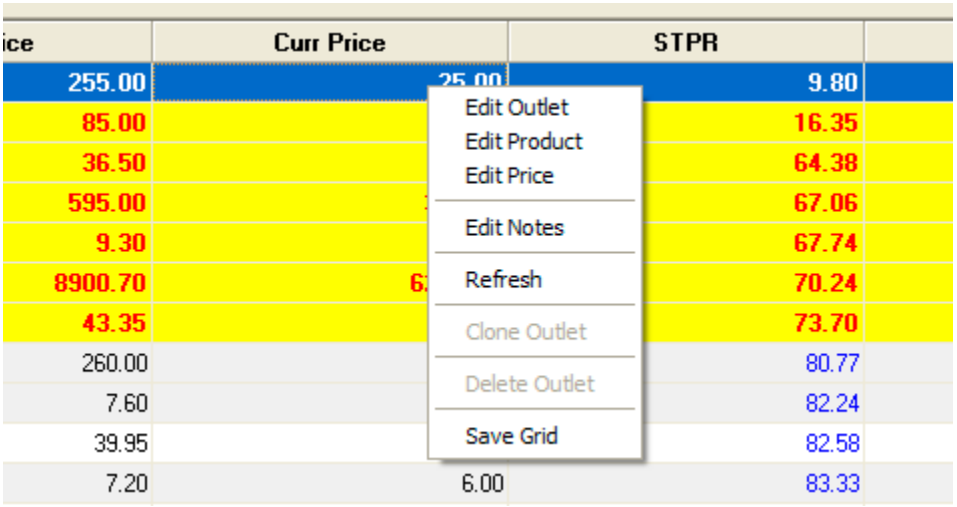


The default method is the Z-score test; the user can also select Box Plots Method or Log-normal Method. The system presents a summary statistical table with the information on short term price relatives including: (1) Number of observations; (2) Minimum value; (3) Maximum value; (4) Mean; (5) Standard deviation; (6) Median; (7) Quartile 1; (8) Quartile 3. The user can select whether

imputed value should be excluded from the imputation. Those outliers will be showed in red with yellow background as showed in the above screen capture.

The first 7 rows are detected to be outliers. Apparently, there is the very strong possibility of a data entry error in product 11.01.12.1_08e in which previous price is 255.00 and current price is 25.00.

To fix the problem, select the record, right click to start the menu, then click Edit Price button.

