

COMPOSITIONAL DATA PACKAGE USER'S GUIDE

February of 2007

CONTENTS

INTRODUCTION	2
FEATURES	2
1. Transformation Menu	5
1.1. Transformations. Unconstrain/Basis	5
1.2. Transformations. Raw-ALR	5
1.3. Transformations. Raw-CLR	6
1.4. Transformations. Raw-ILR	7
2. Operations Menu	8
2.1 Operations. Perturbation	8
2.2 Operations. Power Transformation	8
2.3 Operations. Centering	9
2.4 Operations. Standardization	10
2.5 Operations. Amalgamation	10
2.6 Operations. Subcomposition/Closure	11
2.7 Operations. Rounded Zero Replacement	11
3. Graphs Menu	13
3.1 Graphs. Ternary Diagram	13
3.2 Graphs. ALR Plot	13
3.3 Graphs. CLR Plot	14
3.4 Graphs. Biplot	14
3.5 Graphs. Principal Components	14
3.6 Graphs. ALN Predictive Regions	15
3.7 Graphs. ALN Confidence Regions	16
3.8 Graphs. Balance-Dendrogram	16
5. Descriptive Statistics Menu	20
4.1. Descriptive Statistics. Summary	20
4.2. Descriptive Statistics. Center	21
4.3. Descriptive Statistics. Total Variance	21
4.4. Descriptive Statistics. Variation Matrix	22
4.5. Descriptive Statistics. Atypicality Indices	22
5. Analysis Menu	23
5.1. Analysis. Logistic Normality Test	23
6. Preferences Menu	24
6.1. Preferences. Screen Size	24
6.2. Preferences. Sum Constraint	24
INSTALLATION	25
UNINSTALLATION	31

INTRODUCTION

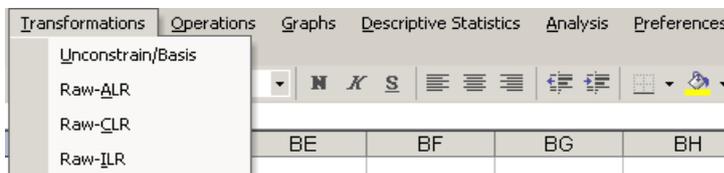
The software that we present at this time is still under construction. The idea is to build a user-friendly application to be used in all fields that need to operate with compositional data.

To use this application you need the Software EXCEL® and to use it it is necessary to open the file CoDaPack.xls with the associated macros and it works using menus that appears in every sheet. When You activate the menus a Visualbasic routine is started and the final results appears in the active sheet when you work, or in a new graphical windows that appears under the active sheet.

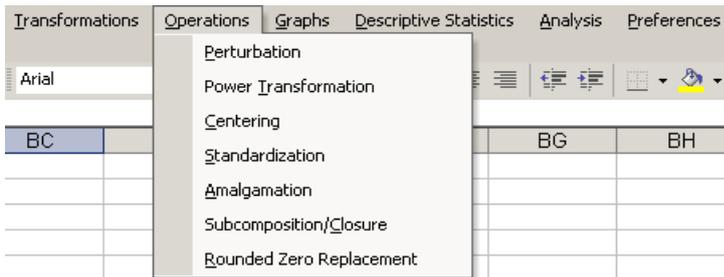
This document has two parts. The first one contains the different features of the application and the second part the instructions to install this software.

FEATURES

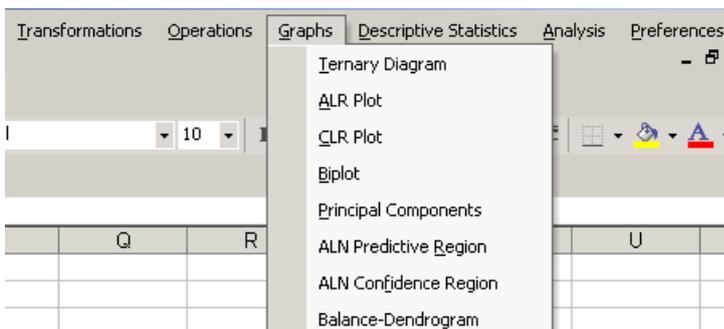
The application has six main menus: *Transformations*, *Operations*, *Graphs*, *Descriptive Statistics*, *Analysis* and *Preferences*. The first one transforms the data from the simplex to the real space or vice versa.



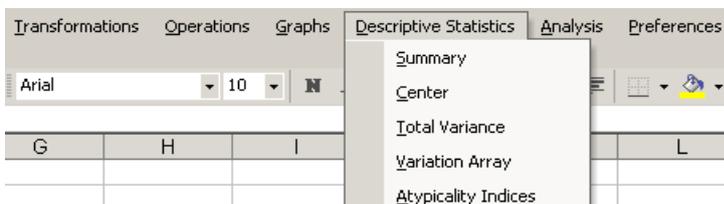
The second one performs some operations of the data in the simplex.



The third one makes some graphical representation in the simplex or in real space.



The fourth one performs some Descriptive Statistics.



The fifth one performs some analysis.



And the last customizes the size of the graphs depending of the screen.



SOME GENERAL CONSIDERATIONS

The web site <http://ima.udg.es/~thio/#Compositional Data Package> contains this freeware package and to install it the user only needs to have Excel installed on his computer.

To use CoDaPack, one has to access Excel and introduce the data in a standard spreadsheet. The observations must be in rows and the variables in columns, and the first row of each column can be used to label the variables or it has to rest blank.

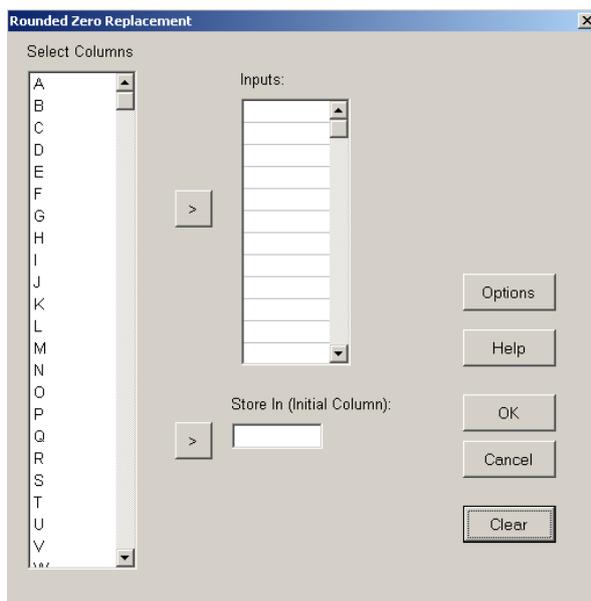
Using menus, one can execute macros to return the numerical results on the same sheet and graphical outputs that appear in independent windows inside Excel. In the present version there exist 6 menus with a total of 27 macros. Each macro asks to the user, which are the data and where to put the results (if there are numerical results). Some of the macros, specially those with graphical output, have an option button to modify the default values.

In Order to make easy the utilisation of CoDaPack, when the user executes a routine it appears on the dialog screen, as a default values, those that he has used the last time.

To make it possible, CoDaPack creates some files with extension "cdp". This files contains all the information needed and are located on the user profile directory, usually in C:\DOCUMENTS AND SETTINGS\....

In order to clear the default values the user can click on the "Clear" button and all the fields of the dialog screen are emptied.

To execute a routine of CoDaPack, user has to select it from the menus with only one mouse ckick A new window appears – standard for all CoDaPack routines, asking which columns of the active sheet to select and where to put the results.



Its left side contains the *Select Columns* structure, the middle the *Inputs* structure and the *Store In (Initial Column)* box and on the right there are buttons: *Help*, *OK*, *Cancel*, *Clear* and in some routines the *Options*. Between the left and the middle part there are two arrows to pass information between them.

When this window is opened the *Select Columns* list contains the first row of each column of the Excel sheet or, for a column without label, the standard letter that identifies the columns of any Excel sheet. Also, if the routine has been executed before, this window contains in *Inputs*, *Store In (Initial Column)* and *Options* the values used on the last execution.

First of all the user should select the parts to be used in the routine, in this case the parts involved in the Rounded Zero Replacement. To do that the user has to mark a row of the *Select Column* list with a single click and then click on the arrow. After that the name of the selected column appears on the middle, inside *Input* structure. The user should repeat this operation in order to select all the parts involved in the routine. If the user wants to unselect a part of *Inputs* structure he has to mark this part inside *Input* structure and click the arrow which now indicates the opposite direction.

Finally the user should select, following the same procedure than before, where to put the results. In order to avoid unnecessary selections the user only needs to select the first column where to put the results and CoDaPack uses the columns it needs starting with the selected column.

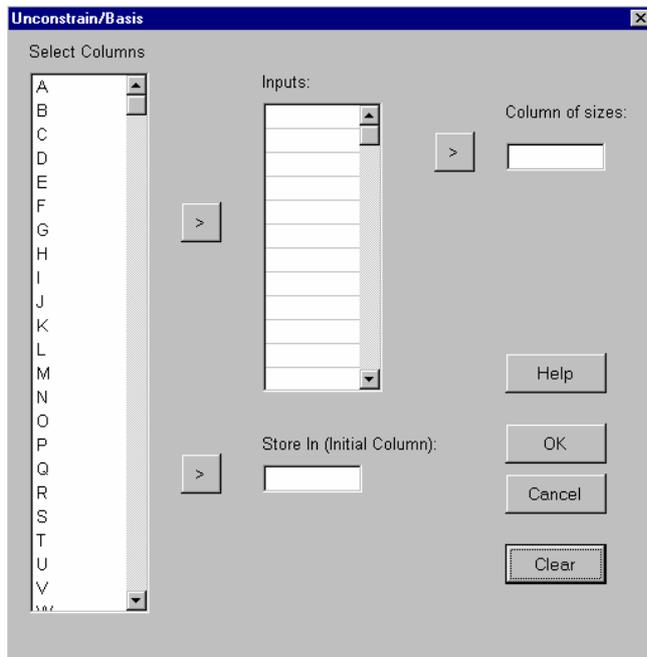
VERY IMPORTANT: At this time, to use the program, it is necessary to copy the data to analyze into the file Codapack.xls.

1. Transformations menu

This menu performs transformations of the data from the simplex to the real space or vice versa.

1.1 Transformations. Unconstrain/Basis

This routine returns the data set unconstrained, that is, for each constrained observation \mathbf{x} , it returns its unconstrained $\mathbf{y} = [x_1w, \dots, x_dw]$, where w is the size or weight of the observation. With this feature the data is transformed from simplex to the real space.



The user has to select the columns to unconstrain, where to put the results and also it has to indicate the column of sizes or weights.

1.2 Transformations. Raw-ALR

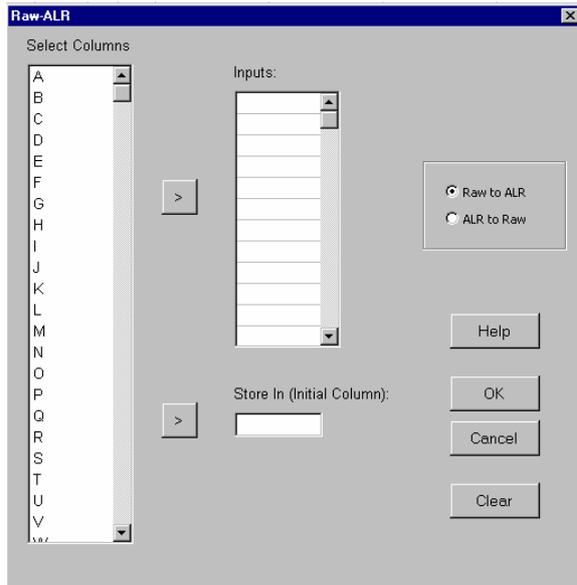
With this feature the data is transformed from simplex to real space according to the additive logratio transformation (alr) or its inverse transformation, that is from real space to simplex, with the generalised additive logistic transformation (agl).

$$\mathbf{y} = \text{alr}(\mathbf{x}) = \left[\ln \frac{x_1}{x_D}, \dots, \ln \frac{x_{D-1}}{x_D} \right], \text{ where } \mathbf{y} \in \mathfrak{R}^{D-1}, \text{ the real space with } D-1 \text{ dimension, and}$$

$$\mathbf{x} = \text{agl}(\mathbf{y}) = \left[\frac{\exp(y_1)}{1 + \sum_{i=1}^{D-1} \exp(y_i)}, \dots, \frac{\exp(y_{D-1})}{1 + \sum_{i=1}^{D-1} \exp(y_i)}, 1 - x_1 - \dots - x_{D-1} \right]$$

Division in the alr transformation is performed with the last component according to the sequence selected by the user.

