Introduction

How to Use this Document

This document is intended to introduce prospective researchers to SYSTAT for Windows, which currently runs under Windows 95/98/NT 4/2000/XP. Those who do not own a personal copy of SYSTAT for Windows may access the software from various Student Technology Centers at IU. Vendor-supplied documents are available for reference at the Swain Hall, Business/SPEA and Education Libraries' Reserve Collections, IUPUI University Library, and at the UITS Center for Statistical and Mathematical Computing (Stat/Math Center). Contact the UITS Stat/Math Center if you are interested in leasing SPSS for Windows to install on IU owned PCs at an educational discount.

What is SYSTAT?

SYSTAT is a comprehensive, general purpose, easy to use, and highly integrated statistical software package most popularly used from microcomputers. SYSTAT includes basic statistics (e.g., descriptive statistics, frequencies, crosstabs, correlations) and advanced statistics (e.g., regression, ANOVA, MANOVA, factor analysis, cluster analysis, discriminant analysis, time series). SYSTAT has an excellent graphics interface for data visualization, is quick and easy to use for exploratory work, and has extensive options for presentation graphics.

SYSTAT for Windows is a full-fledged Windows application; just point and click. You can do most of your work through menus and dialog boxes with SYSTAT for Windows. SYSTAT's Quick buttons give you simple, single-click short cuts to common statistical analysis. In SYSTAT for Windows, each type of task happens in its own window -- Main, Data, and Graph. Each Window has its own menu bar with submenus. The Window menu controls the windows used in SYSTAT. To switch windows, or to make a window active you may select the window of your choice from the Windows' taskbar at the bottom of the screen. SYSTAT output can be viewed onscreen and cut, copied and pasted to other Windows applications. Results and graphs can be edited. You can select from a wide range of fonts, character sizes, and colors. Data may be entered into a spreadsheet-like editor. You can open your data into SYSTAT from free or fixed-format ASCII files, or from spreadsheet or database files in Excel or dBase formats. You can even read SPSS or BMDP system files directly. You can also export files to the same formats.

Hardware and Software Requirements

- Microsoft Windows 95/98/NT 4/2000/XP
- Pentium/clone or above
- 64 MB RAM (minimum)
- 90 MB hard disk space
- CD Drive required for installation of program
- SVGA monitor

How to start a SYSTAT session

To start a SYSTAT session under Windows XP, from any of the UITS Student Technology Centers:

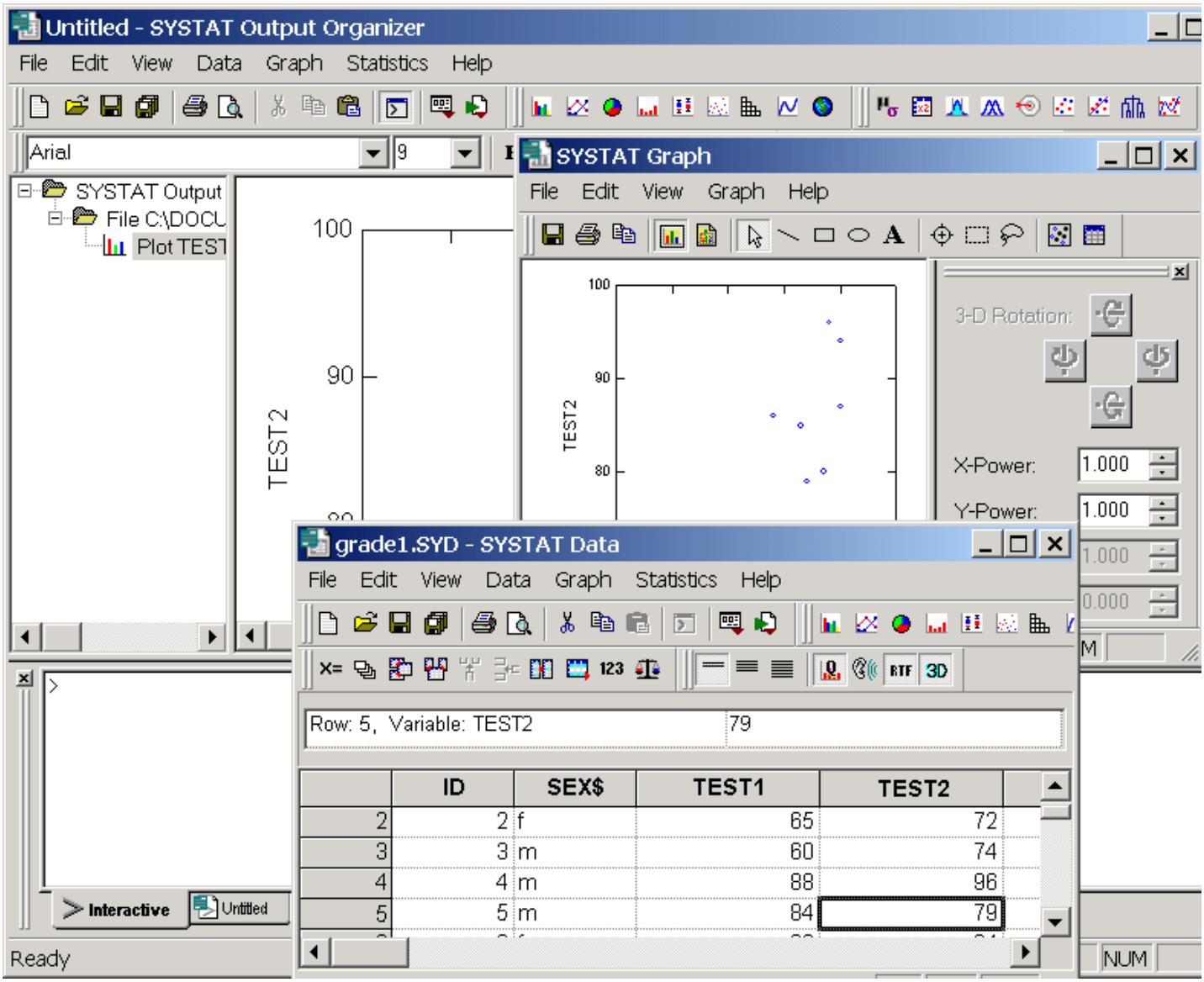
1. Log on to an available workstation
2. Click the Start button
3. Click and drag Programs --> Statistics and Math --> Systat 10.2.01

Now you are ready for SYSTAT computing.

Orientation

The Interface

There are three types of windows in SYSTAT:

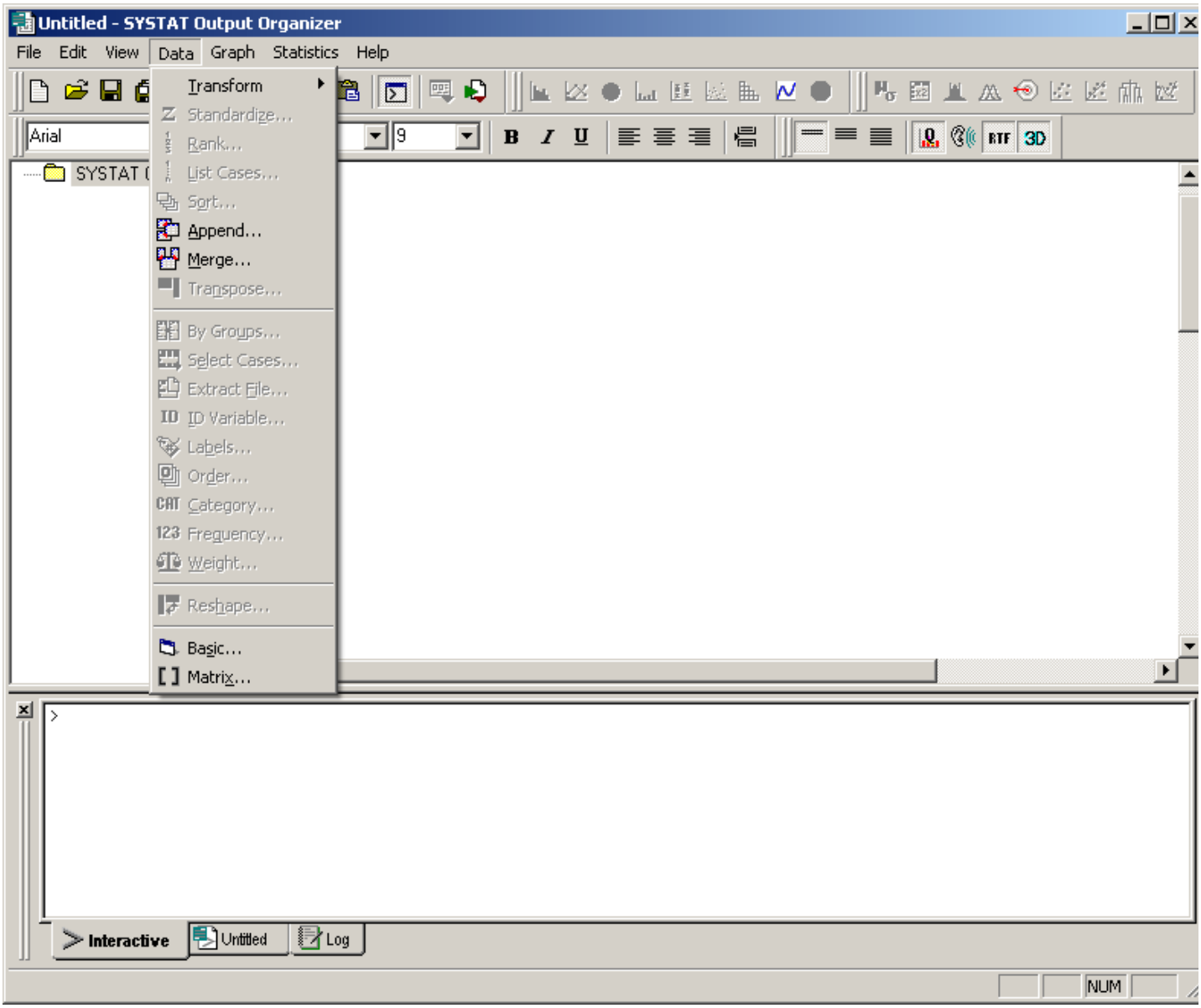


Main window
: The main window is divided into three panes: Output Organizer, Output pane, and Command pane. Output Organizer allows you to navigate through the results of your statistical analysis. Results of the analyses you request with the Statistics and Graph menu items appear in the Output pane. The Command pane is the area for running SYSTAT via its command language.

Data window
: The Data window displays your data in a row-by-columns format. You can use the Data window's Edit menu to cut, copy, delete, and paste rows, columns, and blocks of data. Also, you can use the Data window's Data menu to transform data, search for cases satisfying certain criteria, and select subsets of cases.

Graph window
: Double-clicking a graph in the Output pane opens the Graph Window. 3-D graphs can be rotated with the Dynamic Explorer from View menu using Graph window toolbar.

Menus in SYSTAT for Windows



The top strip of each of SYSTAT's window is a menu bar that contains menus of things you can do.

Each window has its own menu bar with selections appropriate to that window. To use a menu item, click on the menu name, and then click on the item you want. SYSTAT has two most important menus: Graph and Statistics. The Graph menu contains items for producing graphical displays. The Statistics menu contains SYSTAT's statistical analyses.

Customizing your SYSTAT session

Often, you may want to change some of the default option settings of SYSTAT. Select Options from the Edit menu to open the Global Options dialog box. The five tabs-General, Data, Output, Graph and File Locations-control different settings in SYSTAT.

The example outputs in this document are in classic SYSTAT output style which is in ASCII format. SYSTAT generates output in a Rich Text Format (RTF) as default. If you prefer the classic style, you can click on the Output tab of the Global Options dialog box, check the box in front of Use SYSTAT Classic output style, and click on OK. The next output generated will be in the classic output style.

Help during SYSTAT computing

Help is available from each window in SYSTAT for Windows. Select Contents or Search for Help from the pop-up Help menu for a list of topics. Also, The IU Knowledge Base provides answers to the most frequently asked questions. Also, each SYSTAT dialog box has a help button that shows you how to use that dialog box. Clicking Command from a dialog box Help window will provide help on a related SYSTAT command. For help on commands from the command prompt, type

HELP *commandname*

Data Analysis

Organizing your data for analysis

Suppose you have three test scores collected from a class of ten students (5 females, and 5 males) during a semester. The information you have for each student is: identification number, gender, score for test one, test two and test three (the full data set is displayed toward the end of this section for you to enter into the Data window).

