

# USER MANUAL

*Statistical Data Graphics & Analysis for ATARI 8 Bits*  
Written for ABBUC contest 2008



## INDEX

	Page
<b>1.- Introduction</b> .....	<b>2</b>
How do I load onto your computer SDG ATARI? .....	<b>3</b>
<b>2.- Variable EDITOR Module</b> .....	<b>3</b>
<b>3.- Statistics descriptive Module</b> .....	<b>8</b>
<b>4.- Regression Module</b> .....	<b>10</b>
<b>5.- Andeva Module (design of experiments)</b> .	<b>12</b>
a) Fixed Treatments with complete randomization .....	<b>12</b>
b) Model full block .....	<b>13</b>
c) Random Model with factorial structure .....	<b>13</b>
<b>6.- Modules for graphics</b> .....	<b>14</b>
<b>7.- Printing Module</b> .....	<b>14</b>
<b>8.- Module for Files Edition</b> .....	<b>14</b>
<b>9.- Summary commands in SDG</b> .....	<b>16</b>
<b>A.- ANNEXES</b> .....	<b>17</b>

# 1.- Introduction

Graphs and statistics (SDG) is a powerful software and easy to use. Through menu you can create and edit variables, manage files, describe variables in statistical form, perform simple and multiple regression, analysis of variance for experimental designs, two-dimensional and three-dimensional graph equations, build scatter diagrams and histograms. It may also generate random variables, transform variables, variables to convert various formats to interact with various programs, create pie charts, bar, etc.

All results are easy to interpret and the user requires only minimal knowledge of statistics.

SDG consists of several modular programs that interact with each other through interactive menus with the user.

All modules are scheduled in TURBO BASIC with some subroutines into machine code that makes some processes more efficient. All programs are unprotected and can be listed and or modified.

With SDG you can graph their monthly expenses, statisticians calculate averages and other complexes, pooling their data in tables and get immediate histograms, to differentiate between products, predicting its sales or expenses, graph complex mathematical functions, and so on.

## **The configuration that requires you to use SDG in a physical ATARI:**

- 1) A computer ATARI with a minimum of 48 Kb RAM.
- 2) A floppy drive ATARI
- 3) A diskette SDG (A and B sides system or two separate floppy disks) with modular programmes.
- 4) A blank floppy disk to store files: variables (data), graphics, tables, etc.
- 5) A printer EPSON RC-220 or equivalent, optional.
- 6) An optional printer plotter ATARI 1020 for three-dimensional graphics.



The configuration that requires you to use SDG in emulation with Atari800Win Plus:

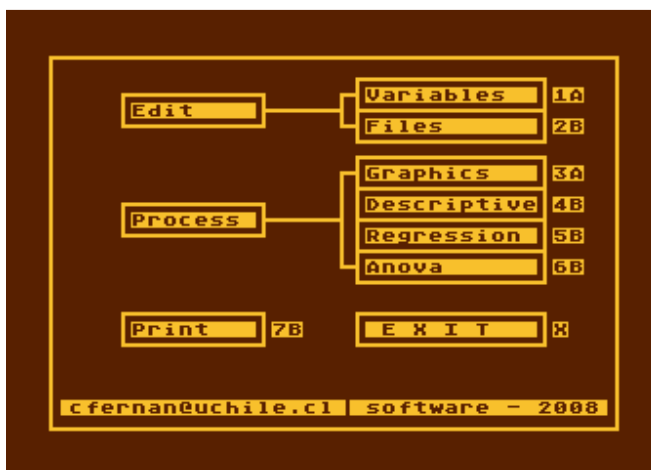
- 1) Load Image `SDG-side-A.atr` in the Disk Drive 1
- 2) Load Image `SDG-Data.atr` in Disk Drive 2
- 3) Reset emulator and disable BASIC. When you start the program, choose Option 2



How do I load onto your computer SDG ATARI?

- 1) Turn on the drive 1 (and 2 drive in case of using two Drives)
- 2) Insert the SDG system disk by side A (or Diskette 1 if the system is recorded in two separate floppy disks).
- 3) Turn on your computer pressing OPTION
- 4) Choose the appropriate option on the screen and follow the instructions below.

## 2.- Variable EDITOR Module



Purpose: To create, modify or add data from a variable. The variables must be created or modified one by one, and before editing another variable, the above is recorded on the disk data.

Once released all relevant variables, the modules can be used for descriptive statistics, graphs, regression, or analysis of variance for design of experiments.

To increase the speed in the edition of variables and allow a greater degree of automation in the subsequent analysis of information, the editor lets you create variables according to their type (discrete or continuous variable), and agree to form that will be transferred to computer (or many variables with little data to be repeated). In general, a variable that is usually discreet is represented by

categories or classes only because it makes no sense to establish ranks continued with their values. The variables are considered continuous if it makes sense to group their values range or ranges. However, even if the variable is discreet, can be classified as if it were ongoing when the number of different values is very high.