The user has to select the columns to transform, where to put the results and also it has to indicate the direction of the transformation.



1.3 Transformations. Raw-CLR

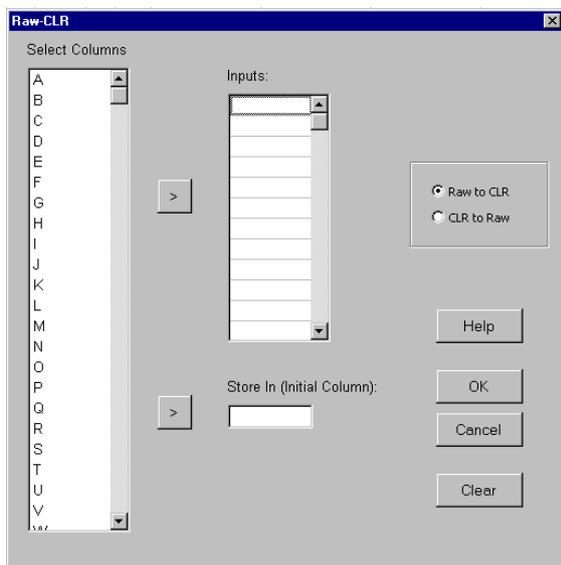
With this feature the data is transformed from simplex to real space according to the centred logratio transformation (clr) or its inverse transformation, that is from real space to simplex, with its inverse (clr⁻¹).

$\mathbf{y} = \text{clr}(\mathbf{x}) = \ln \frac{\mathbf{x}}{g_D(\mathbf{x})}$, where $\mathbf{y} \in \mathcal{R}^D$ and $g_d(\mathbf{x})$ is the geometric mean

$\left(\prod_{i=1}^D x_i \right)^{1/D} = \exp \left[\frac{1}{D} \sum_{i=1}^D \ln x_i \right]$ of \mathbf{x} , and the inverse transformation is

$$\mathbf{x} = \text{clr}^{-1}(\mathbf{y}) = \left[\frac{\exp(y_1)}{\sum_{i=1}^D \exp(y_i)}, \dots, \frac{\exp(y_D)}{\sum_{i=1}^D \exp(y_i)} \right].$$

The user has to select the columns to transform, where to put the results and also it has to indicate the direction of the transformation.



1.4 Transformations. Raw-ILR

With this feature the data is transformed from simplex to real space according to the isometric logratio transformation (ilr) or its inverse transformation, that is from real space to simplex, with its inverse (ilr⁻¹).

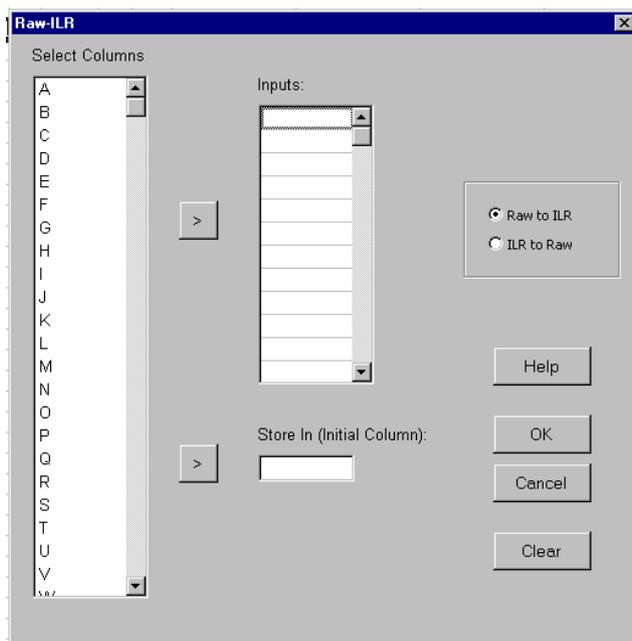
$$\mathbf{y} = \text{ilr}(\mathbf{x}) = (y_1, \dots, y_{D-1}) \in \mathfrak{R}^{D-1}$$

$$\text{where } y_i = \frac{1}{\sqrt{i(i+1)}} \ln \left(\frac{\prod_{j=1}^i x_j}{(x_{i+1})^i} \right) \quad (i = 1, \dots, D-1), \text{ and}$$

$$\mathbf{x} = \text{ilr}^{-1}(\mathbf{y}) = (x_1, \dots, x_D) \in \mathfrak{R}^D$$

$$\text{where } x_i = \frac{f(i-1)}{\sum_{j=0}^{D-1} f(j)} \quad (i = 1, \dots, D), \quad f(i) = \left(\frac{1}{f(i-1)} \exp(\sqrt{i(i+1)} y_i) \right)^{-1/i} \quad \text{and } f(0) = 1.$$

The user has to select the columns to transform, where to put the results and also it has to indicate the direction of the transformation.



2. Operations menu

This menu performs operations inside the simplex. That is, operations where input data and output results are in the simplex.

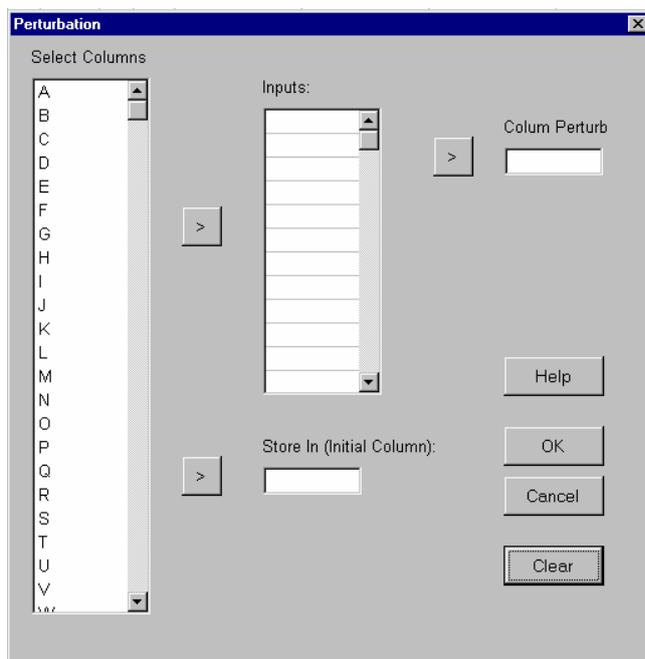
2.1 Operations. Perturbation

With this feature a vector perturbs the data.

Returns a D -composition $\mathbf{y} = \mathbf{p} \oplus \mathbf{x} = C[p_1x_1, \dots, p_Dx_D]$, where C stands for the closure operation

$$C[x_1, x_2, \dots, x_D] = \left[\frac{x_1}{\sum_{i=1}^D x_i}, \frac{x_2}{\sum_{i=1}^D x_i}, \dots, \frac{x_D}{\sum_{i=1}^D x_i} \right], \text{ and } \mathbf{p} \text{ is a given } D\text{-composition.}$$

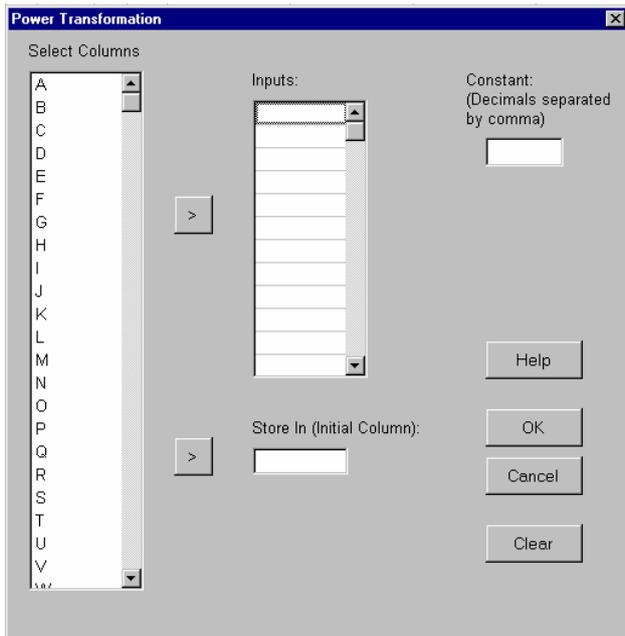
The user has to select the columns to perturb, where to put the results and also it has to indicate the column that contains the vector \mathbf{p} .



2.2 Operations. Power Transformation

This feature applies a Power Transformation to the data. For $a \in \mathfrak{R}$, the power transformation returns $a \otimes \mathbf{x} = C[x_1^a, \dots, x_D^a]$.

The user has to select the columns to power transform, where to put the results and also it has to indicate the constant of the operation (decimals separated by comma).



2.3 Operations. Centering

3) Centering:

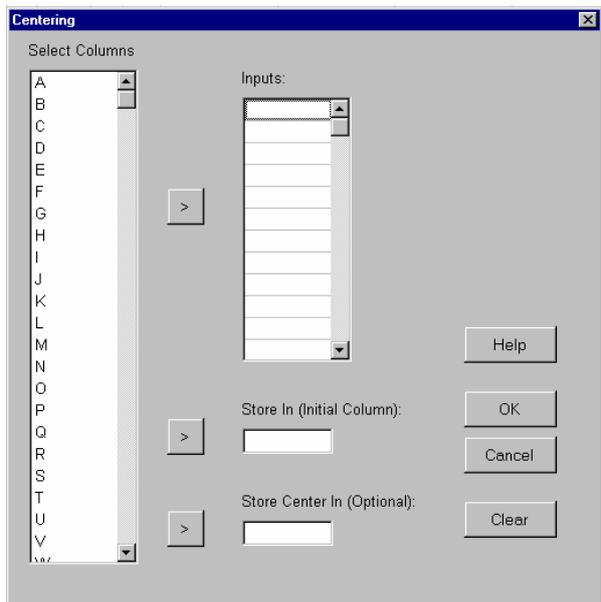
With this feature the data is centered, that is, it is perturbed by the geometrical mean of the data. This

routine centres the data set, that is, it returns the data set \mathbf{Y} formed by the D -compositions

$$\mathbf{y} = \mathbf{g}_N(\mathbf{X})^{-1} \oplus \mathbf{X}, \text{ where } \mathbf{g}_N(x) = \mathbf{g}_N(\mathbf{X}) = \left[\left(\prod_{k=1}^N x_{k1} \right)^{1/D}, \dots, \left(\prod_{k=1}^N x_{kD} \right)^{1/N} \right]$$

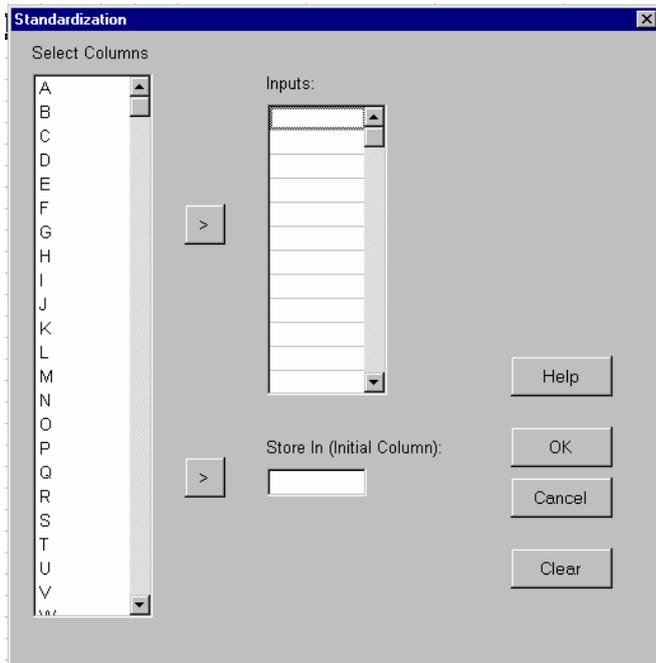
is the vector of geometric means of the data set \mathbf{X} . Thus, the centre of the set \mathbf{Y} is \mathbf{e} , the barycentre of the simplex, e.g. for $D = 3$ the geometric centre of a ternary diagram is [0.333, 0.333, 0.333].

