

Statistics and Data analyses

Ulisse Di Corpo

The menu "Statistics" allows to select groups of records, output lists and reports and produce statistical analyses as frequency distributions, cross tables, correlation tables, factor analysis and cluster analysis.

Record Selection

The first option "Record selection" is used to enter searching conditions and to select groups of records. On selected records it is possible to perform statistical analyses and output the information as lists and reports.

It is possible to select groups of records also in two different parts of the program:

1. in the section "Search" (menu "Data") when searching as a list;
2. in the section "Fast analyses" (menu "Statistics").

When a group of records is selected it stays active until a new selection is performed. When you exit and restart the program the old selection will stay active.

Selecting records

Selection conditions are entered directly on the form. If no conditions are entered all the records will be selected, when conditions are entered only the records which satisfy these conditions will be selected.

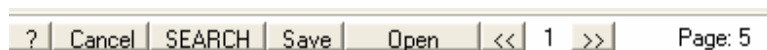
Commands:

1. choose the field in which you want to enter a searching condition, press the "C" to open the window which allows to enter conditions. If no window opens it means that the cursor is positioned on an alphabetical field: in this section it is not possible to enter search conditions in alphabetical fields. If you want to enter search conditions in alphabetical fields use the "list" option in the search menu "Data".
2. In order to remind that in a field a condition is present the program shows the character "=" (equal). If you want to search all the records which do not satisfy the conditions that you

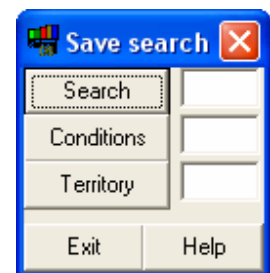
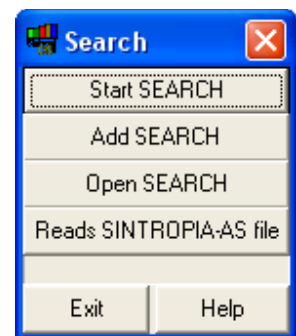
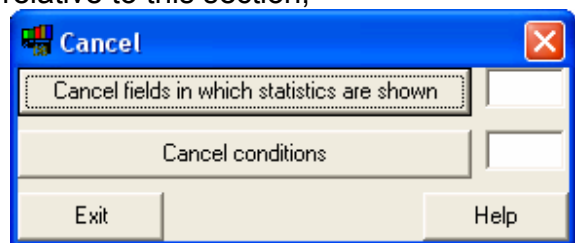
have entered press the key "#" or "=", and the character "# will be shown.

3. If you want to see some very essential distributions, press the key "S", the program will answer with "St", to remind that counting will be performed.
4. If you want to use a quantitative variable as unit, select it and press the key "Q". In order to remind that the variable has been selected as a unit the character Q will be shown. Remember to enter a condition, on the quantitative variable, that only records with values from 0 onwards will be selected.

Below the menu line you will see the following command line:



- The first button "?" opens the content of the help relative to this section;
- The second button "Cancel" opens the window "Cancel" which allows to cancel all the conditions entered in the form and all the fields in which "statistics" (distributions) will be shown. If the form is big it can be good practice to cancel conditions, using this option, before you start a new search.
- When you press the button "Search" the search window will be shown. Usually the option "Start SEARCH" is used. If you want to add a new search to an already existing selection use "Add SEARCH", if you want to Open an already existing search use the option "Open SEARCH", in this case the existing selection will be intersected (AND) to the opened search. The last option "Reads Sintropia-AS file" allows to open a territorial selection performed with the geographical software Sintropia-AS. Once a search is performed it is possible to show the distributions relative to fields previously selected with the key "S" (marked with "St") using the key "D" (distribution).
- Once a search ha been performed it is possible to save it using the button "Save". A window is shown which shows 3 different save options, the first one "Search" saves the search result, the second one "Conditions" saves the search conditions, the third one "Territory" saves the territorial distributions which can be opened by Sintropia-AS and shown as maps.
- Using the button "Open" it is possible to open previously saved searches, conditions, and results of territorial analysis performed with Sintropia-AS.

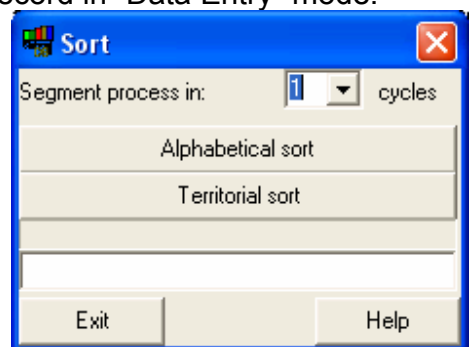
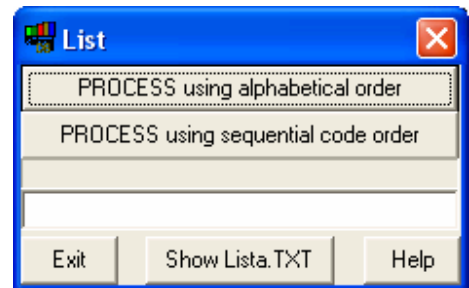
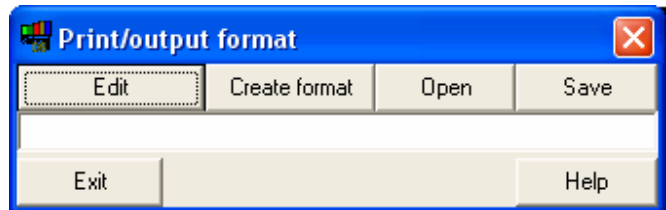


Lists

This section allows to produce lists of the selected records. The command line is:



- The first button “?” opens the help relative to this section.
- The second button “Format” opens a window which allows to define which fields to output and their format. The button “Create format” creates a format with all the fields which you can Edit using the button Edit.
- The third button “List” allows to produce lists, which will be saved in the file Lista.TXT (in the folder of the program). Lists can be produced using alphabetical orders or sequential orders. Alphabetical orders are produced using the button “Sort”.
- The fourth button “Show list” shows the file Lista.TXT. It is important to note that when you select a line the program goes to the corresponding record in “Data Entry” mode.
- The button “Sort” shows the window “Sort”. Choose a field using the key “O”, the program will answer marking the field “O” enter the Sort menu an press the button “Alphabetical sort”. When the number of records is high it can be useful to segment the sorting process. The default value is 1. If sorting time is high, or problems are encountered just increase the sorting segmentation value. Alphabetical sorting treats differently capital and small characters.
- The button “Cancel” cancels the field which has been selected for alphabetical sorting.



Analyses

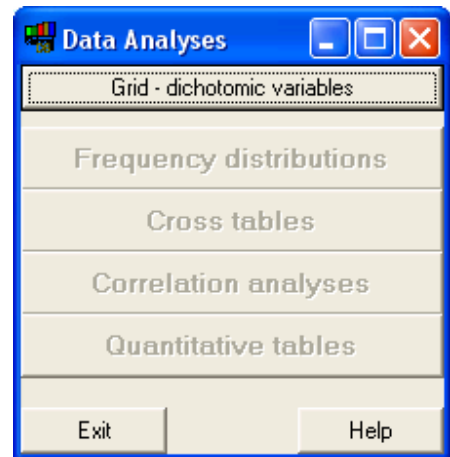
Sintropia-DS offers two types of analyses, traditional and fast. Fast analyses require that the translation grids has been prepared using traditional data analyses. So traditional data analyses are preliminary to fast analyses.

Data analyses

Sintropia-DS processes data using a virtual record in which all data are translated into dichotomic variables (Yes/No). The grid tells how to translate data into dichotomic variables. Data analyses require the existence of the grid and it is not possible to access them if the grid has not been created.

Dichotomic variables

Dichotomic variables (Yes/No, True/False, Present/Absent) are the simplest form of measure. Any type of information, qualitative or quantitative, can be translated in one or more dichotomic variables. With dichotomic variables is possible to merge together information of any type: quantitative and qualitative, nominal, ordinal, integer and real.



Analyses based on dichotomic variables allow to study the most complex natural phenomena, in a similar way to what happens with computer where starting from the bit of information (true/false) it is possible to reproduce the most complex situations.

How information is translated into dichotomic variables

Information can have the form of:

1. nominal variables;
2. ordinal variables;
3. integer variables;
4. real variables.

These 4 types of variables vary according to the mathematical operations which can be performed on them:

1. in the case of nominal variables the value is a modality, for example "Italy", "France", "Germany". With nominal variables it is possible only to count the occurrences of the modality.

2. In the case of ordinal variables the value is a sequence: "First", "Second", "Third"; "Elementary education", "High School", "University". It is possible to divide the sequence into high and low, for example "high education", "low education", or treat each value as a modality (nominal variable).
3. Integer variables (or equal interval variables), for example temperature: 1 degree Celsius, 2 degrees, 3 degrees Celsius. It is possible to say that the distance between 1 and 2 degrees is the same distance that exists between 2 and 3 degrees. Integer values add the possibility of adding and subtracting values. In this way it is possible to calculate mean values and deviations.
4. Real variables differ from integers because their 0 value coincides with the absolute 0. For example using Celsius degrees it is incorrect to say that 2 degrees is twice as much temperature as 1 degree. This because absolute 0 degrees does not correspond to 0 degrees Celsius, absolute zero is equal to -273 degrees Celsius. Real variables add the possibility of using divisions and multiplications.

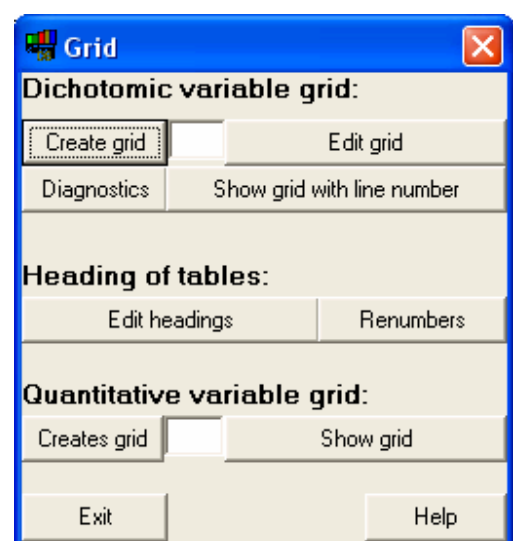
Translation into dichotomic variables:

- in the case of nominal variables, the single modality (for example "Italy") can become a dichotomic variable: yes/no Italy.
- In the case of ordinal variables it is possible to use each modality as a dichotomic variable, but it is also possible to translate the information in the form high/low.
- In the case of integer and real variables it is necessary to find cut-off points. It often happens that the aim of the study is to find the best cut-off points, those which increase correlation values with key variables. For example which is the cut-off between healthy values and pathological values? Dichotomic variables allow to test not only the existence of correlations, but which cut-off point optimize correlations.

Translation grid

Sintropia-DS creates automatically the grid starting from the file Scheda.Txt. To create the grid just press the button "Create grid".

It is usually necessary to edit the grid which has been created automatically. Each line of the grid defines a dichotomic variable. A line is formed of a label, the number of the field, the interval which sets the dichotomic variable true, and of other information which tell if the variable is the first one of an item (group), if it is part of a multiple answers item, if it is the first variable of a synoptic table, if percentages have to use a filter variable as totals.