Your first task is to present the data in a form acceptable to SYSTAT for processing. Before showing you how to enter the data, let us look at what SYSTAT accepts as data.

SYSTAT uses data organized in rows and columns. The rows are called **cases**, and the columns are called **variables**.

Name$	Test1	Test2	Test3
Tim	20	23	24
Hans	21	26	28

A case contains information for one unit of analysis, e.g., a person, an animal, a product, a business. Variables are information collected for each case, such as name, score, age, etc. In the above example there are two cases and four variables. When data are arranged in rows and columns like this and is stored in a file, it is called a 'cases-by variables' or 'rectangular data file.'

SYSTAT accepts numbers and characters as data. In the above example, Test1 is a numeric variable, and Name$ is a character variable. The numbers stored in numeric variables can have up to 15 digits. You can use a negative sign (-) for negative numbers. A character value stored in character variables can have up to 12 characters and can include letters (a-z, A-Z), or any typewriter characters (e.g., ()$&*/). If you include numbers within character values they will be treated as characters. Also, upper and lower case character values are differentiated (e.g., JUNIOR is not the same as junior).

You must assign a unique name for each variable. Variable names may contain up to 12 letters or numbers, and must begin with a letter. The names of character variables must end with a dollar sign ($); the dollar sign counts as one of the 12 letters. Variable names, unlike character values, are not case sensitive. Once you have entered a variable name, you may change it, but you cannot change its type from character to numeric or vice versa. If you forget to put a $ sign at the end of a character variable name, you must type the correct name as a new variable in a new column and later delete the incorrect variable.

When a numeric data is missing, enter a period (.) to flag the position where the value is missing. When a character data is missing enter a blank space surrrounded by single or double quotation marks (' 'or " "). These quotation marks will not show up in the spread sheet. Note that arithmetic involving missing values propagates missing values.

There is no limit to the number of variables or cases your worksheet can contain, but it is limited to the availability of memory or space on the computer.

With the above information in mind, let us assign names to the variables in our example: id, sex, test1, test2,

and test3. Preparing a codebook with various details of your data (e.g., variable name, variable type, variable labels, etc.) is a good practice to organize your data before entering into SYSTAT Worksheet. However, preparing a code book is not mandatory for data analysis. A code book, as shown below, becomes handy especially when you deal with data sets involving several variables with differing variable types, and variable lengths. A sample code book for the data in discussion is shown below for illustration purposes.

```
Var. name         var. type    var.length       var.labels

id                  numeric        2            identification number
sex$              character      1            student gender (f, m)
test1             numeric        2            test one score
test2             numeric        2            test two score
test3             numeric        2            test three score
```

In our example, the only character variable is sex (note the $ sign at the end of the name). All the other variables are numeric. Character variables cannot be used in some of the statistical procedures (e.g., ANOVA, REGRESSION, MANOVA).

Entering Data into SYSTAT Data Window

Let us open a Worksheet and enter the variable names and the data provided.

- Select NewData from the File menu. A blank data window opens.
- Double-click (VAR00001) to open the Variable Properties dialog box.
- Type ID for the variable name, select Numeric as the variable type, and set Decimal places to 0.
- Click OK.

Now, let us create the string variable, SEX$.