The user has to select the columns to center, where to put the results and optionally the column to put the geometrical mean.



2.4 Operations. Standardization

This feature standardizes the data, that is, centers the data and makes the total variance equal to one. This routine returns a sample of D -compositions Y , centred at e and with unit total variance.

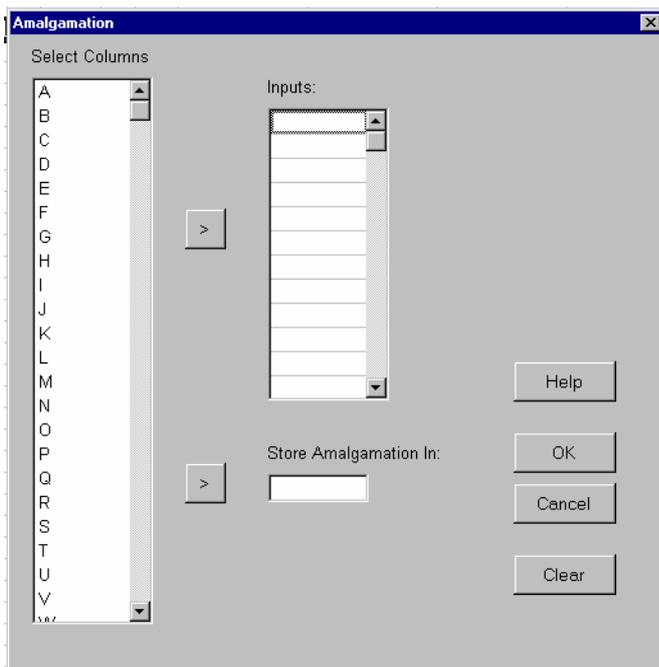


The user has to select the columns to standardize and where to put the results.

2.4 Operations. Amalgamation

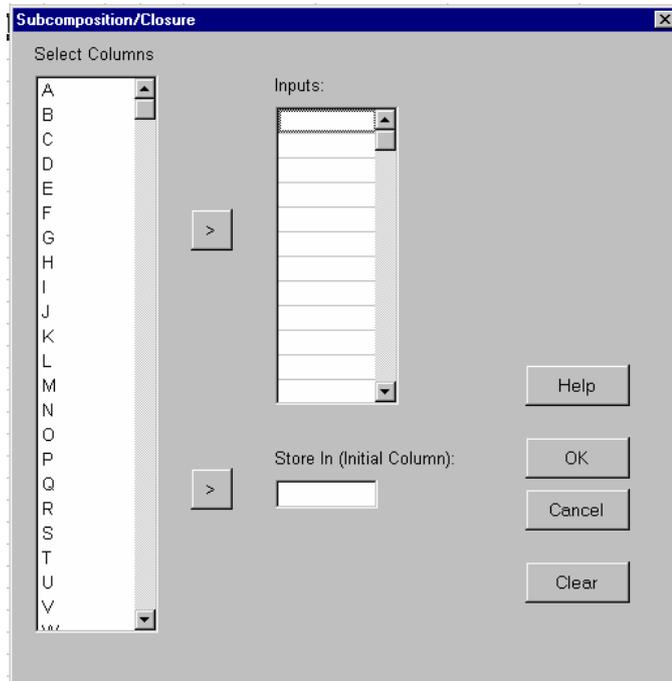
This feature amalgamates some columns of the data. The result of the amalgamation of some of the parts of a D -composition selected by the user is the sum of those parts.

The user has to select the columns to amalgamate and the column where to put the result.



2.6 Operations. Subcomposition/Closure

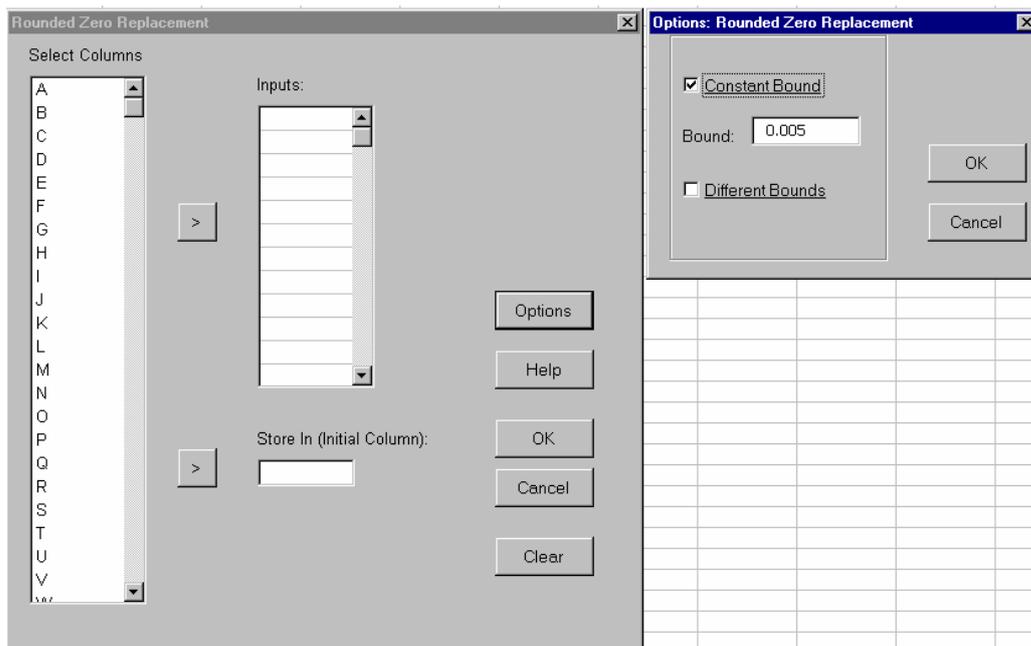
With this feature the data is closed. This routine closes, *i.e.* reproportions, the data, that is, returns $\mathbf{Y} = C(\mathbf{X})$. If we select S parts ($S < D$) a subcomposition with S -parts is obtained.



The user has to select the columns to closure and where to put the results.

2.7 Operations. Rounded Zero Replacement

This feature applies a transformation to the data to avoid the zeros.



Rounded Zero Replacement consist on to substitute an observation \mathbf{x} , with zeros in some parts, by an observation \mathbf{y} using the expression:

$$y_i = \begin{cases} \delta_i & \text{if } x_i = 0 \\ x_i \left(1 - \sum_{x_j=0} \delta_j \right) & \text{if } x_i > 0 \end{cases}$$

where δ_i is the replacement value for the i -th part defined by the user.

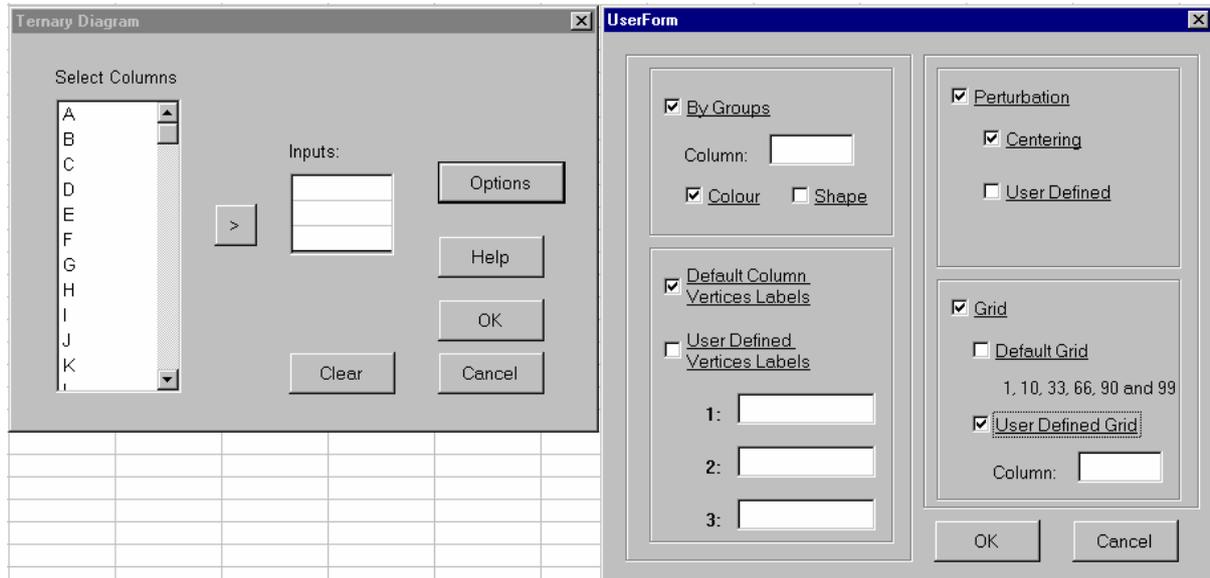
The default constant δ_i is 0.005 but the user can define another constant or a column of constants that has to contain as constants as parts of the composition.

The user has to select the input columns and where to put the results.

3. Graphs menu

This menu enables the user to create two dimensional graphs. The user can customize the appearance of each graph and, in some cases, plot the observations in the graph according to a previous classification:

3.1 Graphs. Ternary Diagram



This feature displays a ternary diagram of 3 selected parts.

There are four options to modify the appearance of the graph:

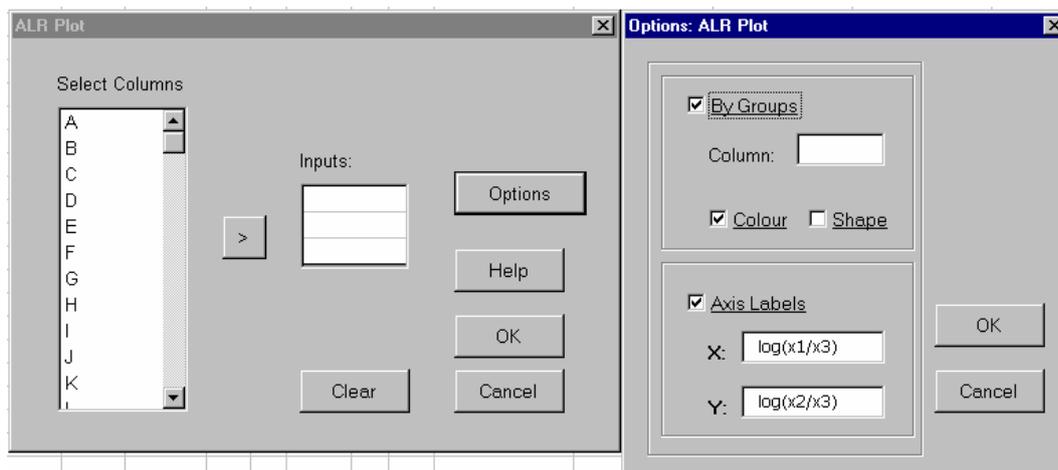
- 1) Differentiate, by colour or by shape, each point depending on a previous classification,
- 2) Label the vertices of the triangle (the default labels are the part names),
- 3) Perturb the data with the inverse of the centre (centring) or with a given vector,
- 4) Display a reference grid of values. The default values of the grid are **1, 10, 33, 66, 90 and 99** but the user can define other values in a column

3.2 Graphs. ALR Plot

This feature displays a plot according to the ALR transformation of the 3 column selected.