Example:

```
1. Active*1=1,1;G0S0
1. Canceled*=2,2;
1. MR*=0,0;
```

The first dichotomic variable “1. **Active**” refers to field number 1 (*1), and it is true if values go from 1 to 1 (1,1;). It is the first variable of an item (G) it is not taken from a multiple answer item (0, otherwise the number of answers would appear), it is not part of a synoptic table (s, otherwise the small character s would be used), it does not refer to a filter variable (0).

The second variable “2. **Cancelled**” refers to the same variable (for this reason the number is not repeated), the interval is (2,2;), it is not the first variable of a group.

The third dichotomic variable “1. **MR**” is relative to the missing answer. With coded answers (which refer to lists) the missing answer is zero, while with quantitative variable the missing answer is -32,000 and with long quantitative variables -9,999,999.

The grid file is saved in the same folder of the archive in a text file named “**grl-sch.txt**”. The headings file is saved in a file named “**tin-sch.txt**”. When you work on the grid you might want to cancel entire items, for example you don’t want the list of the provinces. If variables are erased from the grid they must also be erase from the headings. Heading are organized in the following way:

nnn <- progressive number of the heading

```
xxxxxxxxxxxxxxxxxxxxx <- text (it can take more lines)
/// <- end
```

For example:

```
001
1. Record status:
///
002
2. Date:
///
003
3. Sex:
///
004
5. Geographical area:
///
```

When you cancel or add a heading it is necessary to renumber them. For this purpose it is possible to use the button “renumbers”.

When *grl-sch.txt* file exceeds 32k (32,000 characters) it is necessary to use a word processor and save as a text file. Word of Microsoft often adds special characters which interfere with Sintropia-DS, and does not allow to share open files with other programs, so it usually is more convenient to use Wordpad.

Editing the file *grl-sch.txt* you might make some mistakes, as forgetting a comma, etc. The button diagnostics finds out these errors. It is good habit to use this option before moving to data analyses.

Quantitative variables grid

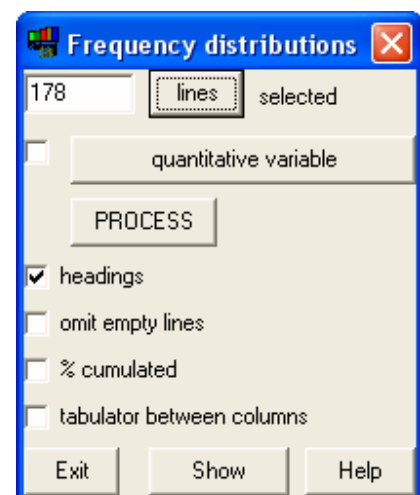
With quantitative variables it is possible to calculate sums, average value and deviation. The grid is very simple, the first value is relative to the field, the second to the label, the third to the status (synoptic table and beginning of a group). “S” marks the start of a synoptic table, s the continuation and G the beginning of a group. For example:

```
33*doctors *SG
34*nurses *s
35*volunteers *s
```

Frequency distributions

Select the lines, process, and show the results. It is also possible to produce frequency distributions weighting each record, for example with population loadings. Options:

- Headings: dichotomic variables are usually grouped according to items and fields. For example: Sex, Geographical area, etc.
- Omit empty lines. When records are few but variables are numerous it is easy to have empty lines (where frequency equals to zero).
- % cumulated. When dichotomic variables are taken from nominal variables, percentages are usually calculated comparing frequencies with the total; but when dichotomic variables originate from ordinal or quantitative variables, it can be interesting to show cumulated percentages.
- Tabulator between columns. When the output will use characters different from courier it can be useful to separate columns using tabulators.

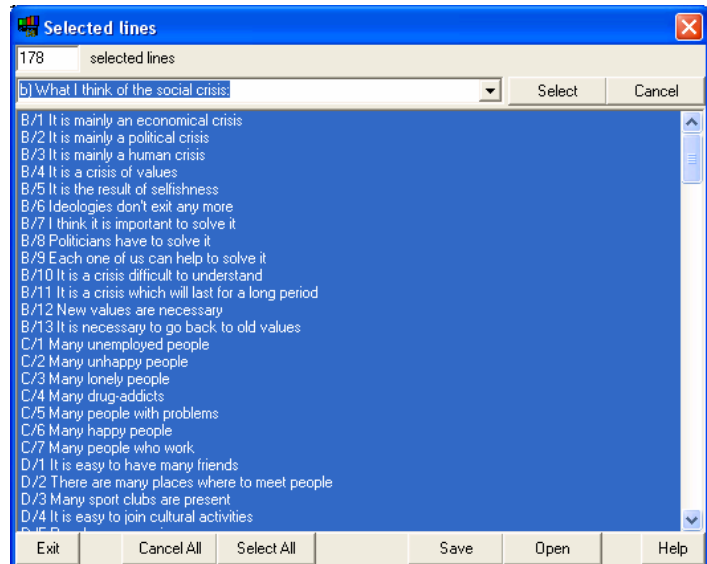


Selection of lines

This options allows to select dichotomic variables by:

- *Selecting them directly with the mouse;*
- *Using the shift key combined with vertical arrows keys;*
- *Choosing the heading and pressing the Select or Cancel button;*
- *Using the Cancel All or Select All buttons at the end of the window..*