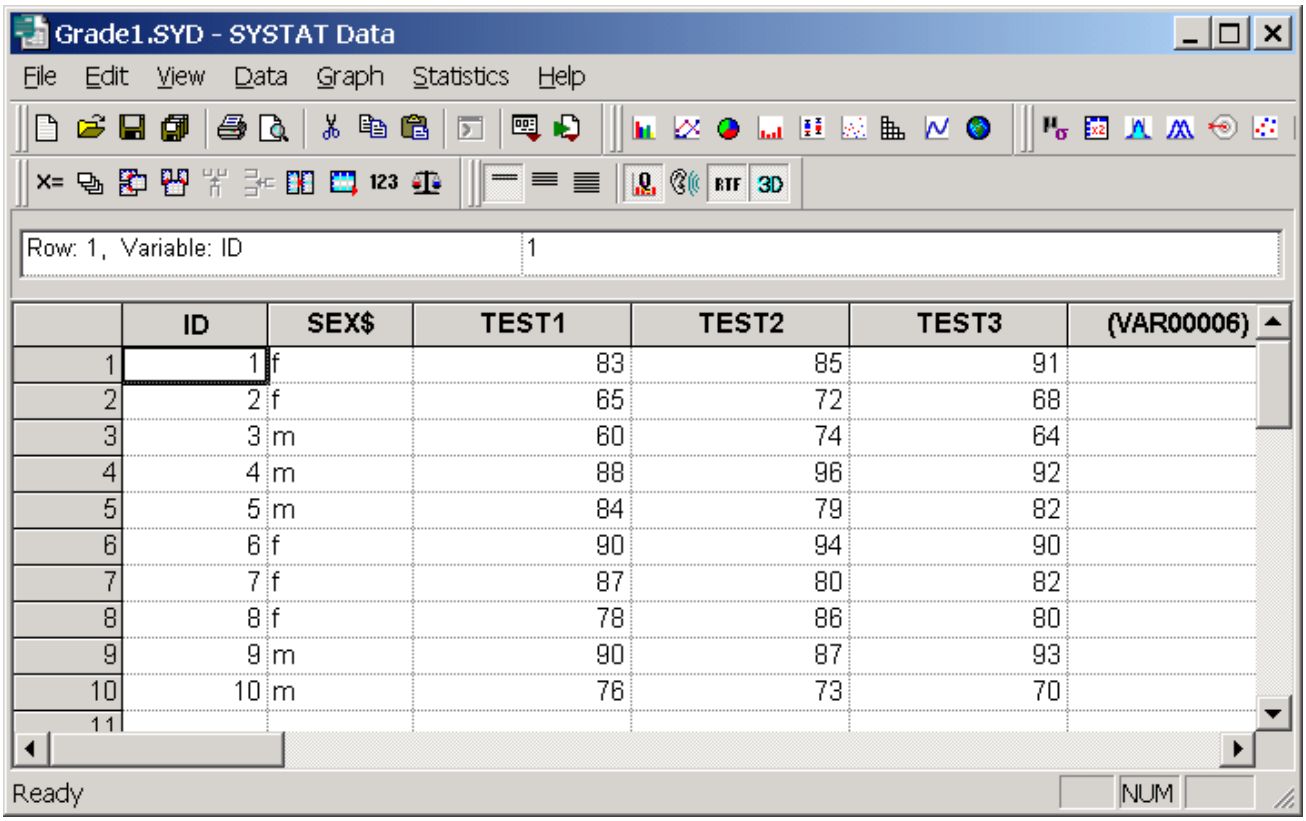
- Double-click (VAR00002) to open the Variable Properties dialog box.
- Type SEX$ for the variable name and select String as the variable type.
- Click OK.

Define the remaining three numeric variables, TEST1, TEST2, and TEST3, the same way the variable id was defined.

Now, the top row now should have five variable names. Let us key in the data.

- Move the pointer to the cell immediately below the variable ID, by clicking it.
- Type in the data given below, pressing [Enter] or [Tab] after each value. To move down columns, press the down arrow.

If the NumLock is on, you can use the keypad to type in numbers. Once you finish entering the data your Worksheet will appear as shown below.



	ID	SEX$	TEST1	TEST2	TEST3	(VAR00006)
1	1	f	83	85	91	
2	2	f	65	72	68	
3	3	m	60	74	64	
4	4	m	88	96	92	
5	5	m	84	79	82	
6	6	f	90	94	90	
7	7	f	87	80	82	
8	8	f	78	86	80	
9	9	m	90	87	93	
10	10	m	76	73	70	
11						

Creating a New Variable

Let us create a new variable, total, to represent the mean score from the three tests for each student he/she took during the semester.

- From the Data menu (Data window) select TransformLet

In the box on the left side of the screen, under the title variable, type in the new variable name total. Select Function Type: Multivariable. Now use the tab key to move the pointer to the box on the right side of the screen, under the title variable or expression, type in the expression: avg(test1 test2 test3)(note that this task can be performed by pointing the mouse to the function box and selecting the 'AVG( )' function and then highlighting and selecting the three variables.)

- Click OK

The variable TOTAL appears as the last column on the Worksheet.