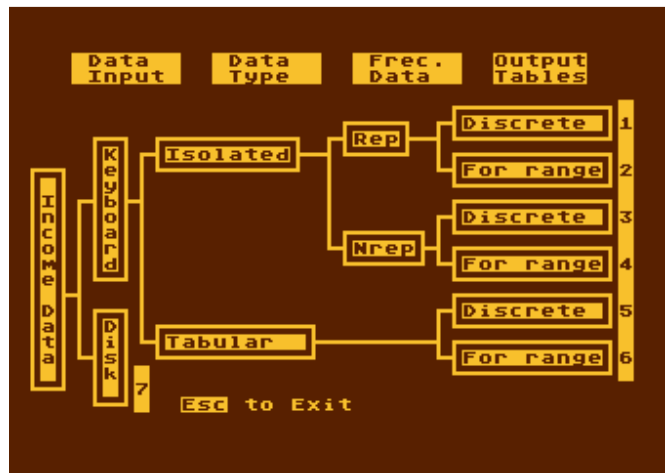
Options 1 and 3 are similar, they differ only in how they are introduced or edit data.

Options 2 and 4 are similar, they differ only in how they are introduced or edit data.

**The SDG EDITOR distinction between:**

a) Data isolated with values that are very repeated in the sample and to be grouped into categories with only tables (no class intervals). In this case the variable is considered Discrete. The publication of such data is done with option 1

Example: Variable daughter. The number of children from 27 families. As we enter the 27 data of this variable (daughter), it is easier to enter using the OPTION 1 program, ie variable discrete values with repeated.



2	2	2	2	3	1	1	2	3	2	1	2	3	2	1	2	3	0	2	2	0	0	0	2	1	1	1
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Note that the value 0 has a frequency of 4, the value 1 a frequency of 7, the value 2 a frequency of 12, and worth 3 a frequency of 4, ie, with many repeated values. 1 with the option of SDG facilitates the entry or editing of this type of variable.

When the descriptive process of variable with SDG (tabulation and calculation of parameters), will be something like this:

Daughter	Frequency	Frec. Acum.	Frec. %	Frec.Acum. %
0	4	4	14.81	14.81
1	7	11	25.93	40.74
2	12	23	44.44	85.19
3	4	27	14.81	100
Totales	27	-	100	-

→ Note that the categories are not intervals (Discrete Variable)

b) Data isolated with values that are not (very) repeated in the sample and to be grouped into categories with only tables (no class intervals). In this case the variable is also considered Discrete. The publication of such data is done with Option 3.

Example: Height. The height of 13 floors (m) measures with a decimal precision. Note that can be described discrete variables even if they have decimals. Although you can also use option 1, it is easier to use option 3

1.6	1.5	1.5	1.7	1.8	1.7	1.0	1.5	1.8	1.0	1.3	1.7	1.7
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

When the descriptive process of variable with SDG (tabulation and calculation of parameters), will be something like this:

HEIGHT	Frecuency	Frec. Acum.	Frec. %	Frec.Acum. %
1.0	2	2	15.38	15.38
1.3	1	3	7.69	23.08
1.5	3	6	23.08	46.15
1.6	1	7	7.69	53.85
1.7	4	11	30.77	84.62
1.8	2	13	15.38	100
Totals	13	-	100	-

↳ Note that the categories are not intervals (Discrete Variable)

Observations: The variable can be discreet when considering their classification in frequency tables does not involve the use of intervals or ranges. The classification tables in a unique class of discrete variables are possible only if the number of values "different" observed in the variable does not exceed 20.

c) Data isolated with values that are repeated in the sample and to be grouped into categories with tables intervals. In this case the variable continue is considered. The publication of such data is done with Option 2. We can also use option 4, but I would write one to one data.

Example: Variable weight. The weight (kg) of 40 persons:

45	45	45	50	50	52	40	50	48	48	38	45	50	40	50	45	45	50	37	37
45	37	37	39	50	37	42	48	45	45	45	50	52	37	37	45	45	45	48	38

When the descriptive process of variable with SDG (tabulation and calculation of parameters), will be something like this:

WEIGHT (kg)	Average	Frecuency	Frec. Acum.	Frec. %	Frec.Acum. %
Lim. Inf. - Lim. Sup.					
$36 \leq x \leq 39$	37.5	10	10	25.00	25.00
$39 < x \leq 42$	40.5	3	13	7.50	32.50
$42 < x \leq 45$	43.5	13	26	32.50	65.00
$45 < x \leq 48$	46.5	4	30	10.00	75.00
$48 < x \leq 51$	49.5	8	38	20.00	95.00
$51 < x \leq 54$	52.5	2	40	5.00	100
Totals	-	40	-	100	-

→ Note that the categories are expressed by intervals (continuously variable). SDG automatically calculates the required number of intervals and lower and upper limits for each category.