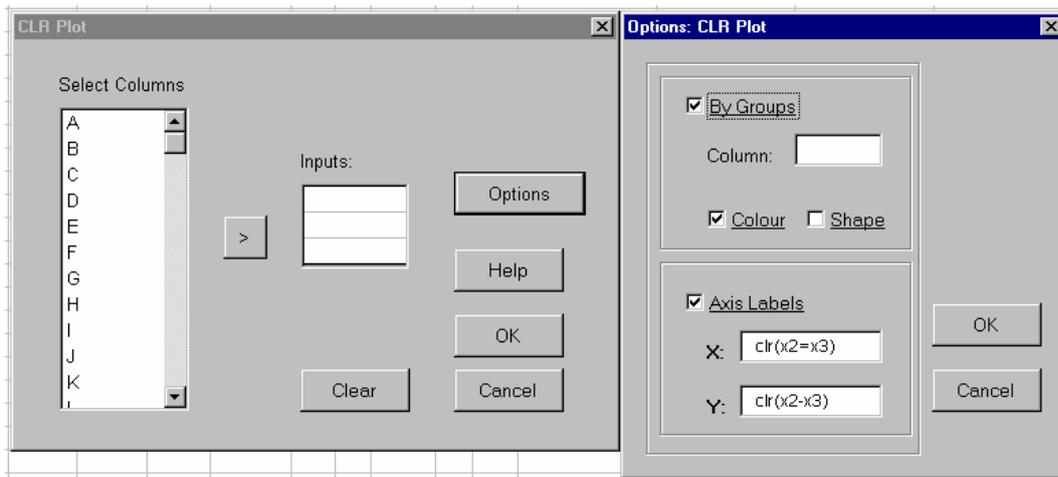
There are two options to modify the appearance of the graph:

- 1) differentiate, by colour or by shape, each point depending on a previous classification
- 2) To label the axis (The default labels are **$\log(x1/x3)$** and **$\log(x2/x3)$**).



3.3 Graphs. CLR Plot

This feature displays a plot according to the centred logratio transformation (clr) of 3 selected parts.



There are two options to modify the appearance of the graph:

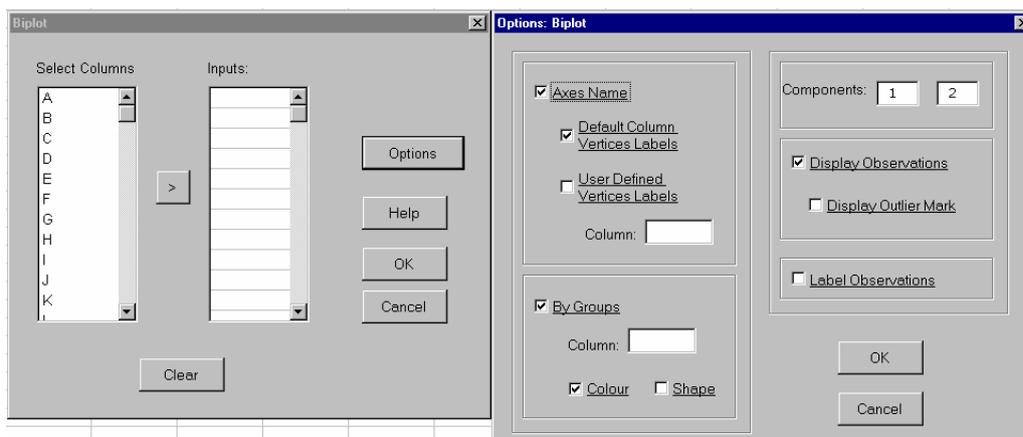
- 1) differentiate, by colour or by shape, each point depending on a previous classification
- 2) Label the axis (The default labels are ILR1 and ILR2).

3.4 Graphs. Biplot

This feature performs a Biplot of selected parts.

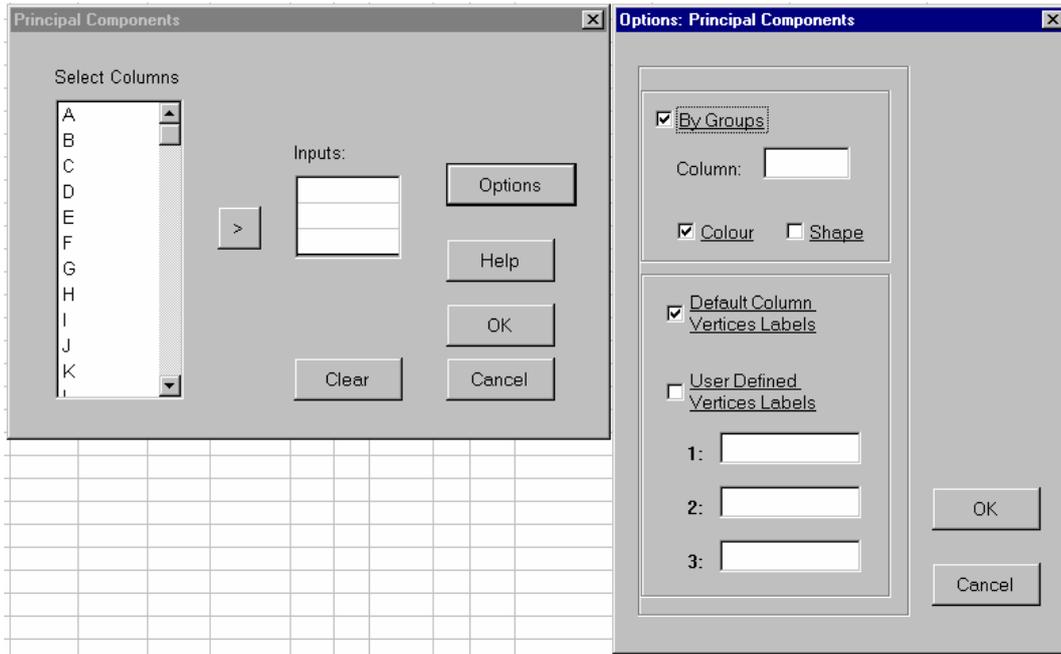
There are six options to modify the appearance of the graph:

- 1) indicate a column with the labels of the axes,
- 2) differentiate, by colour or by shape, each point depending on a previous classification,
- 3) choose the factor plane indicating which parts to display,
- 4) label the observations (the default is not label),
- 5) display or not the observations (the default is yes), and
- 6) display with a different mark the observations that are outliers (the default is not mark).



3.5 Graphs. Principal Components

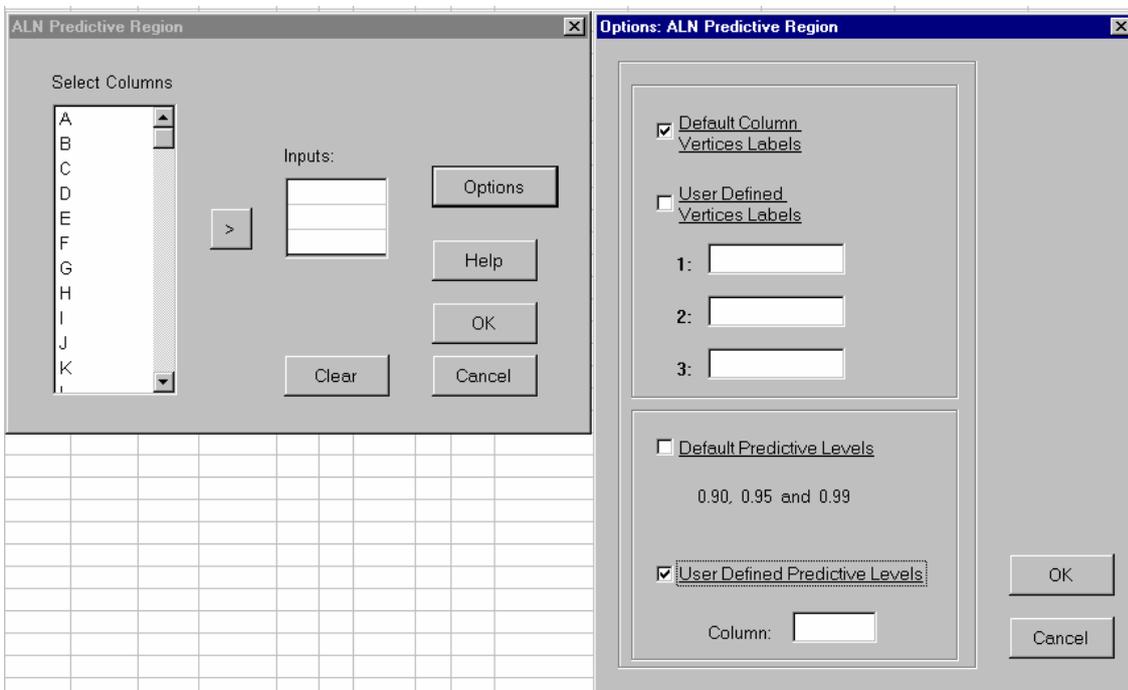
This feature calculates the two compositional Principal Components for a 3-part composition of 3 selected parts and displays the result in a ternary diagram.



There are two options to modify the appearance of the graph:
 1) differentiate, by colour or by shape, each point depending on a previous classification, and
 2) label the vertices of the triangle (the default labels are the part names)

3.6 Graphs. ALN Predictive Region

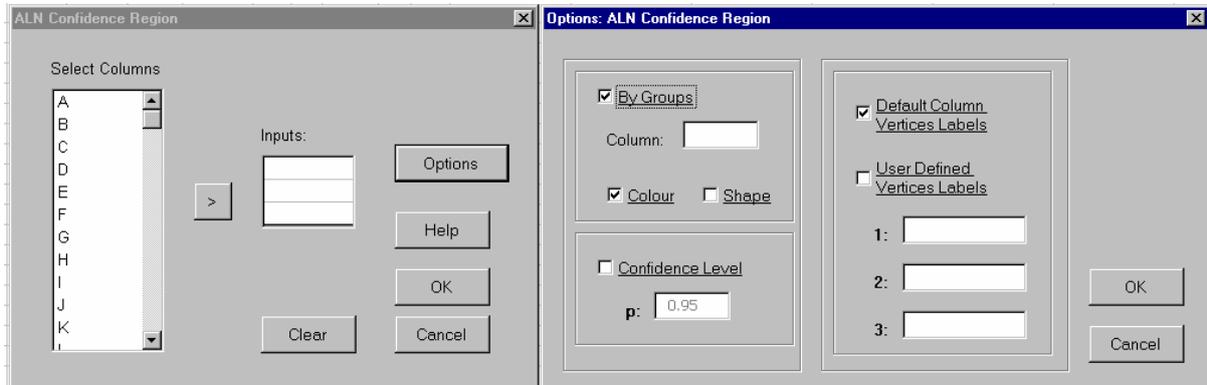
This feature calculates the Additive Logistic Normal Predictive Region of the selected parts and displays the result in a ternary diagram.



There are two options to modify the appearance of the graph:
 1) label the vertices of the triangle (the default labels are the part names), and
 2) choose the default predictive levels (the default levels are 0.90, 0.95 and 0.99)

3.7 Graphs. ALN Confidence Region

This feature calculates the Additive Logistic Normal Confidence Region for the mean vector of the selected parts and displays the result in a ternary diagram.