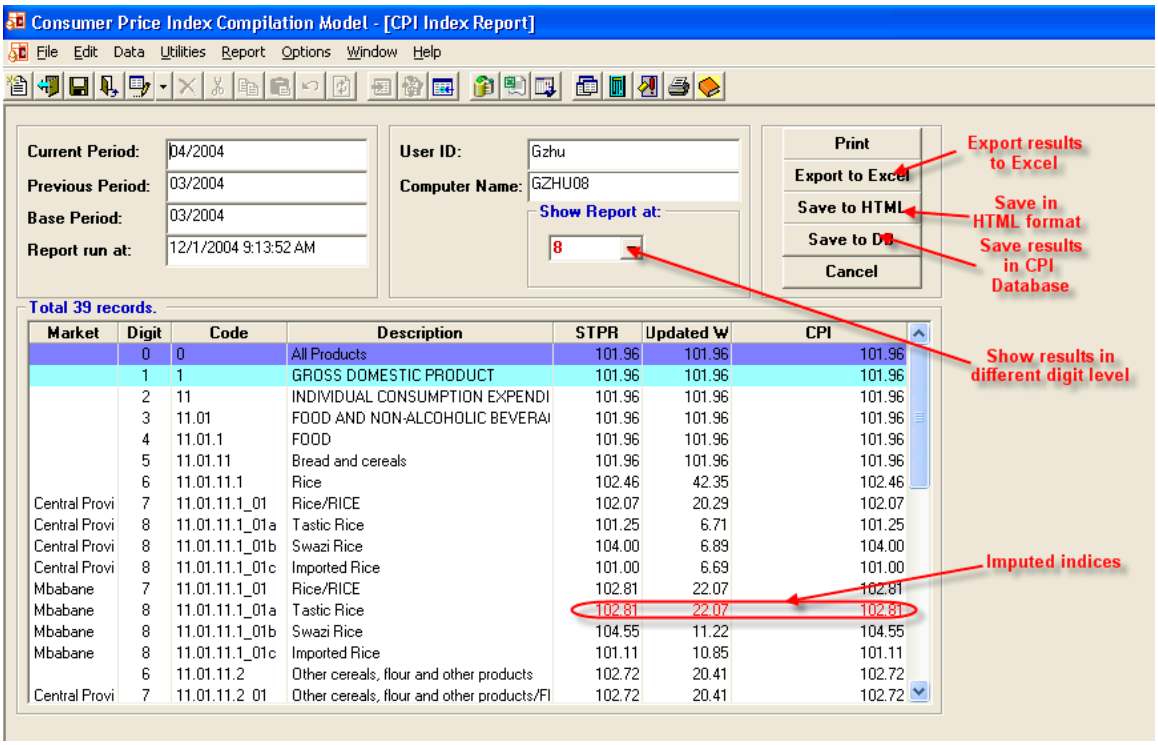


Index Compilation

The user defines the compilation method by choosing one of the options: **By Product**, **By Region** or **By Outlet**. The equally-weighted elementary level compilation always uses the geometric mean. The user can also choose either Laspeyres-type or Geometric Laspeyres-type formula in the compilation at the weighted level.

If **Modified/Two-stage** is checked, the system uses the previous period Long Term Price Relatives ($LTPR_{t-1}$), which is price increase from period 0 to period t-1, in the calculation (please refer the Chapter 3, Methodology for the details of formula). Alternatively, base period price P_{t0}, previous period price P_{t-1} and current period price, P_t will be used for the imputation process.

Imputation results are shown in following screen:

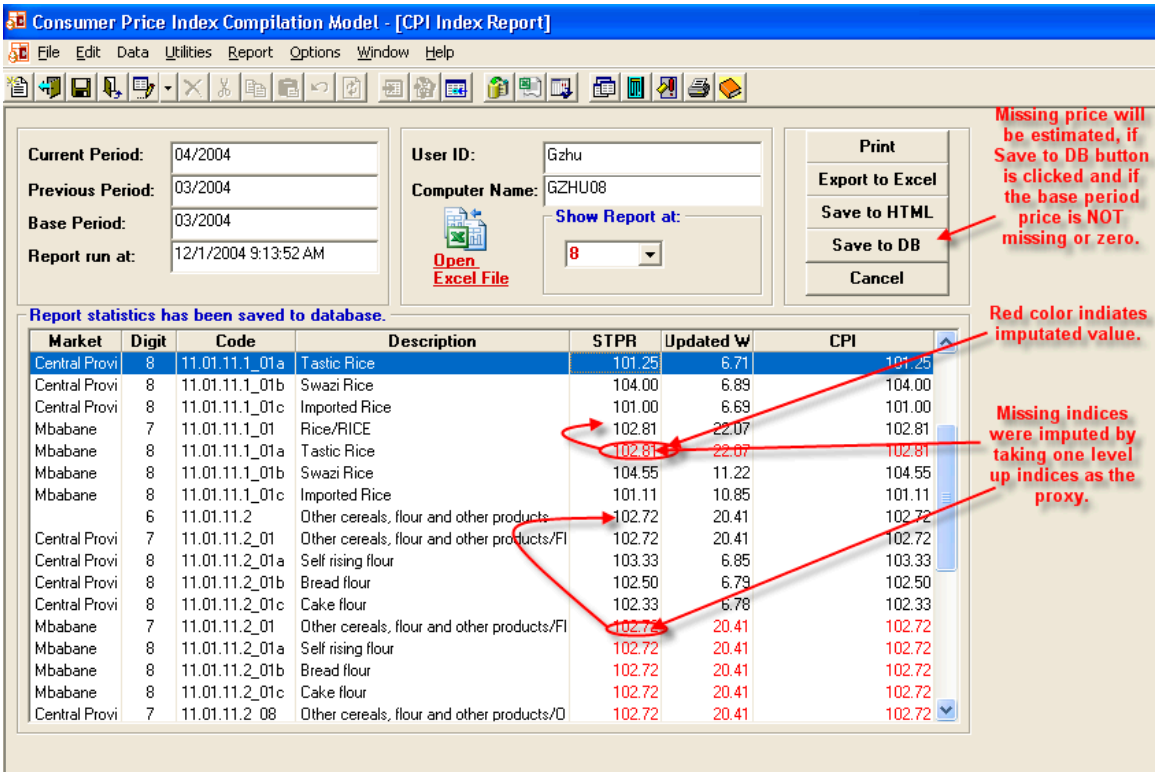


The system generates three sets of results: (1) **STPR**, Short Term Price Relatives, which is the index of consumer price from last period to current period. (2) **CPI**, which is also Long Term Price Index, indicates accumulated price increase since period 0. (3) **Updated W** is updated cost weight.