- d) Data isolated with values that are not (very) repeated in the sample and to be grouped into categories with tables intervals. In this case the variable is considered Continuing. The publication of such data is done with Option 4. It also could use the option 2 but would have to write one to one data over a frequency value of one for each input.

Example: Variable height (in centimeters) of 30 persons:

146	155	145	167	178	177	160	145	178	180	173	157	177	165	171
148	175	165	176	149	150	162	163	166	147	151	154	177	157	152

When the descriptive process of variable with SDG (tabulation and calculation of parameters), will be something like this:

HEIGHT (cm)	Average	Frecuency	Frec. Acum.	Frec. %	Frec.Acum. %
Lim. Inf. - Lim. Sup.					
$145 \leq x \leq 151$	148	8	8	26.67	26.67
$151 < x \leq 157$	154	5	13	16.67	43.33
$157 < x \leq 163$	160	3	16	10.00	53.33
$163 < x \leq 169$	166	4	20	13.33	66.67
$169 < x \leq 175$	172	3	23	10.00	76.67
$175 < x \leq 181$	178	7	30	23.33	100
Totals	-	30	-	100	-

Note that the categories are expressed by intervals (continuously variable). SDG automatically calculates the required number of intervals and lower and upper limits for each category.

e) The data come from a table with discrete classes or categories. In this case use the Option 5.

Example: Classification of the number of children in 100 families:

Number of children	Number of families (frequency)
Clase 1: 0	12
Clase 2: 1	33
Clase 3: 2	30
Clase 4: 3	15
Clase 5: 4	8
Clase 6: 5	2

f) The data come from a table with classes or categories of intervals. In this case the use OPTION 6

Example: Classification of salaries of 300 employees.:

SALARIES		Number of employees (frequency)
Lim. Inf.	Lim. Sup.	
Clase 1:	$0 \leq x \leq 20000$	100
Clase 2:	$20000 < x \leq 40000$	80
Clase 3:	$40000 < x \leq 60000$	70
Clase 4:	$60000 < x \leq 80000$	30
Clase 5:	$80000 < x \leq 100000$	20

Observations: The module Descriptive Statistics processed properly only when all these tables have equal intervals breadth. The upper limit of the intervals is included in each class, however the lower limit of the intervals is excluded from each class, except in Class 1.

Options 1, 2 and 3, 4 respectively, are equivalent and differ only in how you can enter data.

When the module is used descriptive statistics to analyze the data, must be chosen carefully between options 1, 2, 3, 4, 5 or 6, because it depends on which tables are incorporated into classes resulting unique or Ranges (Intervals).

The modules to build chart, making regressions, or to design experiments, they can only use variables entered with options 1, 2, 3 or 4.

Options 1 to 6 of the editor of variables are used only to enter variable (data) through the keyboard.

Option 7 is used to edit variables taped drive through any of the options described above,

Note: If you mistakenly entered the publisher, to leave him with enough pressure <Esc>.

Example: Enter through the editor the following variables (will be used later to exemplify processes). As there are few data and are not repeated, use option 3 of the publisher. There are two variables: An independent variable YEARS (YEARS) and a dependent variable production (PROD):

Var1: YEARS	2000	2001	2002	2003	2004	2005	2006	2007
Var2: PROD	5	10	165	25	38	52	68	84

To edit the variables above, follow these steps:

- 1) Turn on the drive Atari 1050
- 2) It floppy system by the side A
- 3) When the home screen of SDG, 1 choose whether their data are saved on a diskette that will drive the 1 (exchange with the floppy System), or 2 if their data is saved on a diskette that will drive the 2 (Another floppy 1050).
- 4) When displayed the main menu, choose Option 1 (EDITOR variable).
- 5) Choose option 3 Editor Variables (Keyboard - Discretos - not to repeat).
- 6) Enter the number of data containing the variable: 8.
- 7) Enter the data for one to one variable YEARS:  
 2000<return>  
 2001<return>  
 2002<return>  
 2003<return>  
 2004<return>  
 2005<return>  
 2006<return>  
 2007<return>
- 8) Check your data and verify they are correct.
- 9) Accept the income of the variable (Press A)
- 10) Insert the disk data and press <return>.
- 11) If you have not created variables that floppy, the directory will appear empty. Press "E" to finish watching the directory.
- 12) Enter the name of the variable: write YEARS and press <return>.

Repeat entire process prior to the variable PROD (Production). Finally, return to the main menu by pressing <return>.

### 3.- Statistics descriptive Module

STATISTICS OF POSITION AND DISPERSION	
ARITHMETIC AVERAGE	1.59259259
GEOMETRIC AVERAGE	0
MEDIAN	2
QUARTILE 1	1
QUARTILE 3	2
MODE 1	2
MINIMUM VALUE	0
MAXIMUM VALUE	3
RANGE	3
MEAN DEVIATION	0.8091168096
VARIANCE	0.83401922
STANDARD DEVIATION	0.9132465275
VARIATION COEFFICIENT	57.3433867
SLOPE OF PEARSON	-0.4461089067
KURTOSIS PERCENTILIC	-0.0963333334
1TB   2HP   3HA   4PM   5PE   6SV   7PR   8EX   <-->	

This module can describe any statistically variable hospitalized with options from 1 to 6 Editor variables. Handing the following information:

- a) Table of frequencies with complete and absolute frequency percentage, and accumulated partial .