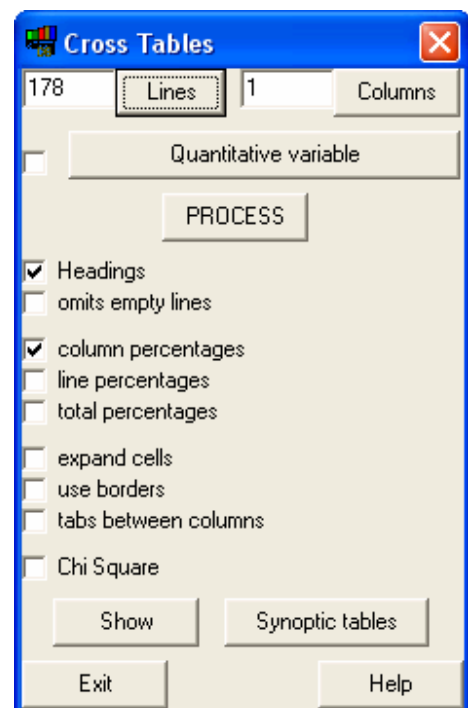
Usually all dichotomic variables are selected, but in some cases only a particular subset of variables is required, and selecting can take time, for this reason it is possible to save and open selections, using the “Save” and the “Open” buttons.



Cross tables

In order to produce cross tables it is necessary to select lines and columns. It is possible to produce cross tables weighting each record, for example with population loadings. The options present in this window are:

- **Headings:** dichotomic variables refer to items, fields, and are usually grouped according to these fields. For example: Sex, Geographical area, etc.
- **Omit empty lines.** When records are few but variables are numerous it is easier to read tables in which empty lines (where frequency equals to zero) are omitted.
- **Column percentages.** The primary aim of cross tables is that of comparing columns in order to find differences/relations among variables. Usually these comparisons are carried out on columns percentages values.
- **Line percentages.** Also row percentages allow to compare variables, it is though common to use column percentages.
- **Total percentages** are rarely used, but in some case it



- might be interesting to compare single cell data with the total value.
- When cross tables are produced using a quantitative unit, it might be useful to expand the dimension of the cells.
 - Enables cells borders using text characters.
 - Tabulator between columns. When the output will use characters different from courier it can be useful to separate columns using tabulators.
 - Chi Square. To assess if the difference which has been observed (for example among columns) is meaningful Chi Square index is used. When Chi Square is followed by (*) the risk of error that we accept stating that a relation is meaningful is less than 5%, when Chi Square is followed by (**) the risk is less than 1%.

Example:

```
col. 1 = Male
col. 2 = Female

b) What I think of the social crisis:

      1      2      Total
B/1 It is mainly an economical      24      37      68
      % Col.      6.63      8.26      8.04
B/2 It is mainly a political c      168      200      385
      % Col.      46.41      44.64      45.51
      Total (records)      362      448      846
```

This example is taken from the archive "Monteoro" present in Sintropia-DS

The first two lines:

```
col. 1 = Male
col. 2 = Female
```

show the labels of the columns of the table, in this example the first column is relative to males and the second to females.

The heading of the tables follows (headings are stored in the file "*tin-sch.txt*"):

```
b) What I think of the social crisis:
```

Then the heading of the columns of the table is shown: `1 2 Total`, the meaning of these numbers is found in the first lines where, labels of the columns are listed. Then cell values follow:

```
B/1 It is mainly an economical      24      37      68
      % Col.      6.63      8.26      8.04
B/2 It is mainly a political c      168      200      385
      % Col.      46.41      44.64      45.51
```

Line labels are truncated, if necessary it is possible to modify them editing the file "*grl-sch.txt*".

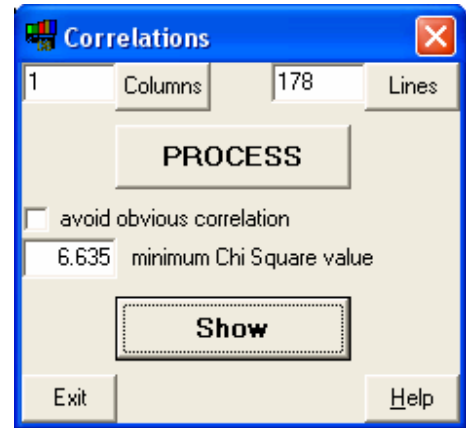
The table is closed with a line of totals: *Total (records)* 362 448 846
 In parenthesis (**record**) reminds that the total is relative to the number of records. This is important, because when items allow multiple answers summing percentages values does not add to 100%.

Correlation analysis

Correlations allow to cross each dichotomic variable with all the other dichotomic variables of the form. Correlation tables:

- **Show the typology** of each dichotomic variable;
- *Provide the same information which is found in cross tables, but, in a concise and more immediate way.*

In order to produce correlation tables it is just necessary to select the columns variables and the line variables. Correlations are just a different way to look at cross tables, if you have just processed a cross table you can directly show the results in the form of correlations.



When showing correlations it is possible to omit obvious correlations (dichotomic variables which originate from the same item always show negative correlations) and choose the lowest Chi Square value.

A Chi Square of 3.841 corresponds to a risk of 5%, a Chi Square of 6.635 to a risk of 1% and a Chi Square of 10.827 to a risk of 1/1000. In most book of statistics it is possible to find the Chi Square distribution. If you will ever need to look for other values remember that the dof value (degrees of freedom) is equal to 1.

How to read correlation values

Chi2	rPhi	(% Yes	% No	
846.00	1.000	(48)	(100.00%/	0.00%)	F/5 I feel useless
453.25	0.732	(47)	(85.87%/	12.40%)	F/4 I feel depressed
248.14	0.542	(50)	(66.17%/	12.40%)	F/7 I am pessimist
195.39	0.481	(49)	(69.38%/	21.11%)	F/6 I feel lonely
162.52	0.438	(46)	(77.30%/	33.77%)	F/3 I feel anxious
155.86	0.429	(51)	(65.74%/	22.69%)	F/8 I don't like my body

Annotations in the image:

- Column variable:** Points to the variable names in the rightmost column of the table.
- Concomitancies between column variable and correlated variables:** Points to the Chi2 and rPhi values.
- Line variables:** Points to the variable names in the rightmost column of the table.
- Chi Square:** Points to the Chi2 values.
- r.Phi:** Points to the rPhi values.
- Sequential position of the variable in the Grid:** Points to the numbers in parentheses (e.g., 48, 47, 50, 49, 46, 51).