Saving the Data File

Now that you have entered the data into the Worksheet, save it on to a diskette on drive A (those who are working from a personally owned PC may want to save it to a hard drive).

Now to save the data file:

- Select Save As... from the File menu.
- Type a:\grade1.SYD as filename (specify appropriate pathname if you are using alternate location to store the file)
- Click Save

Reading a Data File in Text Format

In some situations you may have a data file created using other software applications (e.g., Lotus, Excel), a

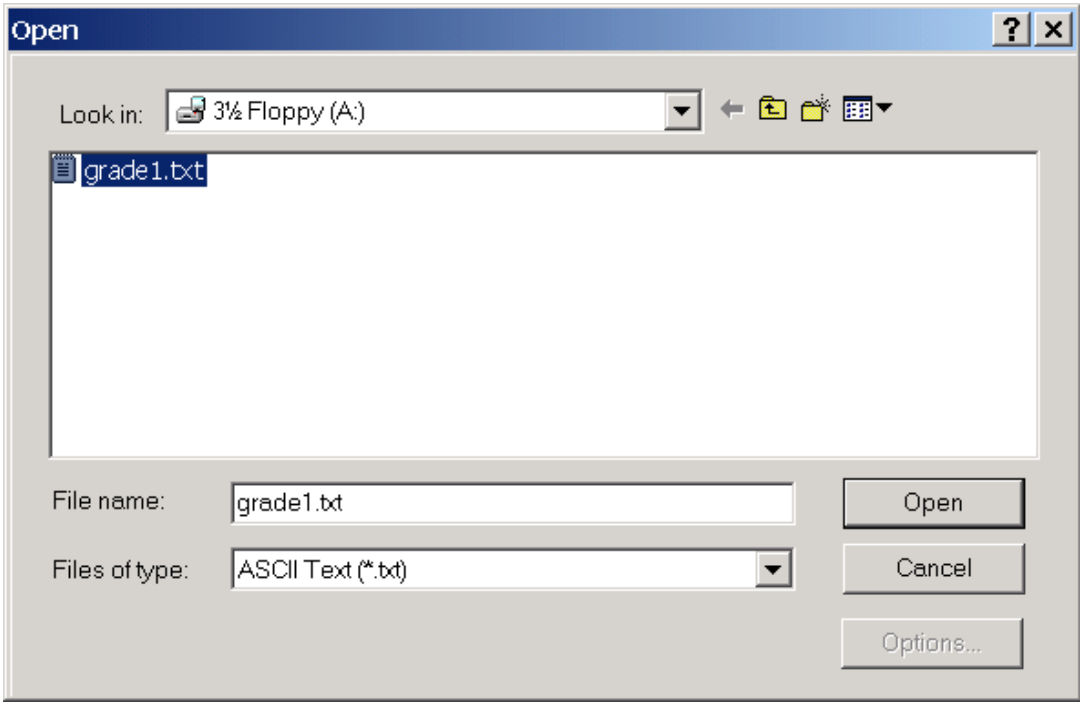
data editor or a word processor. In such a situation you do not have to enter your data again into the Worksheet. You can import such a file into SYSTAT. SYSTAT can read ASCII data files in free-format (each value is separated by spaces, commas, or tabs) or fixed-format (each value appears in the same place in every case).

Suppose the data file we discussed above is stored on drive A as an ASCII (text) file, grade.txt, as follows:

```
id sex$ test1 test2 test3
01  f    83    85    91
02  f    65    72    68
03  m    60    74    64
04  m    88    96    92
05  m    84    79    82
06  f    90    94    90
07  f    87    80    82
08  f    78    86    80
09  m    90    87    93
10  m    76    73    70
```

To import the data into SYSTAT:

- From the File menu select OpenData



- Select ASCII text from the list of file types located at the bottom corner of the dialog box.
- Type in a:\grade.txt as file name.
- The contents of the data file will be displayed in the Data window by clicking Open.
- Select Save As... from the File menu in the Data window. Type in a:\grade1.SYD as filename for the file to be saved in SYSTAT format.
- Click Save

To control for the number of decimal places displayed for the imported file, Double-click variable to open the Variable Properties dialog box and set Decimal places to 0.. If you do not give the columns names, Systat will automatically generate names like VAR(1), VAR(2), etc. You may rename, add/delete variables, cases to/from an existing SYSTAT file. Earlier we showed how to create a new variable, total, as the mean value for the three test scores. Follow the same procedure to create the new variable and save your new data set.