If the variable is discreet, the maximum number of categories or classes is 20. When the variable is continuous, the number of intervals is calculated according to the Rule of Sturges, who said:

Number of intervals =  $1 + 3.3 * \text{Logbase10}(N)$ , where N is the number of isolated data.

Example: If  $N = 200$ , then number intervals =  $1 + 3.3 * \text{Logbase10}(200) = 9$

- b) Histograms of frequency partial and cumulative.
- c) Position parameters: Arithmetic mean, Geometric mean, median, mode (one or more), Percentiles, Minimum and maximum value.
- d) Dispersion parameters: Variance, Standard deviation, rank, and coefficient of variation.

The variance calculated corresponds to that of the population, and the Standard deviation is the square root of the variance earlier. The coefficient of variation is calculated as the standard deviation divided by the arithmetic mean.

e) Bias Pearson

f) kurtosis percentile

At the end of this manual specifies all the formulas used by SDG.

Examples:

Studying the behavior of descriptive:

- a) the variable PROD defined anteriormente with SDG
- b) The variable X as defined below.

Use Editor variables (option 4) to enter the following values of variable X:

12.23	13.00	18.07	15.36	17.23	12.22	22.34	26.21	18.78	11.09
15.22	18.76	12.25	14.56	34.24	17.23	16.33	16.00	18.45	23.67

For both variables, once recorded, insert the floppy on the side B and choose Option Descriptive (4B). Then follow the instructions on screen. When asked if he recorded the results and you answered yes, the original data will be recorded and ordered not ordered (as admitted). These data will be recorded as a screen and can only be read later with the printing module (7B).

When the results appear on screen Module descriptive, they will be distributed over six screens, the first two for the table of frequencies, for the following parameters, and the last three for Percentiles.

The menu commands appears on the bottom line of the screen, and allows the following:

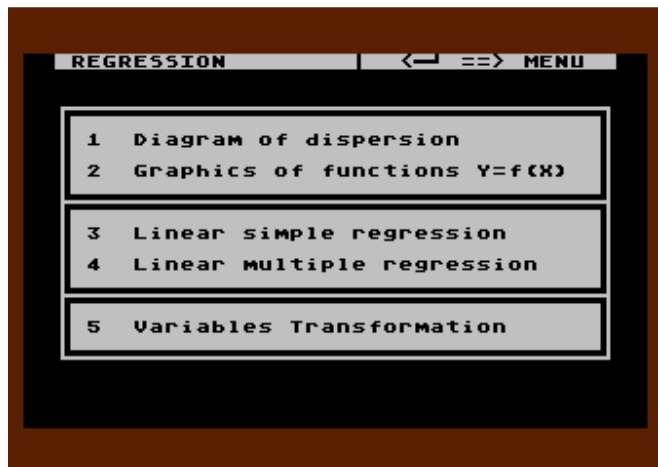
- 1) Show from the beginning the table frequencies.
- 2) Show Histogram Frequency Partial
- 3) Show Histogram Frequency Cumulatively
- 4) Show Parameters
- 5) Show the first screen Percentiles
- 6) Record six screens described above.
- 7) Print the reultados.
- 8) Exit to Main Menu

The left and right arrows are used to move half of screens to the left or right respectively.

To print or save the results, previously always press Option 1.

## 4.- Regression Module

The module is loaded from the Main Menu B-side of the disk system. Insert the disk data and press <return>, you'll see that in diquete data are recorded by the two variables you: YEARS and PROD. Press "E" to finish watching the directory and enter as it prompted the dependent variable and the independent variable PROD YEARS. Then choose the menu screen the regression model most appropriate, or adjust all models and then pick the best of them according to the results (Coefficient of Determination and analysis of variance).



While it can be done using multiple regression models simple, you can transform any variable (add a constant, calculate logarithms, etc.). And then perform regression with the variables processed. For example, if you want to adjust the model is  $Y = A + B / \text{SQR}(X)$  transforms the variable X as follows: first calculate the square root of the original variable and save this transformation with some name in the floppy data. Then transform the variable created previously by calculating the reciprocal of it.

Notación	X	→	SQR(X)	→	1/SQR(X)
Guardar como	Var1	→	Var2 (o var 1)	→	Var3 (o var2 o var 1)

Later, make a linear regression of the dependent variable "Y" on the new variable "1/SQR(X)".

Virtually all models of equations are linearization through transformations. The same method is applicable to the case of multiple regression.