The first line shows the label of the column variable, in this case “I feel useless”, the following lines show the correlated variables, sorted by strength of the correlation. The first line shows the most correlated variable. As it is possible to see it is a obvious correlation with itself, the correlation value obtain is the highest possible ($r\Phi = 1.000$, Chi Square = number of counted records), concomitances are equal to 100% and discordances to 0%.

The second line shows the correlation among “I feel useless” and “I feel depressed”:

```
453.25 0.732 ( 47) ( 85.87%/ 12.40%) F/4 I feel depressed
```

Among the interviewed people which say “I feel useless” 85.87% also say “I feel depressed”. Chi Square value is extremely high 435.25 (6.635 is the mark for an error lower than 1%).

Sample dimensions

When the sample is limited Chi Square values are smaller and it is difficult to obtain statistical significant correlations. Interesting correlations start when the sample is over 100 cases. More cases are used (records), more correlations are found. It is though important to say that when a very high number of cases is used (for example 10,000 subjects) even very minor correlations become significant. It is then important to consider also the strength of the correlation which is provided by $r\Phi$ values.

Sample requirements

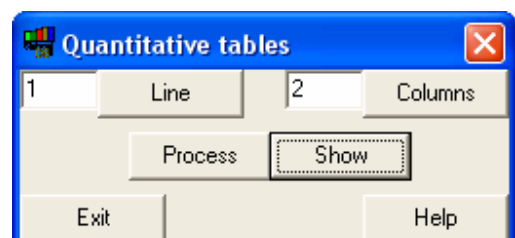
In order to study correlations it is necessary to maximize the variability of the information. In other words it is not necessary to use “representative samples”, but samples in which the traits distribute equally among the two values of the dichotomic variable (yes/no). For example if we want to study what is correlated to a specific pathology we will need 50% with the pathology and 50% without it. The aim of a correlation study is to find correlations and to specify when and how they exist. It is wrong to think that the aim is that of finding correlations which can be generalized to all the population.

Quantitative tables

For each quantitative variable (lines) the total value, the number of cases, the mean values and the deviation are calculated. For example:

Column:

- 1 - TOTAL
- 2 - Male
- 3 - Female



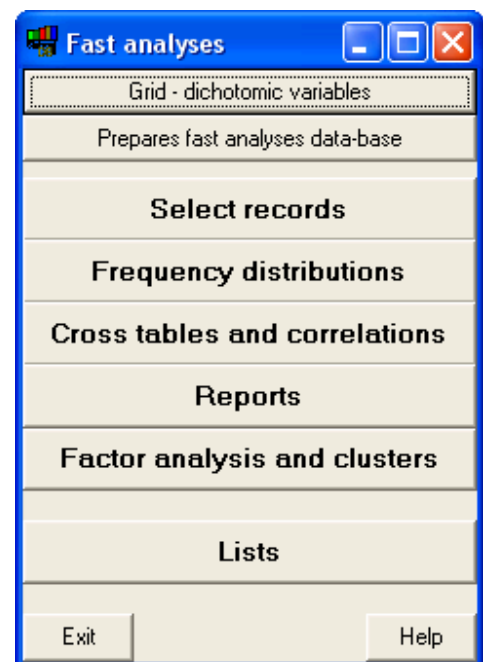
Age :	1	2	3
Total.....:	14105	6101	7372
Cases.....:	846	362	448
Average on total.....:	16.67	16.85	16.46
Average on cases.....:	16.67	16.85	16.46
Deviation.....:	1.65	1.73	1.51

Fast analyses and factor analysis

Sintropia-DS dichotomic data analyses optimize the use of computer binary structures, in this way it is possible to produce instant analyses (fast analyses) even when working with archives of millions of records (in this free version the limit is set to 500,000 records).

In order to activate fast analyses it is necessary to copy the dichotomic grid using the first option “Grid – dichotomic variables”, and use the option “Prepare fast analyses data-base” each time data or grid are updated or changed. Once the grid is ready and data-base updated it is possible to:

1. **select records.** This section allows to select records just choosing the dichotomic variables as searching conditions.
2. produce **frequency distributions.** This window is practically identical to the one discussed in “Data analyses”, the only difference is that there is no “PROCESS” button and results are immediately shown.
3. **Cross tables and correlations.** Cross tables and correlations are now united in the same window, as they are different ways to see the same data analysis.
4. **Reports.** It is possible to organize standard queries and tables, and select them from a list.
5. **Factory analysis and cluster analysis.** This window allows to group variables on the basis of their reciprocal correlations. These variables are then used to group records in homogeneous groups (clusters).
6. **Lists.** Selected records can be outputted in the form of lists.