There are three options to modify the appearance of the graph:

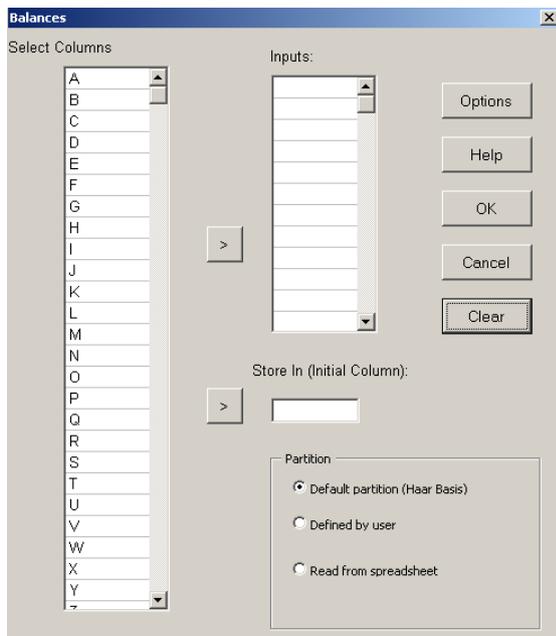
- 1) perform an ALN Confidence Region for each group defined by a column,
- 2) label the vertices of the triangle (the default labels are the part names), and
- 3) define the confidence level (the default is 0.95)

3.8 Graphs. Balances- Dendrogram

Input

The input required consists of the list of the parts to be included in the analysis, the sequential binary partition into groups of parts desired, and the output of interest.

Main Window



After the invocation of the Balance-Dendrogram routine a new window appears. This window contains on its left side the *Select Columns* structure, on the middle the *Inputs* structure and the *Store In (Initial Column)* box, and on the right there are buttons: *Options*, *Help*, *OK*, *Cancel* and *Clear*. Between the

left and the middle part there are two arrows to pass information between them. Also, in the lower right part of the window there are some buttons to specify the way of definition of a partition.

Once all necessary options have been selected, pressing the *OK* button the Balance-Dendrogram routine starts the execution.

Definition of the partition

A partition is a hierarchical grouping of parts of the original compositional vector, starting with the whole composition as a group and ending with each part in a single group. First the compositional vector is divided into two non-overlapping groups of parts. In a similar way, each of these two groups is divided again, and so on until all groups contain only a single part. If D is the number of parts of the original composition, the number of steps of partition is $D-1$. CoDapack includes three different ways to define a partition:

1) Read the partition from the spreadsheet

In this case, the first column of the spreadsheet that contains the definition of the partition has to be indicated. To do so, the corresponding row of the *Select Column* list is marked with a single click and then it is selected with a click on the arrow.

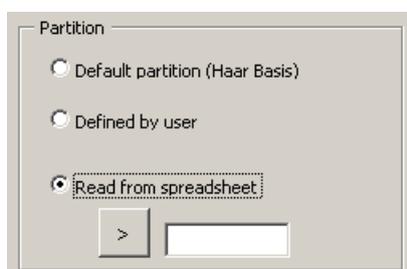
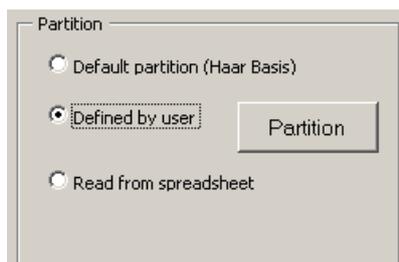


Figure contains an example of a partition. The first row represents the labels of parts of the compositional vector and these labels should be the same as the labels of the parts of the original composition. If there are no labels, this row should remain empty. In this case, CoDaPack understands that the first column of the partition corresponds to the first part selected as the input composition, and maintains the order of the other parts. At each order of partition, +1 means that the part is assigned to the first group, -1 to the second, and 0 that this part is not in the group which is divided at this order.

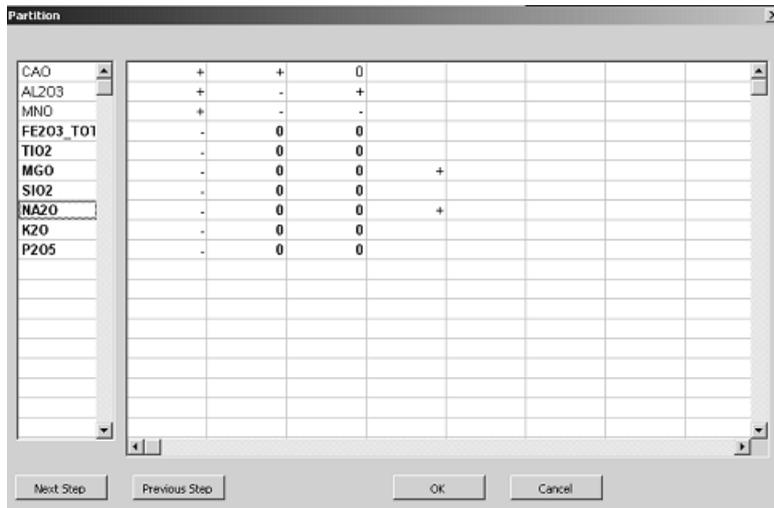
	BT	BU	BV	BW	BX	BY	BZ	CA	CB	CC
SiO2	TiO2	Al2O3	Fe2O3	TcMnO	MgO	CaO	Na2O	K2O	P2O5	
	1	1	1	1	1	-1	-1	-1	-1	-1
	1	1	1	-1	-1	0	0	0	0	0
	1	1	-1	0	0	0	0	0	0	0
	1	-1	0	0	0	0	0	0	0	0
	0	0	0	1	-1	0	0	0	0	0
	0	0	0	0	0	1	1	1	-1	-1
	0	0	0	0	0	1	1	-1	0	0
	0	0	0	0	0	1	-1	0	0	0
	0	0	0	0	0	0	0	0	1	-1

2) Enter a partition using an auxiliary window



Activating this option, a new button appear, and clicking on it a new window. This window has two grids and four buttons. The left grid contains the labels of each part of the original compositional

vector. To define the partition, every time the user marks with a single click the label of one part, a + sign appears in the partition grid at the cell corresponding to this part in the current step. At each step of partition, a + sign means that the part is assigned to the first group, a - sign to the second, and 0 that this part is not in the group which is divided at this order. At each step it is only possible to introduce + signs.



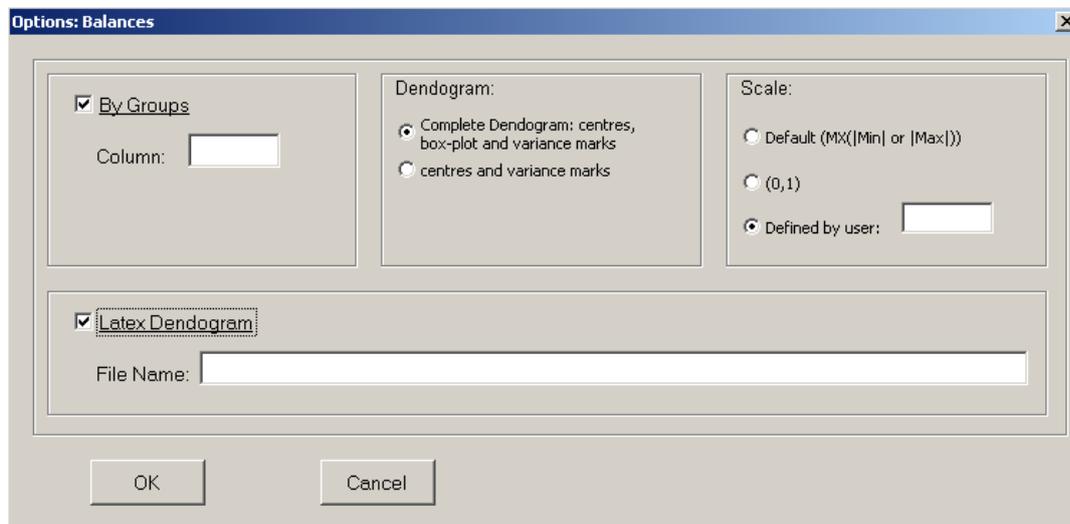
To remove a + sign from the current step it is necessary to mark the cell of the current step of the partition grid that contains this + sign with a single click. To finish a step, press the *Next Step* button: all - and 0 signs are automatically placed where they correspond. At each step it is only possible to divide one group. This group is marked in bold characters on the label grid and on the partition grid. In order to facilitate this task, when *Next Step* button is pressed, all the information (labels and partition) is reordered in such a way that the next parts to divide appear all together. To eliminate some steps of the partition, press the *Previous Step* button as many times as required.

3) Default partition - Haar basis

The default partition is defined by the Haar basis. It consists in separating, at each step, the parts in the middle. Figure shows the corresponding +1, -1 and 0 codification.

Options Window

After pressing the *Options* button of the main form, the *Balance-Dendrogram* routine shows a new window which includes different options mainly focused on the appearance of the dendrogram. They can be selected all together, or only some of them.



1) By Groups: Is it possible to include in the same dendrogram information about different groups. For every group a balance analysis using the same partition is performed and a dendrogram is drawn with the information of the variance on the balance components associated with each binary partition with a different colour for each group according to the scale of the first group. The default option is without groups. If it is activated, the column letter of the *Excel* sheet indicating the group of each observation has to be indicated.

2) Dendrogram: By default the dendrogram contains, for each balance, the sample centre, a box-plot of quantiles, and the proportion of the sample total variance associated with it. This variance is drawn as vertical coloured lines. If groups are defined, the variance is drawn in a different colour for each group. Centres and box-plots of each binary partition are placed horizontally inside the dendrogram. The box-plots show the quantiles 5, 25, 50, 75 and 95. If two groups are defined, CoDaPack draws in blue only a half of the box-plot of the first group (below the horizontal line) and in red a half box-plot of the second group (above the horizontal line). If there are more than 3 groups, CoDaPack only draws the box-plot of the first group.

3) Scale: There are three scale options to draw the horizontal segments: a) Each horizontal segment is scaled as a part of the real axis of the balance defined by the interval $(-u, u)$, where u is the maximum value of minimum or maximum data, in absolute value, corresponding to all parts, that is $u = MX(|Min|or|Max|)$. This is the default option. b) Each horizontal segment is scaled in a 2-part simplex S^2 , i.e. in the segment $(0,1)$. c) Each horizontal segment is scaled as a part of the real axis of the balance. This option requires a u value to define an interval $(-u, u)$ on which the balance is drawn.

3) Scale: There are two scale options to draw the horizontal segments: a) Each horizontal segment is scaled in a 2-part simplex S^2 , i.e. in the segment $(0,1)$. This is the default option. b) Each horizontal segment is scaled as a part of the real axis of the balance. This option requires a u value to define an interval $(-u, u)$ on which the balance is drawn.

4) Latex Dendrogram: The Balances-Dendrogram routine can create a LaTeX code to draw the dendrogram. This dendrogram is drawn with a clockwise rotation of 90° with respect to the original dendrogram. That is, horizontal lines appear as vertical and viceversa. By default no LaTeX dendrogram is drawn. If this option is selected, the name of the file has to be indicated, with its complete path if different from the active directory.

4. Descriptive statistics menu

This menu returns characteristic values for a data set, like:

4.1 Descriptive Statistics. Summary

Performs five descriptive statistics: two of logratios (Variation Array and CLR Variance) and three compositional descriptive statistics (Center, Min, Max and quartiles)

1) Variation Array: Returns a matrix where upper diagonal contains the logratio variances and the lower diagonal contains the logratio means. That is, the ij -th component of the upper diagonal is $\text{var}[\ln(\mathbf{X}_i/\mathbf{X}_j)]$, where and ij -th component of the lower diagonal is $E[\ln(\mathbf{X}_i/\mathbf{X}_j)]$, where $(i,j=1,2,\dots,D)$.