Observation: To record the results of regression, pre-press option 1 which rewrites the screen and press option 6.

Also you can see the chart all the sampling points and adjusted model with options 3, 4 or 5. Only you can save or print figure high resolution obtained with the option 5.

The estimated values of the equation adjusted obtained with option 2.

The graphics produced by the form of regression (regression simple charts and diagrams scatter graph functions) while increasing its resolution greater the time remaining on the screen.

#### Graphics produced by regression simple:

**SELECT:** amending intensity of background color.  
**OPTION:** changes color intensity points.  
**START + SELECT:** chart records.  
**START + OPTION:** graphic prints.  
**START:** back to previous screen.

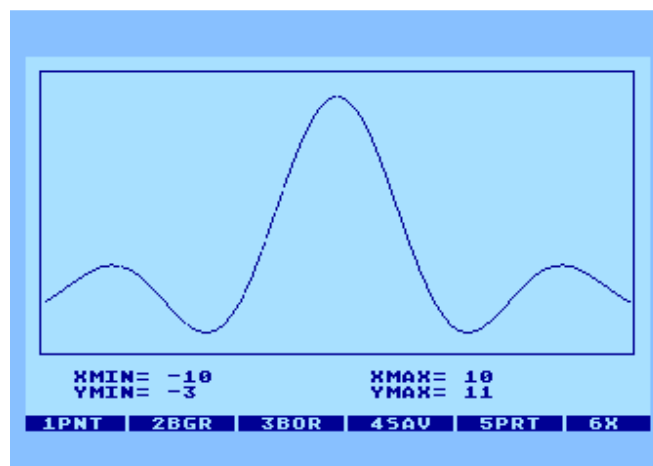
Graphics functions and scatter diagrams:

When the graphic has reached the desired resolution pressures START and you'll see a menu on the bottom line of the screen that lets you:

**1, 2 and 3: modify the intensity of background color, and border points respectively.**  
**4: graphic record.**  
**5: graphic print.**  
**6: leave the graph.**

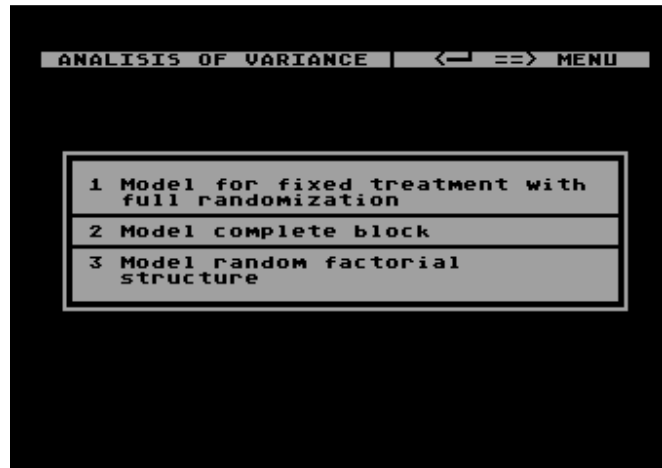
The possibilities are many delivery this module, and only practice with the different options will allow you to take full advantage. All screens are sufficiently interactive with the user.

All programs are modular unprotected, so you can review and modify some of its characteristics. Any problem that has to run the programs, you can reset with SYSTEM RESET and re-run with RUN.



## 5.- Andeva Module (design of experiments)

The module calculates and constructs the table Analysis of Variance (Andeva) to study the existence of significant differences between different treatments. You have the option of bringing their design of experiments to one of the following models:



### a) Fixed Treatments with complete randomization.

This model is used when the treatments were randomly assigned to the experimental units.

Example: Suppose you have four treatments (A, B, C and D) relating to four different drugs to combat disease. The drugs are tested in groups of people, repeating the process five times. The response of the drug was measured by calculating the percentage of people who Sanaa. The results were as follows:

Treatment	A	B	C	D
	40	50	80	30
% of people who heal (5 repetitions for each drug)	30	70	80	40
	20	60	90	50
	10	40	50	50
	50	30	100	30

The Analysis of Variance helps us to determine with any degree of confidence (as proof F) if possible to say that the effects of drug A, B, C or D are significantly different.

MODEL TOTALLY RANDOM			
ANALYSIS OF VARIANCE			
F.V.	g.l.	S.C.	C.M.
Trat	2	1579.2667	789.63335
Err.	27	5458.2	202.155555
Tot.	29	7037.4667	

F Calculate= 3.90606802

1 SAVE    2 PRINT    3 EXIT

To solve this problem, you must enter each of the treatments through EDITOR variable SDG. In other words, you must create 4 variable data with 5 each. This can be used for instance option 3 Editor.

After creating the variables in the data disc, go to the Main Menu, where Choosing 6, which bear the form of analysis of variance. Then choose

option 1 of the following menu that appears on screen (Model treatments fixed completely randomized). Put your floppy and enter data from one to one variable (treatments) released earlier.