CPI indices will be plugged into the imputation as $LTPR_{t-1}$ next period if **Two-Stage Index** box is checked.

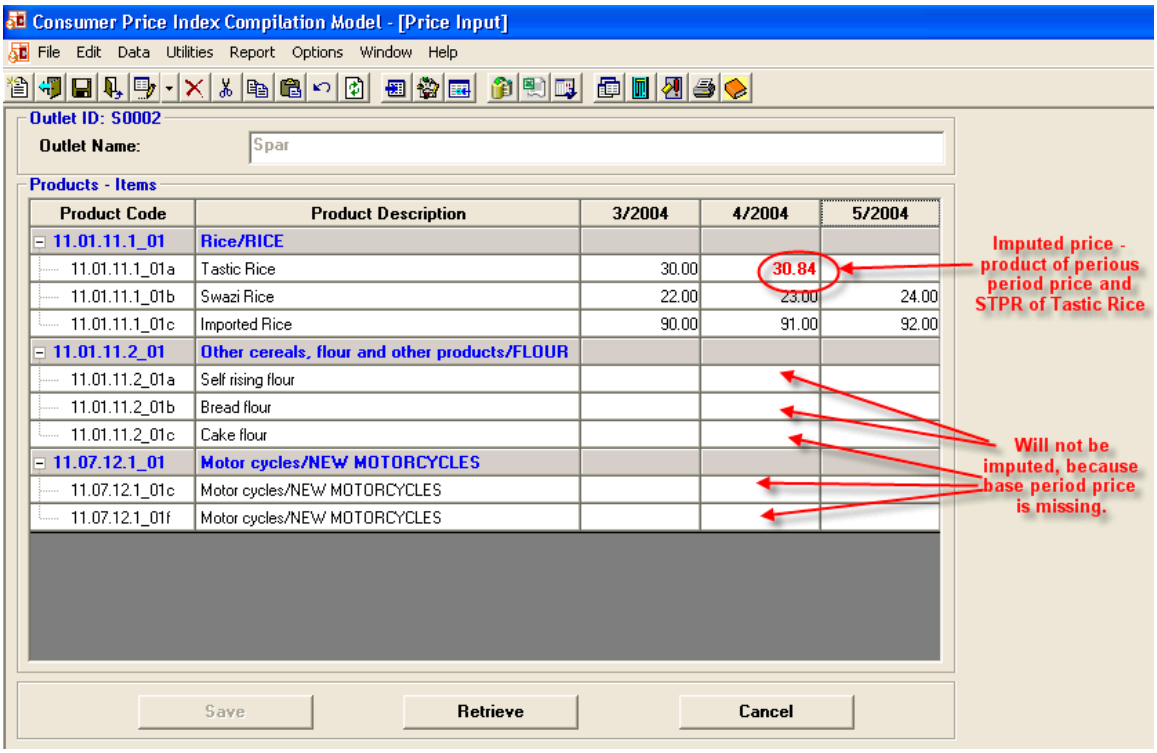
Imputation results can be exported to the Excel easily. The user can click **Export to Excel**, Export to Excel button, then an Excel icon will appear in the screen.

Click the link, imputation results will be exported to the Excel spreadsheet.



Missing price indices are imputed by taking the indices from their parent group. If parent group indices are missing, the system has to go one level up again, until it finds a calculated index. Imputed indices are indicated in red. When the indices are saved to the database, the missing prices are automatically imputed and stored in the database with a flag. Missing prices will also be shown in the price edit screen in red.