2) CLR Variance: Returns the sum of logratio variances that involve each part. The sum of all CLR

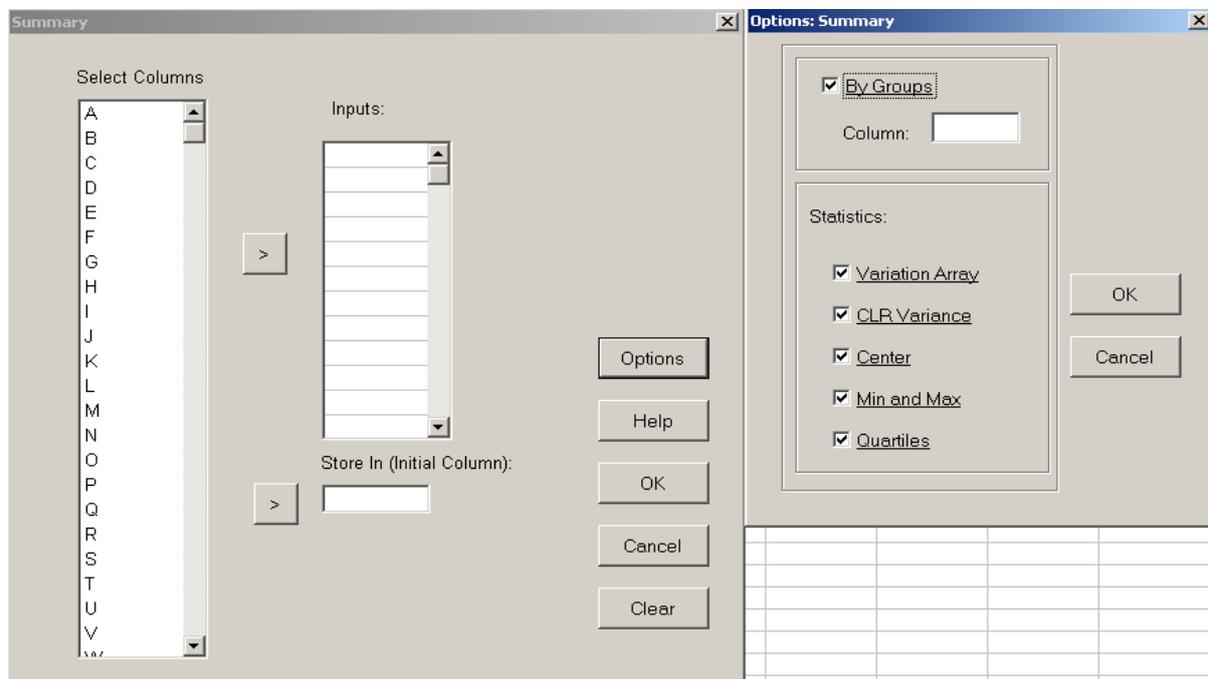
$$\text{Variances is the Total Variance. So CLR-Variance} = \frac{\sum_{j=1, i \neq j}^D \text{var}[\ln(\mathbf{X}_i/\mathbf{X}_j)]}{2D}$$

3) Center: Returns centre of the data set, that is, $\hat{\xi} = C[g_1, g_2, \dots, g_D]$, where

$g_i = \left(\prod_{k=1}^N x_{ki} \right)^{1/N}$ symbolizes the geometric mean of part \mathbf{X}_i in data set \mathbf{X} . The data set \mathbf{X} has been previously closed.

4) Minimum and Maximum: For each part of the data set \mathbf{X} returns the maximum and the minimum of the closed data set $C(\mathbf{X})$.

5) Quartiles: For each part of the data set \mathbf{X} returns Q1, the median and Q3 of the closed data set $C(\mathbf{X})$.



The user has to select the columns to closure and where to put the results.

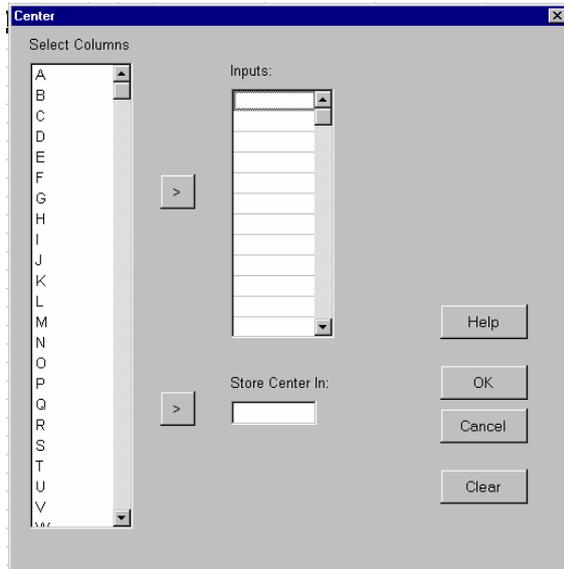
There are two options on this routine:

- 1) perform the statistics for each group defined by a column,
- 2) the user can choose which descriptive wants (at least one must be chosen).

4.2 Descriptive Statistics. Center

With this feature the user obtains the centre of the data set, that is, $\hat{\xi} = C[g_1, g_2, \dots, g_D]$ where

$g_i = \left(\prod_{k=1}^N x_{ki} \right)^{1/N}$ symbolizes the geometric mean of part \mathbf{X}_i in data set \mathbf{X} .



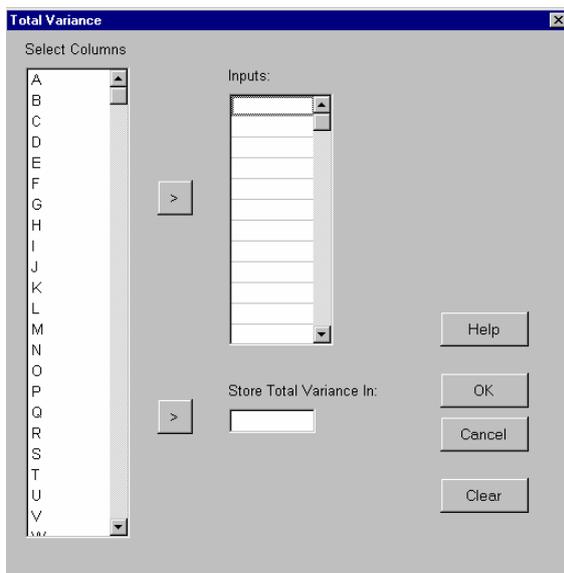
The user has to select the columns to calculate the center and where to put the results.

4.3 Descriptive Statistics. Total Variance

With this feature the user obtains the total variance of the selected columns. The routine returns the sum of all the elements in the upper diagonal variation matrix divided by D , that is

$$\text{totvar}(\mathbf{X}) = \frac{\sum_{i=1}^{D-1} \sum_{j=i+1}^D \text{var}[\ln(\mathbf{X}_i/\mathbf{X}_j)]}{D} = \frac{1}{N} \sum_{i=1}^N d_a^2(\mathbf{x}_i, \hat{\xi}),$$
 where d_a is the Aitchison

distance. Equivalently, total variability $\text{totvar}(\mathbf{X})$ can be calculated as the sum of the diagonal elements of the covariance matrix of the clr-transformed data set, $\text{clr}(\mathbf{X})$.

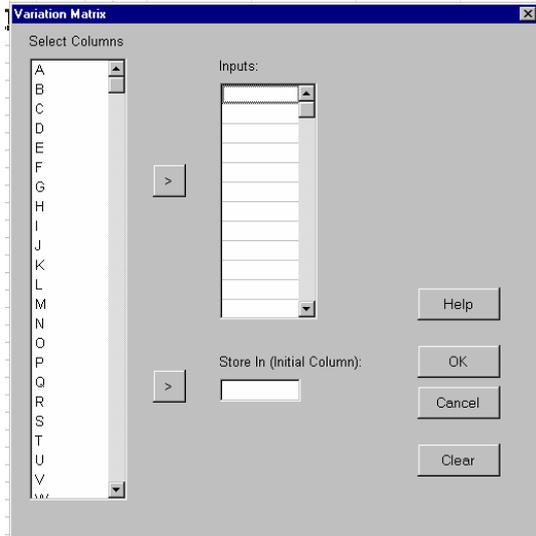


The user has to select the columns to calculate the Total variance and where to put the results.

4.4 Descriptive Statistics. Variation Array

With this feature the user obtains the Variation array of the selected columns. It returns a matrix where upper diagonal contains the logratio variances and the lower diagonal contains the logratio means.

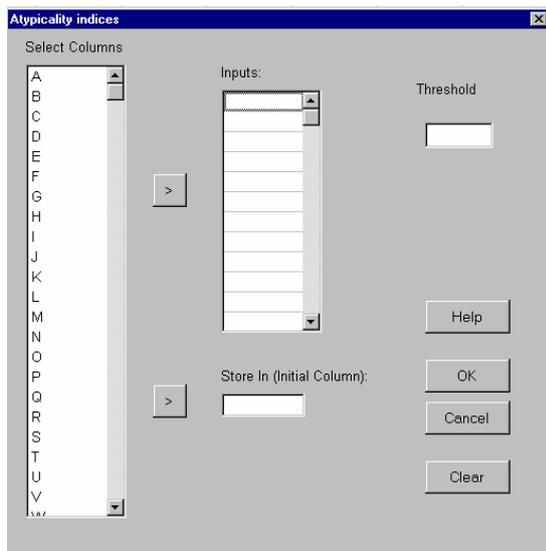
That is, the ij -th component of the upper diagonal is $\text{var}[\ln(\mathbf{X}_i / \mathbf{X}_j)]$, where $(i,j=1,2,\dots,D)$. and ij -th component of the lower diagonal is $E[\ln(\mathbf{X}_i / \mathbf{X}_j)]$, where $(i,j=1,2,\dots,D)$.



The user has to select the columns to calculate the variation array and where to put the results.

4.5 Descriptive Statistics. Atypicality Indices

With this feature the user obtains the atypical observations and their indices under the assumption of Additive Logistic Normal distribution of the selected parts.



The user has to select the columns to calculate its atypical observations and where to put the results. Also has to indicate the Threshold of atypicality (usually 0.95).

5. Analysis menu

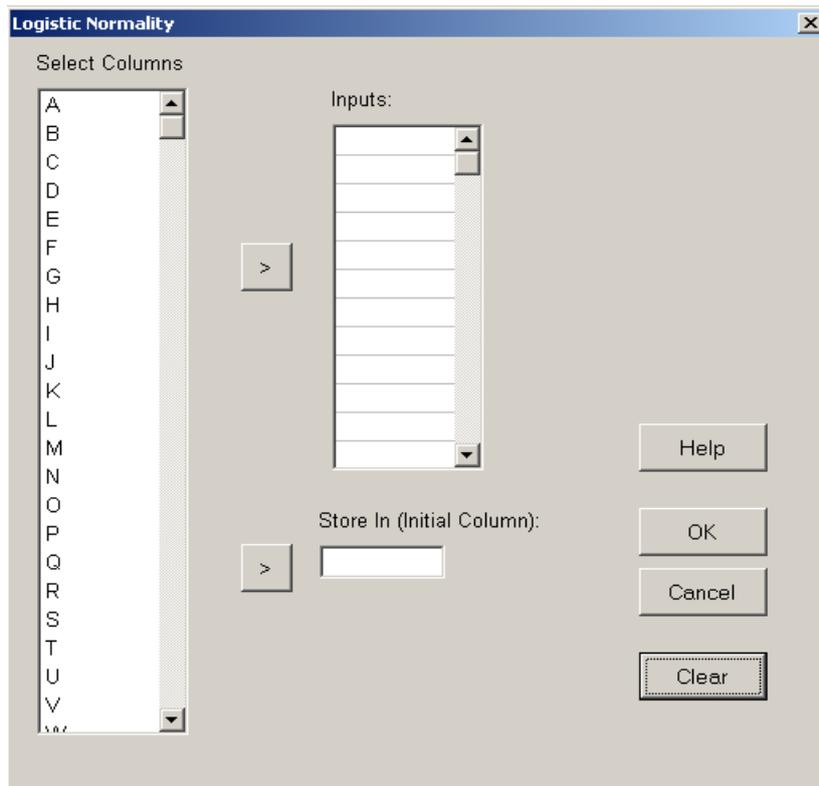
This menu at this moment only performs:

5.1 Analysis. Logistic Normality Test

This feature performs a Test for:

- 1) All marginal, univariate distributions (with a total of D tests)
- 2) All bivariate angle distributions (with a total of $1/2D(D-1)$ tests)
- 3) The D -dimensional radius distribution.