- b) **Model full block.** When there is no uniformity for all experimental units, can be separated into blocks, each with characteristics uniform. So for example, a soil fertility of different stripes, each of the strips to form a bloc passes, and within each bloc, the experimental units are assigned randomly.

Example

Blocks	Treatments				
	1	2	3	4	5
Block I	2	4	12	9	3
Block II	3	6	14	6	6
Block III	4	8	15	4	3
Block IV	3	10	11	5	4

Similar to previous point, each treatment (1 to 5) should be paid in advance with Editor Variables, and then use the module Andeva.

- c) **Random Model with factorial structure.** When each treatment consists of the variation of 2 or more factors, the design has factorial structure. By example, if one takes 2 factors, the first with two different levels and the second with three levels, treatments will be trained  $2 \times 3 = 6$ .

RANDOM MODEL FACTORIAL			
ANALYSIS OF VARIANCE			
F.V.	g.l.	S.C.	C.M.
Trat	3	1579.2667	526.422233
A	1	-4505.7333	-4505.7333
B	1	-8871.6333	-8871.6333
AB	1	14956.6333	14956.6333
Err.	36	5458.2	151.616666
Tot.	39	7037.4667	
F Calculated treatm= 3.47206047			
F Calculated fact A= -29.71792889			
F Calculated fact B= -58.51357594			
F Calc. factor AB= 98.64768626			
1 SAVE 2 PRINT 3 EXIT			

Example:

Suppose that a trial is conducted with two fertilizers (A and B) to test its effectiveness in separately and together (interaction). Using three doses for each fertilizer (A1, A2, A3 and B1, B2, B3), the

combination therapies will be formed among them: A1B1, A1B2, A1B3, A2B1, A2B2, A2B3, A3B1, A3B2 and A3B3 ( $3 \times 3 = 9$  treatments). Suppose you are four replicates per treatment, and results in production are as follows:

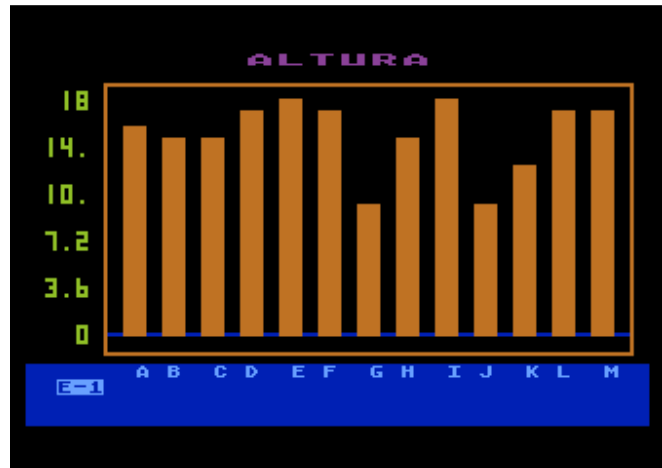
	A1B1	A1B2	A1B3	A2B1	A2B2	A2B3	A3B1	A3B2	A3B3
PRODUCTION	2	4	4	4	5	4	4	6	10
(ton/acre). 4	3	3	6	3	3	8	3	5	4
repetitions per	1	2	4	2	7	4	5	8	8
treatment	2	5	2	1	5	8	4	7	10

The 9 treatments are paid in advance as variables through Editor Variables, and then uses the module Andeva.

## 6.- Modules for graphics

The graphics are the most common statistical bar (simple and grouped) and circulars (sectoral or cakes).

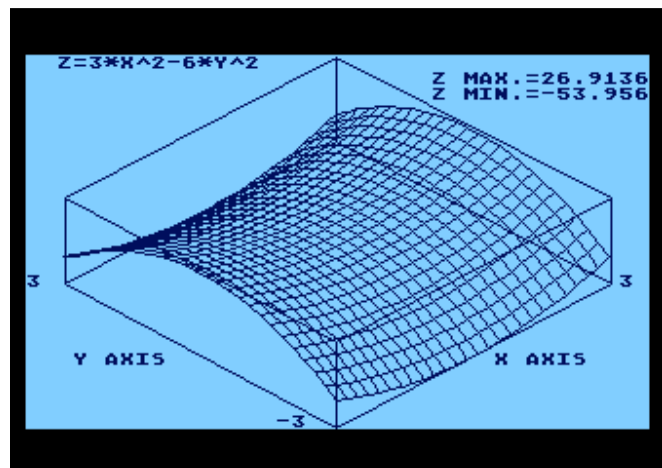
Any variable can be created with SDG graficada with this module, provided it is within their capabilities. The pie charts require a maximum of one variable, however, the bar allowed up to three variables (all with the same number of data for each figure represents a bar). These graphs can be recorded and printed on paper.



This module also allows the variables are entering through the keyboard directly without passing through the publisher of variables.

Another interesting option is to graph functions of three variables in three-dimensional shape.