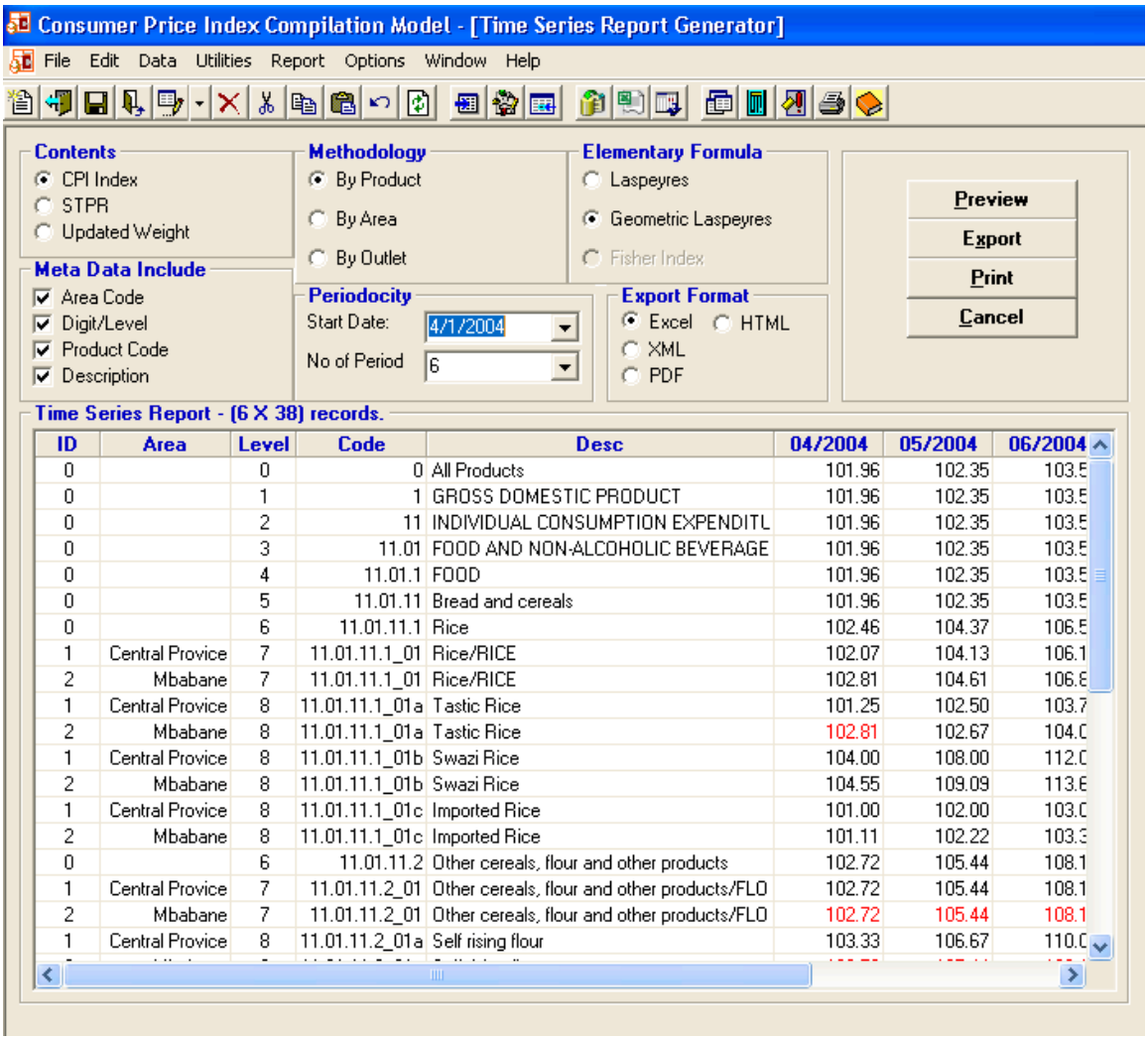
Current period missing prices are imputed by multiplying the last period price by the STPR of that variety. If the last period's price is missing, then the imputed price will take the product of the base period price and overall CPI index for the higher product group.



Time Series Report

Indices in time series format (one index for multiple periods, each column indicates a period) can be generated. Details of such report are showed in following screen.

Time series presentation will give the user one index at a time for the multiple periods. The user can select which, if any, meta data they would like to have meta data appear in their report. The time series report can be readily exported to an Excel spreadsheet.



References

- Consumer Price Index Manual: Theory and Practice (2004), *International Labor Organization*. http://www.ilo.org/public/english/bureau/stat/guides/cpi/ and at http://www.imf.org/external/pubs/cat/longres.cfm?sk=17165.0

- COICOP, Detailed structure and explanatory notes, Statistics Division, UN. http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=5&Lg=1

Index