For each kind of test Anderson-Darling, Cramer-von Misses and Watson tests are performed.



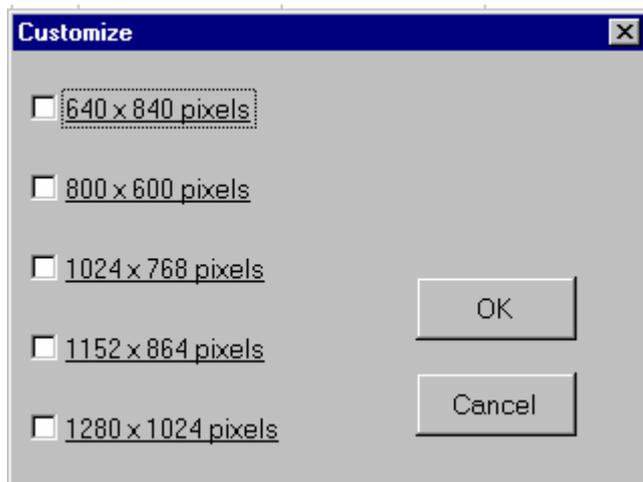
6 Preferences menu

This menu returns characteristic values for a data set, like:

6.1 Preferences. Screen Size

With this feature the user indicates which resolution of screen has in order to obtain complete graphs. The default value is 1152x864 pixels.

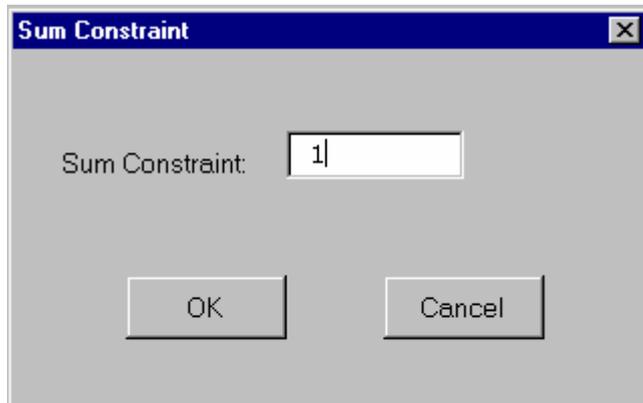
This routine writes the file "screensize.cdp" that is read by all the graphical routines.



6.2 Preferences. Sum-constraint

With this feature the user indicates which is the constant used by the user to closure the data. The default value is 1.

This routine writes the file "sumconstraint.cdp" that is read by others routines.



INSTALLATION

The installation of the program must be done only once and then every time you open excel the program will be used without any installation.

There are two ways to install the macros: Automatic and Manual. We recommend the Automatic installation that just consists on execute a macro. If this macro doesn't work (if the EXCEL® used in your country is not compatible with this macro) you should use Manual installation.

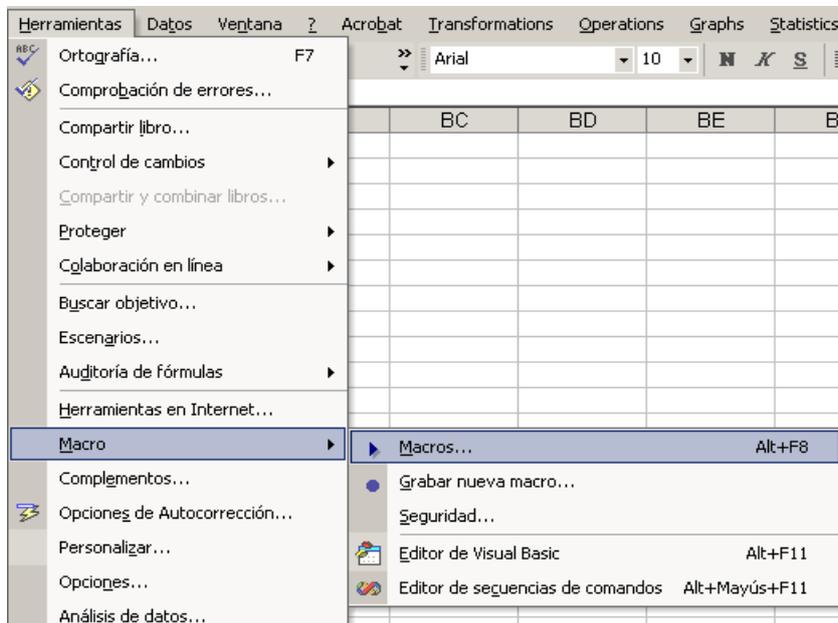
AUTOMATIC INSTALLATION

STEP 1

Open the file Codapack.xls inside EXCEL®, charging the macros

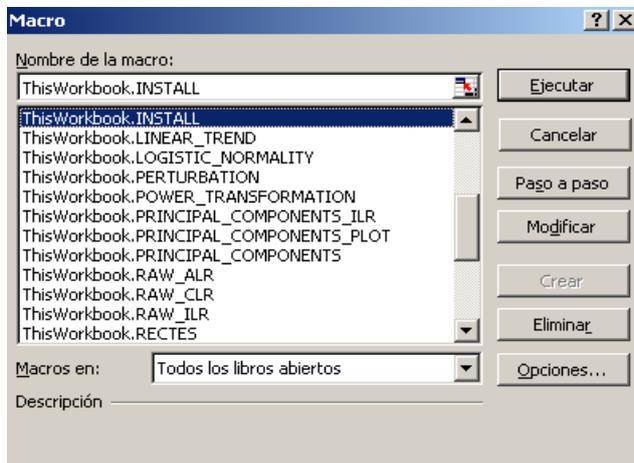
STEP 2.

Open the menu **Tools** (herramientas) and select the option **Macros...** (Macro...) inside the submenu **Macro**. An alternative to the menus is to press Alt+F8.



STEP 3

Execute the macro "ThisWorkbook.INSTALL"



MANUAL INSTALLATION

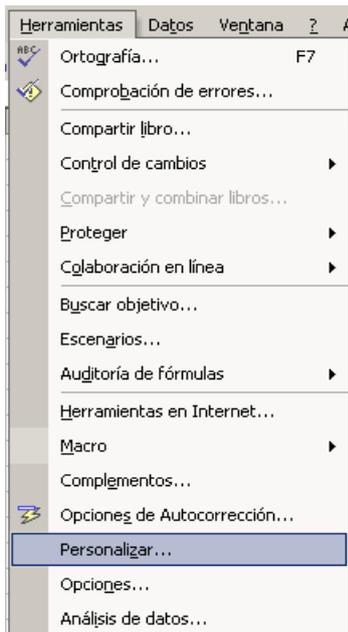
STEP 1

Open the file Codapack.xls inside EXCEL®, charging the macros

STEP 2. How to create a new menu

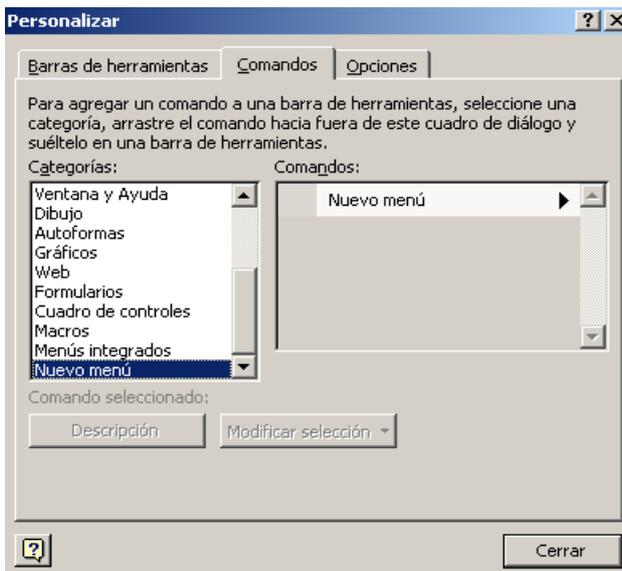
STEP 2.1

Open the menu **Tools** (herramientas) and select the option **Customize** (Personalizar...)



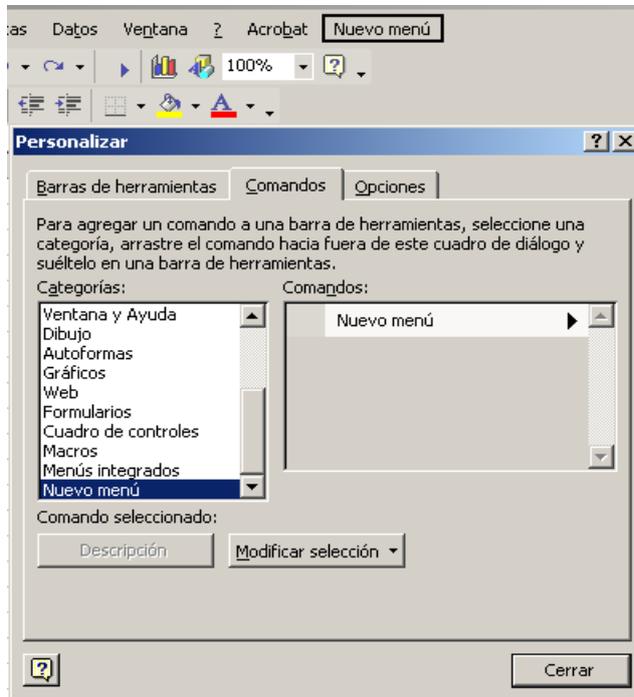
STEP 2.2

Select the **Commands** (Comandos) menu and the category **New Menu** (Nuevo menú)



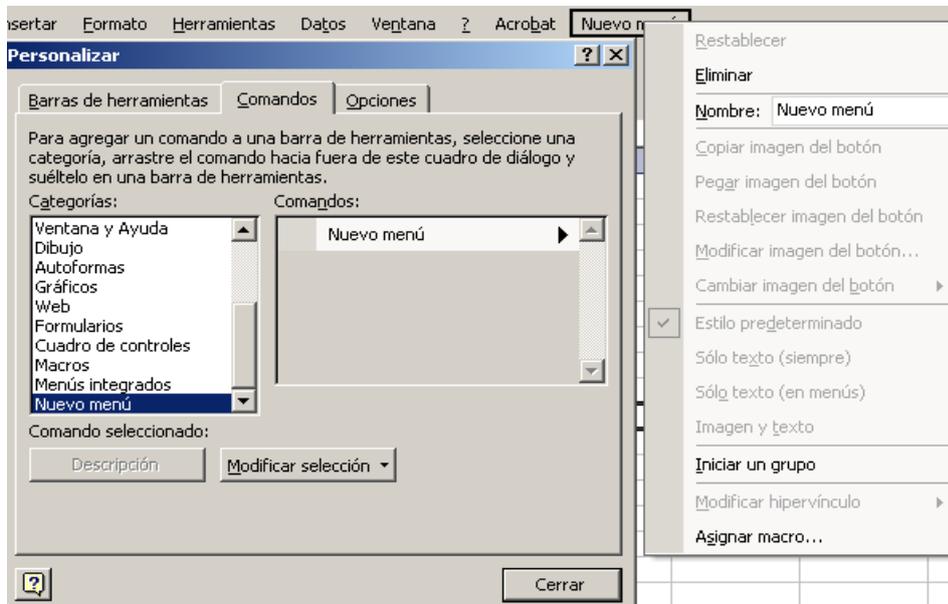
STEP 2.3

Select with the mouse **New menu** (Nuevo menú) and drag it trough the top of the window near the other menus



STEP 2.4

Select it with the right button of the mouse and change the **name** (nombre) by p.e. *Transformations*.