Observations: These graphics capability are not creción the author of SDG. The bar charts and circulars come from the software "GRAFIQUELO", and three-dimensional graphics are known as "Z\_PLOTTER". These programs have been adapted to use modular SDG.



## 7.- Printing Module

It serves to recover on the screen and / or any file printer (graphic or not) product of the process of recording them from any form of SDG.

## 8.- Module for Files Edition

Lets see any floppy directory, rename, delete, protect or unprotect files. Also used to format disks. In other words, it avoids having to leave the DOS to conduct these operations.

It also has another option to convert the variables created with SDG in other formats of the same SDG or to make them interchangeable with other programs. This option allows the following:

- Make variables with DIF format to format SDG
- Make variables with SDG format to format DIFF
- Make variables with SDF format to format SDG
- Make variables with SDG format to format SDF
- Make variables with SDD format to format SDC
- Make variables with SDC format to format SDD

Format:

SDG: Format used by the program itself.

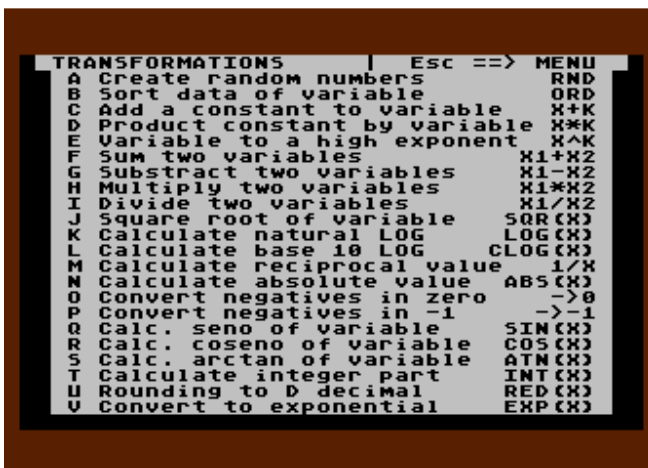
DIFF: Data Interchange Format.

SDF: Standard Data Format (Column pure data)

SDD: Discrete Variable created with the Editor SDG

SDC: Variable continua created with the Editor SDG

The transfer to DIF format makes the variables created with SDG can be used by other programs like VisiCalc for example. They can also create data with other programs and through the DIF format, can be caught by SDG.



## 9.- Summary commands in SDG

Click 1	Click 2	Result in the Editor of Variables
1	1	Create, edit and burn discrete variables with values that are repeated.
1	2	Create, edit and burn continuous variables with values that are repeated.
1	3	Create, edit and burn discrete variables with values that are not repeated constantly.
1	4	Create, edit and burn continuous variables with values that are not repeated constantly.
1	5	Edit and recorded discrete variables that are tabulated in frequency tables to be processed by the module Descriptive Statistics.
1	6	Edit and recorded discrete variables that are tabulated in frequency tables to be processed by the module Descriptive Statistics.
1	7	Edit any variable disk previously recorded through Module Editor Variables SDG.

Click 1	Click 2	Click 3	Result in the File Editor
2	D	-	Displays directory of any disc
2	B	-	Clears files of any disc.
2	R	-	renames files from one disk
2	P	-	Protects files from one disk
2	E	-	Unprotected files from one disk
2	C	1	Converts variables with DIF format to format SDG
2	C	2	Converts variables with SDG format to format DIFF
2	C	3	Converts variables with SDF format to format SDG
2	C	4	Converts variables with SDG format to format SDF
2	C	5	Converts admitted as discrete variables to be processed by the processor as continuing Description of SDG
2	C	6	Converts admitted as a continuous variable to be processed by the processor as discrete Description of SDG

Click 1	Click 2	Click 3	Result in Gráphics
3	D	B	Create Bar graphs, capturing data variables recorded in Disk.
3	D	C	Create pie charts, capturing data variables recorded in Disk.
3	T	B	Create Bar graphs, entering data directly from the keyboard.
3	T	C	Create pie charts, entering data directly from the keyboard.
3	T	T	Create thridimensional graphics to enter with the keyboard the function of three variables.

Click 1	Result in Graphics
4	Process the variables entered with descriptive statistics.