Grid – dichotomic variables

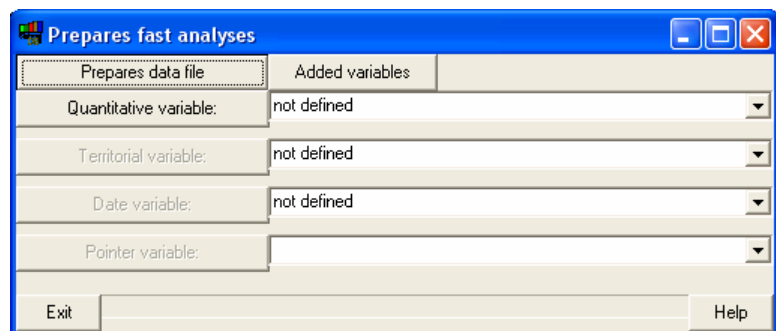
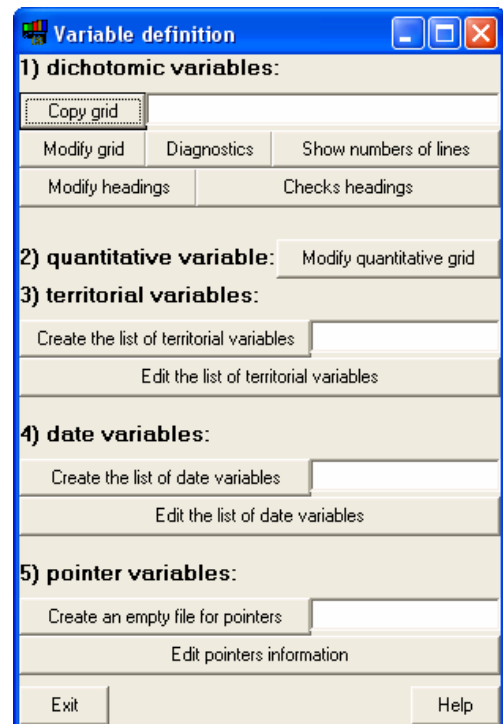
The first time you enter the fast analyses it is necessary to copy the grid. Usually fast analyses are used more to produce correlation tables and factor analyses. This means that labels used in “Fast analyses” can have lengths up to 65 characters which are very easy to read. In fast analyses the file with the grid is named “*grl-dat.txt*” and the file with the headings “*tin-dat.txt*”.

Quantitative, territorial, date and pointer variables can also be used in fast analyses, but they are not necessary.

Pointer variables are required only in those cases in which the data structure is divided among more archives with different statistical units.

Prepare fast analyses data-base

In this section the fast analyses database is prepared. Usually only dichotomic variables are present, but if quantitative, territorial or date variables exist it is necessary to select the variable which will be used in fast analyses. Once you have selected the variables (if necessary) just press the button “Prepares data file”. It is also possible to create new variables (Added variables) as the result of conditions entered on one or more original variables.

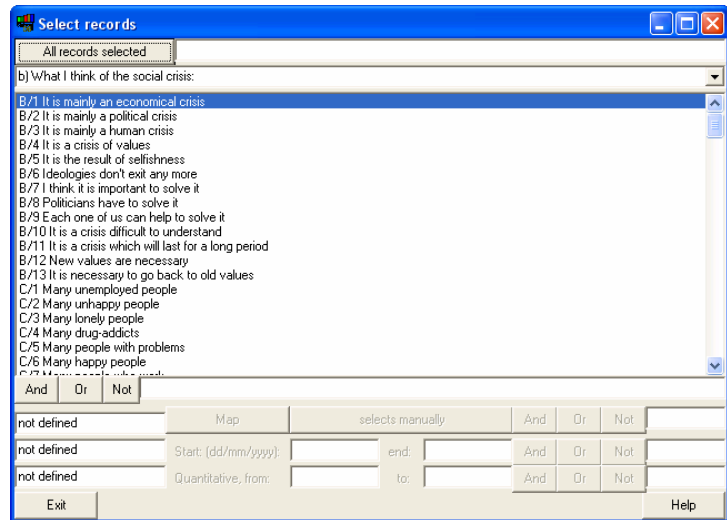


Select records

Start with the button “All records selected”, so that you know from which situation you are starting selecting records.

Just selecting the dichotomic variable you need and choose if you want to perform an AND, OR or NOT search.

When you choose AND, among the records which are already active, only the records which show the variable which you have selected will be left active. When you choose OR the records which show the selected variable will be added to the already active records. When you choose NOT all the records which show the selected variable will be canceled from the already active records.



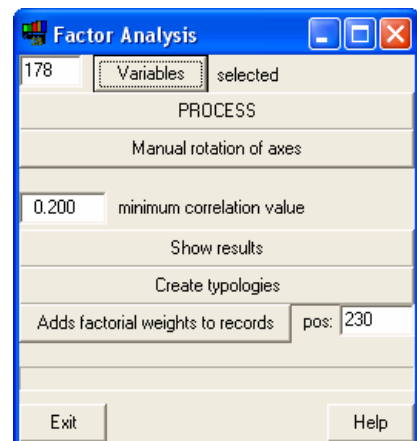
When you choose AND, among the records which are already active, only the records which show the variable which you have selected will be left active. When you choose OR the records which show the selected variable will be added to the already active records. When you choose NOT all the records which show the selected variable will be canceled from the already active records.