STEP 3. How to assign a new function in a menu

STEP 3.1

Open the menu **Tools** (herramientas) and select the option **Customize** (Personalizar...) as in step 2.1

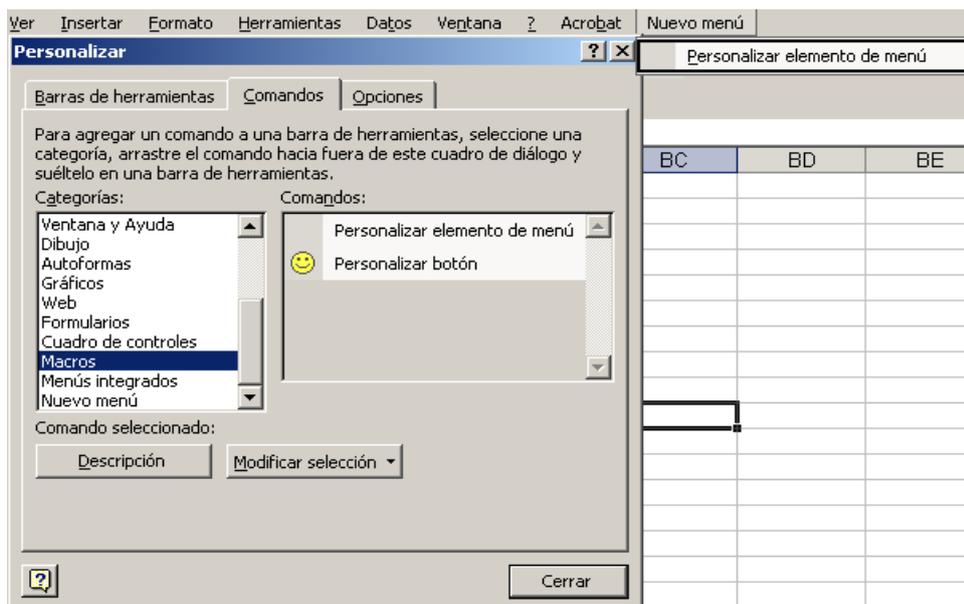
STEP 3.2

Select the **Commands** (Comandos) menu and the category **Macros**



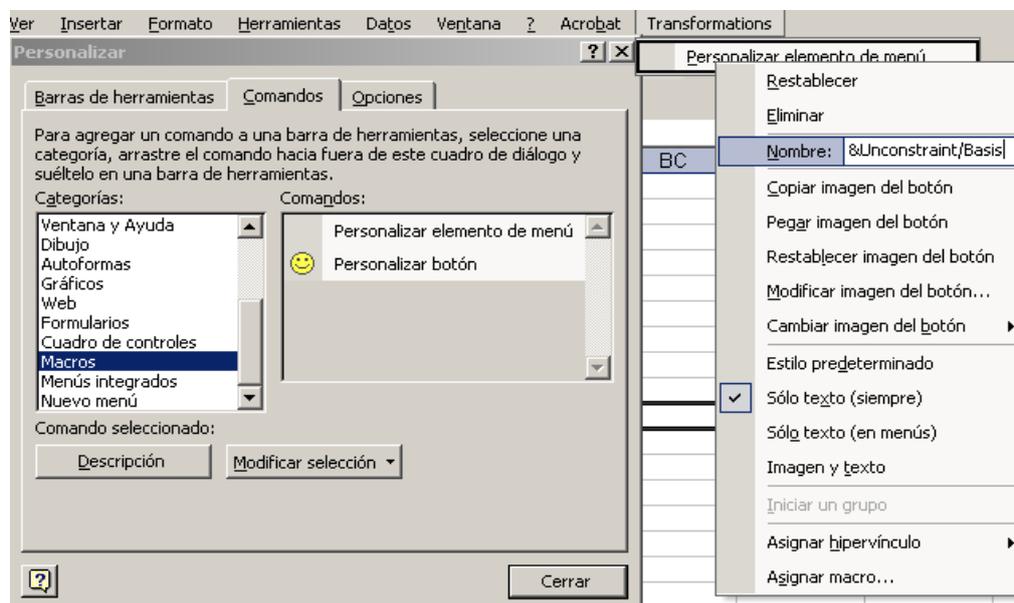
STEP 3.3

Select with the mouse **Custom Menu Item** (Personalizar elemento de menú) and drag it behind the menu that it belongs



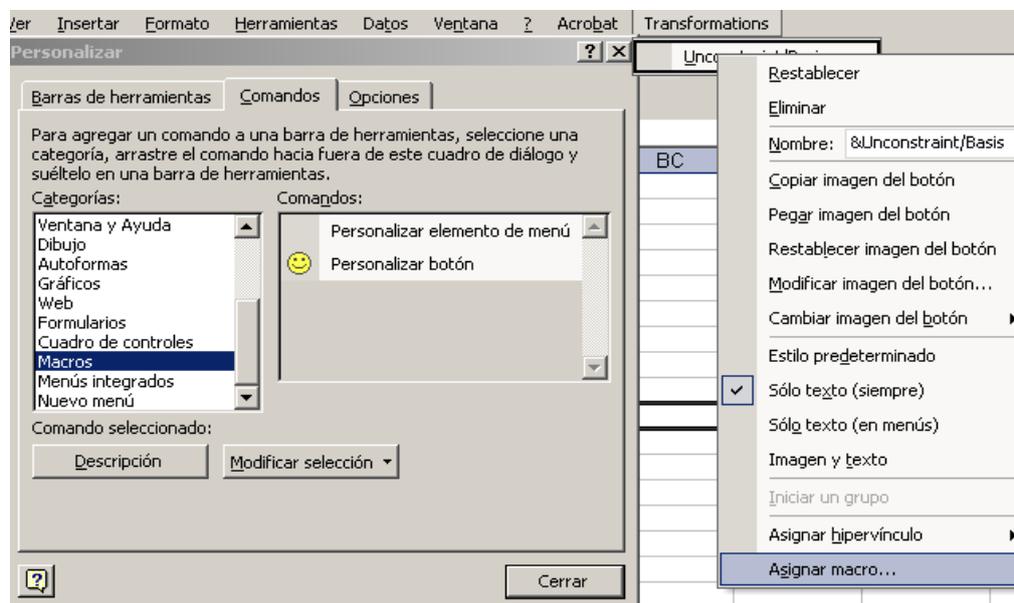
STEP 3.4

Select it with the right button of the mouse and change the **name** (nombre) by p.e. *Unconstrain/Basis*.



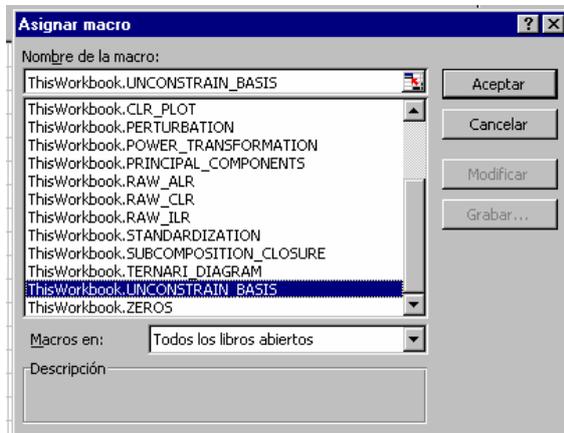
STEP 3.4

Select on the same menu **Assign macro...** (asignar macor).



STEP 3.5

And assign the corresponding macro **assign macro** to the submenu



Repeat STEP 3 as many times as macros has the menu and repeat STEP2 and as STEPS 3 as needed to complete all the other menus

UNINSTALLATION

There are two ways to uninstall the macros: Automatic and Manual. We recommend the Automatic uninstallation that just consists on execute a macro. If this macro doesn't work (if the EXCEL® used in your country is not compatible with this macro) you should use Manual uninstallation.

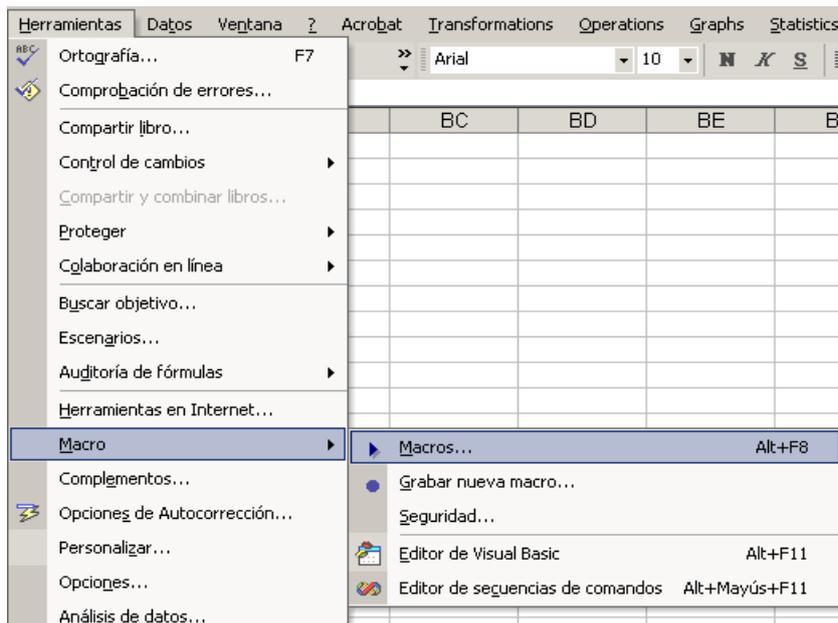
AUTOMATIC UNINSTALLATION

STEP 1

Open the file Codapack.xls inside EXCEL®, charging the macros

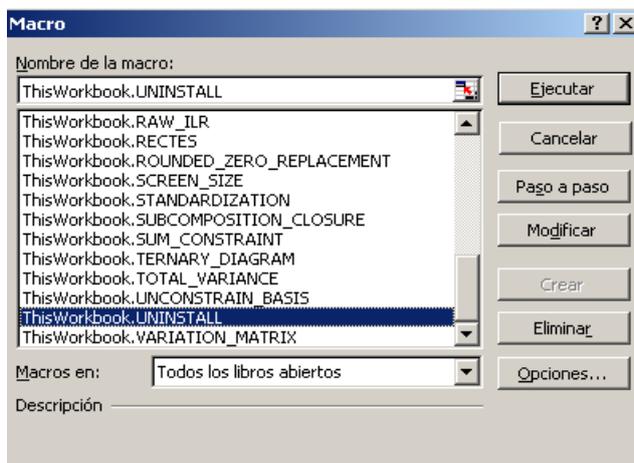
STEP 2.

Open the menu **Tools** (herramientas) and select the option **Macros...** (Macro...) inside the submenu **Macro**. An alternative to the menus is to press Alt+F8.



STEP 3

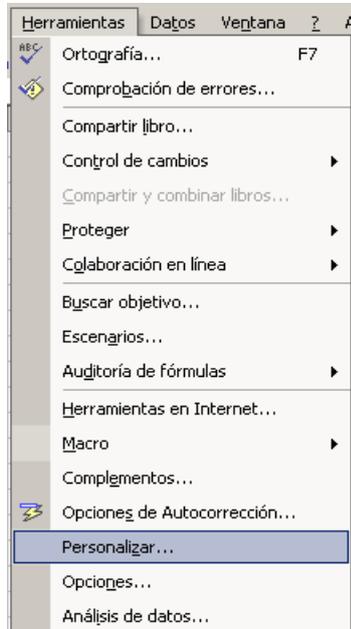
Execute the macro "ThisWorkbook.UNINSTALL"



MANUAL UNINSTALLATION

STEP 1

Open the menu **Tools** (herramientas) and select the option **Customize** (Personalizar...)



STEP 2

Drag all menus outside the top of the window