Click 1	Click 2	Click 3	Results in Graphics, Regressions and transformations.
5	1	-	Graf Dispersion Diagram between two variables recorded in Disk
5	2	-	Graf functions of two variables entered with the keyboard.
5	3	-	Make simple linear regression between two variables recorded on the Data Disk.
5	4	-	Make Multiple Linear regression between 2, 3, or 4 variables recorded on the Data Disk.
5	5	A	Generates variables with random data.
5	5	B	Generates another variable ordering growing data in the form of a variable saved on disk.
5	5	C	Generates a variable adding a constant to another variable of data recorded on disk.
5	5	D	Generates a variable multiplying a constant per another variable of data recorded on disk.
5	5	E	Generates another variable bringing to power a variable data recorded on disk.
5	5	F	Generates another variable adding one to one of the data of two variables recorded on disk.
5	5	G	Generates another variable by subtracting one-one data recorded on disc of two variables.
5	5	H	Generates another variable multiplying one-one data of two variables recorded on disk.
5	5	I	Generates another variable dividing one-one data of two variables recorded on disk.
5	5	J	Generates another variable by calculating the square root of the data variable saved in a disk.
5	5	K	Generates optionally another variable calculating the natural logarithm of data variable saved in a disk.
5	5	L	Generates optionally another variable calculating the Briggs logarithm of data variable saved in a disk.
5	5	M	Generates another variable by calculating the reciprocal of the data variable saved in a disk.
5	5	N	Generates another variable by calculating the absolute value of a data variable saved on disc.
5	5	O	Generates another variable becoming zeros the negative values of a variable saved to disk.
5	5	P	Generates another variable becoming -1 the negative values of a variable saved to disk.
5	5	Q	Generates another variable calculating Sin(x) to a variable data recorded on disk.
5	5	R	Generates another variable calculating Cos(x) to a variable data recorded on disk.
5	5	S	Generates another variable calculating ArcTan(x) to a variable data recorded on disk.
5	5	T	Generates another variable calculating the whole party to a variable data recorded on disk.
5	5	U	Generates another variable rounding to the decimal specified recorded in disk.
5	5	V	Generates another variable calculating Exp(x) to a variable data recorded on disk.

## ANNEXES

Reference on some formulas used by SDG:

### A) Descriptive Statistics

$$\text{Arithmetic Mean} \quad \mu = \frac{\sum_{i=1}^n X_i}{n}$$

$$\text{Median} \quad Me = L_i + \frac{\frac{N}{2} - F_{(i-1)}}{f_i} \cdot A$$

$$\text{Mode} \quad Mo = L_i + \frac{f_i - f_{(i-1)}}{(f_i - f_{(i-1)}) + (f_i - f_{(i+1)})} \cdot A$$

$$\text{Geometric Mean} \quad Mg = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdot \dots \cdot x_n}$$

$$\text{Variance} \quad \sigma^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{n}$$

$$\text{Standard Deviation} \quad \sigma = \sqrt{\sigma^2}$$

$$\text{Coefficient of Variation} \quad CV = \frac{\sigma}{\mu}$$

$$\text{Pearson Bias} \quad S = \frac{\mu - Mo}{\sigma}$$

$$\text{Percentilic Kurtosis} \quad K = \frac{P_{75} - P_{25}}{2(P_{90} - P_{10})} - 0.263$$

$$\text{Percentile} \quad P_r = L_i + \frac{\frac{r}{100}n - F_{(i-1)}}{f_i} A$$

Range	$R = (\text{Higher data}) - (\text{Smaller data})$
-------	----------------------------------------------------

Other measures of position can be obtained from percentiles, for example:

Quartile 1	$Q_1 = P_{25}$
------------	----------------

**Notation:**

$x_i$  : values of a variable (*data*) :  $x_1, x_2, x_3, \dots, x_n$

$n$  : Number of data in the variable

$L_i$  : Lower Limit in the  $i$ -interval of frequencies table.

$F_{(i-1)}$  : absolute frequency accumulated in the interval  $(i - 1)$

$f_i$  : absolute frequency parcial in the interval  $(i)$

$f_{(i-1)}$  : absolute frequency parcial in the interval  $(i - 1)$

$f_{(i+1)}$  : absolute frequency parcial in the interval  $(i + 1)$

$A$  : Size intervals of the frequencies table

**A) Regression**

If the model  $\mathbb{Y} = \alpha + \beta x + \epsilon$ , then  $a$  and  $b$  are estimators of  $\alpha$  and  $\beta$  respectively. The coefficients  $a$  and  $b$  of the equation  $Y = a + bx$  obtained using the method of least squares. Thus:

$$b = \frac{n \sum_{i=1}^n (x_i y_i) - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \qquad a = \bar{Y} - b \bar{X}$$

In a similar manner calculated parameters in multiple linear regression.

The variance analysis for regression is used to determine the extent to which the regression model used explains the phenomenon being studied across the variables used. It also measures the significance of each variable.

**The value of F calculated to be compared with the value of  $F$  Snedecor in the distribution of probabilities F given in statistical tables in books.**

**We must employ the necessary degrees of freedom (given in the analysis of variance), and a confidence level desired (90, 95 or 99%).**

**If the value of F calculated on the analysis of variance is greater than the F to the table of probabilities, then the ratio is significantly different from zero, ie should be part of regression model.**

**The variance analysis to design experiments is studied in a similar manner, except that in this case is discussed if there are significant differences between treatments, or between factors if the design has factorial structure. In the latter case shows whether or not there is significant contributions of the various factors and their interactions in the response observed in the experiment.**