Frequency distributions, cross tables and correlation tables

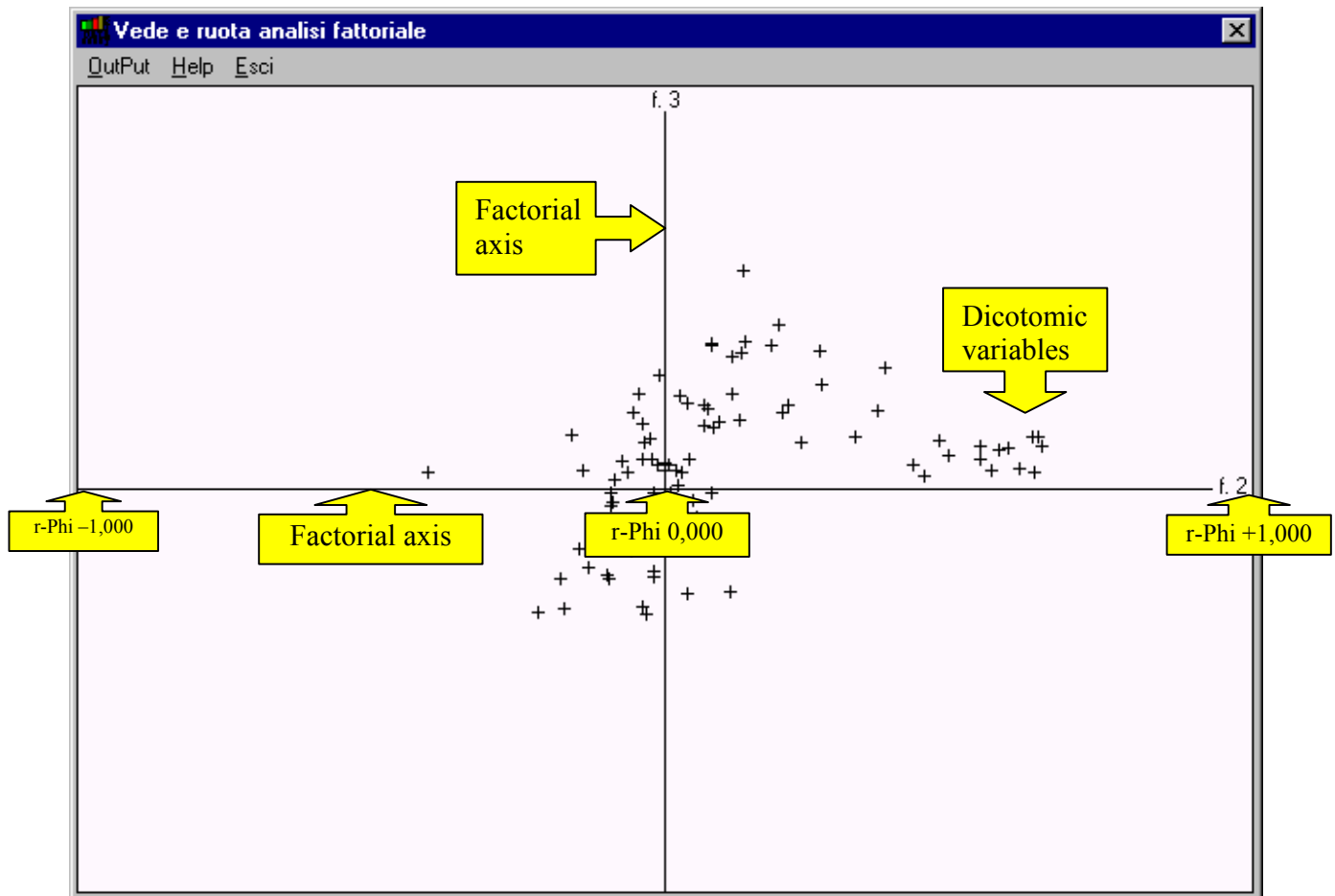
Frequency distributions, cross tables and correlations are produced using windows which are totally similar to those seen in “Data analyses” the only difference is that there is no “Process” button.

Factory analysis

Factor analysis is capable of revealing patterns and structures responsible for the observed single relations. In short, the purpose of factor analysis is to find a new set of variables, fewer in number than the original variables, which express that which is common among the original variables. In medium-large size studies thousands of significant correlations are obtained. When analyzing this large amount of relations it is difficult to find the underlying structures. Factor analysis solves this problem by creating new variables that are not correlated (factorial axes) and for which we know the correlation with the original variables. Factorial axes have the property of not being correlated and therefore each axis intersects the other axes at point 0, at a right angle. This results in a Cartesian space that has as many dimensions as there are factorial axes. The



correlation between the original variables and the factorial axes are used as factorial coordinates. In this way, by intersecting any two factorial axes, it is possible to place every variable on a diagram.

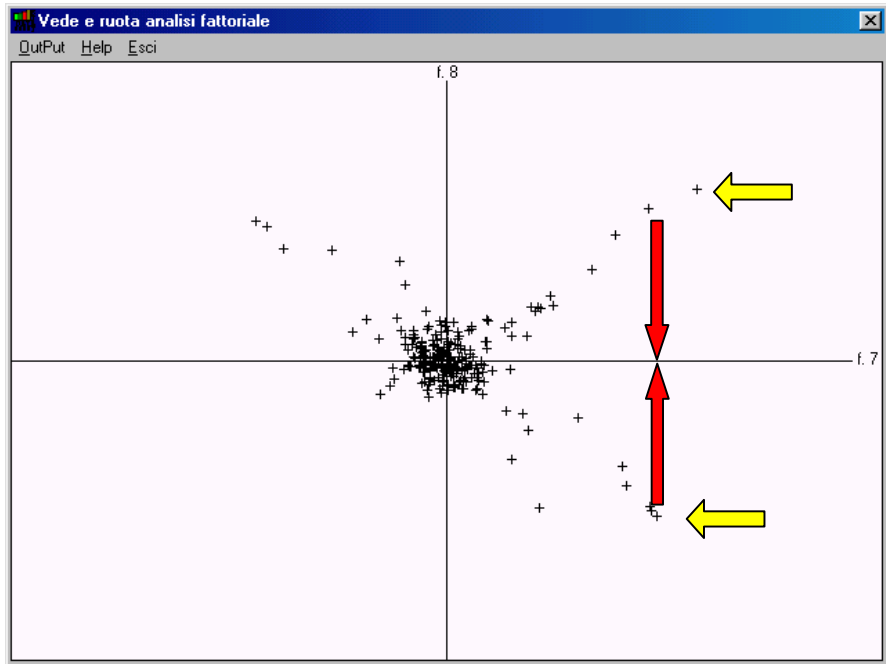


The representation of the original variables on factorial planes permits to delineate structures of variables which are coherent among themselves. These structures are called "factors", as they identify (in general) the "conceptual factors" which describe the phenomena under study.

The geometrical view of factor axes introduces an important topic of factor analysis which is rotation.

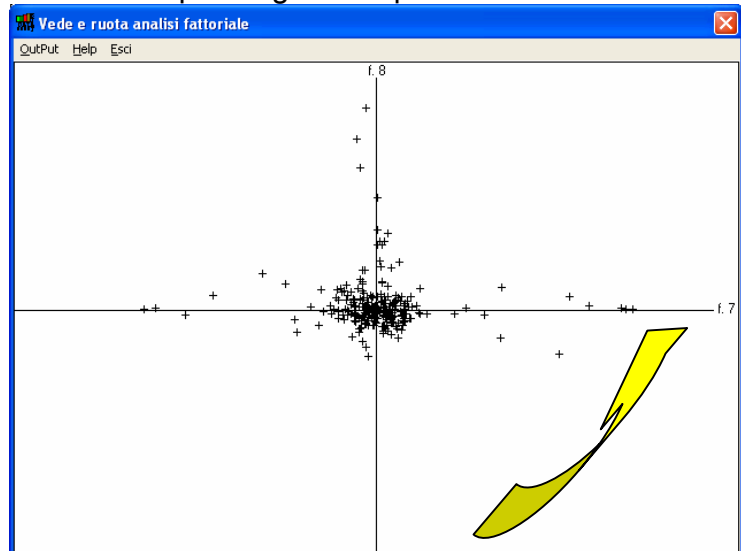
As it is possible to see:

the structures detected by the two yellow arrows at the right are not correlated (they form a 90° angle) but their projection on the horizontal axis coincide (red arrows). In the same way variables that are seen close together on the factorial plane might be at a 90° angle (therefore not correlated) when using another factorial axis. If the factorial planes are rotated in such a way that factorial structures tend to coincide with factorial axis the risk of seeing together variables which are not



correlated is reduced and usually solved. It is therefore necessary to rotate factor axis before reading factor results. Only after rotation the reading of factor results can be considered scientific, otherwise the results will vary each time depending on the position of the axis.

In order to rotate axes choose the "rotate button" which will show the intersection for factor 1 (f.1) and factor 2 (f.2). To increase factor axis click with the left button of the mouse on (f.1) or (f.2) to decrease the factor axis click on the opposite side of the axis. To rotate factors click on the point which you want to rotate on the horizontal axis. It is important to check all the intersections among factors as the automatic rotation which is performed often misses structures which instead should be placed on the axes. Once you have ended rotating factors it is possible to show the results.



Factor analysis results are divided into:

1. variance of factors;
2. variance of variables;
3. factors;

Usually the attention is focused on the third point, the correlation among dichotomic variables and factor axes, but let us describe also the other 2 points:

----- VARIANCE OF FACTORS -----

Fat.	1:	0.506
Fat.	2:	0.178
Fat.	3:	0.174
Fat.	4:	0.142

The sum of the variance of factors is set to 1. In this example the first factor contains more than 50% of the variance explained by factors.

----- VARIANCE OF VARIABLES -----

123 -	0.674	N/11	I am not spontaneous
60 -	0.646	H/4	My father is very affectionate
134 -	0.643	P/1	It is a couple
47 -	0.633	F/4	I feel depressed
69 -	0.630	H/13	My family is very affectionate
16 -	0.628	C/3	Many lonely people
36 -	0.626	E/8	I watch a lot of TV
61 -	0.625	H/5	I trust my mother
4 -	0.621	B/4	It is a crisis of values

It is important to remember that the technique used in Sintropia-DS (Common factor analysis) uses only the common part of variance (covariance) and therefore only part of the variance of each variable is translated into factors.

Common factor analysis treats the system as an "open system", it finds factors by using the common part between the variables (covariance). The solution in common factors results in two spaces: the factorial space of the common factors, and the space of the single factors (the part of each variable which is not explained by common correlations).

General variance: 0.570

Tells how much of the original variance is explained by common factors.

----- Factor: 1

1	0.631	(47)	F/4	I feel depressed
2	0.585	(48)	F/5	I feel useless
3	0.577	(51)	F/8	I don't like my body
4	0.568	(49)	F/6	I feel lonely
5	0.513	(123)	N/11	I am not spontaneous
6	0.487	(50)	F/7	I am pessimist

----- Factor: 2

1	0.475	(61)	H/5	I trust my mother
2	0.469	(62)	H/6	I trust my father
3	0.446	(63)	H/7	I feel relaxed with my parents
4	0.425	(59)	H/3	My mother is very affectionate
5	0.422	(58)	H/2	My father pays attention to my activities
6	0.417	(60)	H/4	My father is very affectionate
7	0.415	(66)	H/10	I have a beautiful relationship with my mother
8	0.396	(67)	H/11	I have a beautiful relationship with my brothers
9	0.383	(64)	H/8	If I have problems I can count on my father
10	0.379	(80)	I/1	I have beautiful memories

Variables are grouped together into factors. It is important to say that these lists can contain errors (due to rotation problems), and it is therefore useful to remove those variables which we consider not coherent with the results.

Factor weights – cluster analysis

Once you select factors which are interesting and remove variables which are not coherent it is possible to attribute a value between 0 and 100 to each record according to how much each factor is present. If you have selected 3 factors, first create 3 quantitative variables at the end of the form, then save the 3 factors with the selected variables in a .txt (text) file and before pressing the button "Add factorial weights to record" enter the position of the first field in which factor loadings will be saved. When factorial loadings have been saved they can be treated as any other variable of the form: used in statistical analyses and used to select records and homogeneous groups (clusters).