Similarly, if you have files created using spread-sheet software or SPSS system files you can import them to SYSTAT by choosing the appropriate file type from Open a File dialog box.

Descriptive Statistics

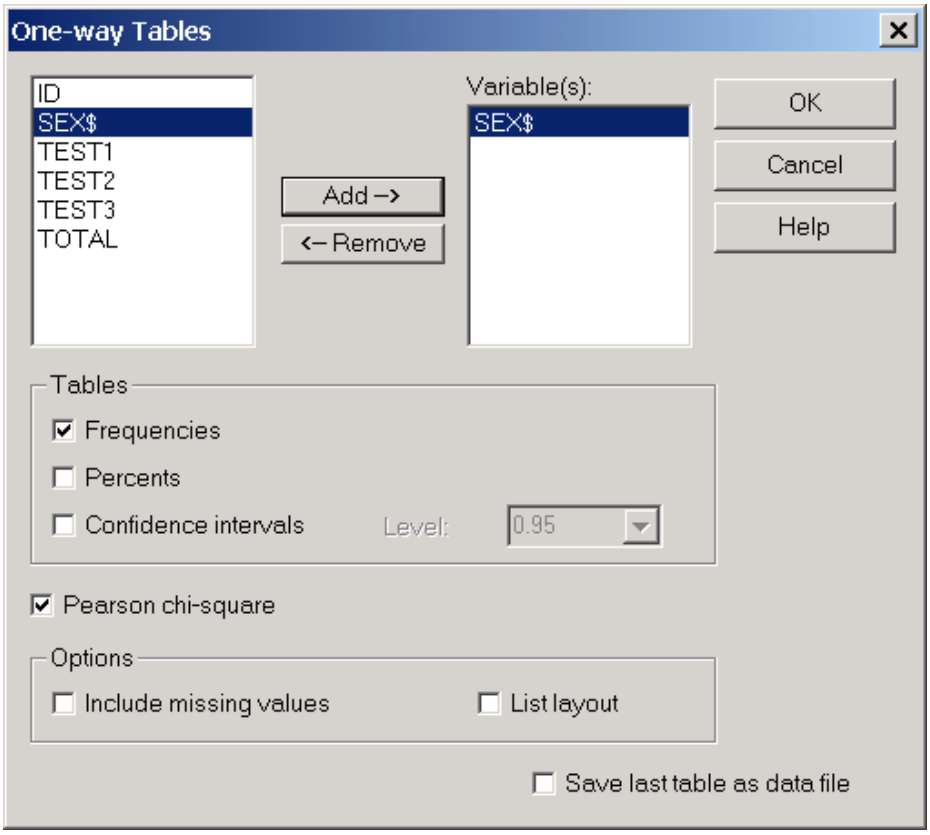
Now that we have created a data set, let us run a few descriptive statistics on the variables in the data set. (The worksheet will still be displayed in the Data window. If not, use the File menu to open the data set, grade1.SYD, you saved.)

Of the variables in our data set, SEX$ is a categorical variable, and TEST1, TEST2, TEST3 and TOTAL are continuous variables. We will use the One-way (Statistics/Tables/Crosstabs/One-way) command to obtain frequency counts for the variable SEX$, and Basic Statistics (Statistics/Descriptive Statistics/Basic Statistics) command to obtain descriptive statistics

To run the One-way procedure:

- Select Tables from the Statistics menu
- Select Crosstabs from the Tables menu
- Select One-way... from the Crosstabs menu

A dialog box titled One-way Frequency Tables appears.



- Highlight the variable SEX$
- Click the Add--> button.

The variable appears in the box on the right. In this situation we are using a single variable to create a one-way table for illustration. However, you may add a column variable to create a two-way table by selecting Two-way from the Crosstabs menu. SYSTAT can also create multiway tables.

- Choose Frequency from the list of options provided at the bottom of the dialog box by clicking on the box left to it. An sign appears on any item selected.
- To deselect an item click on the box with the sign. Deselect Pearson chi-square.
- Click OK

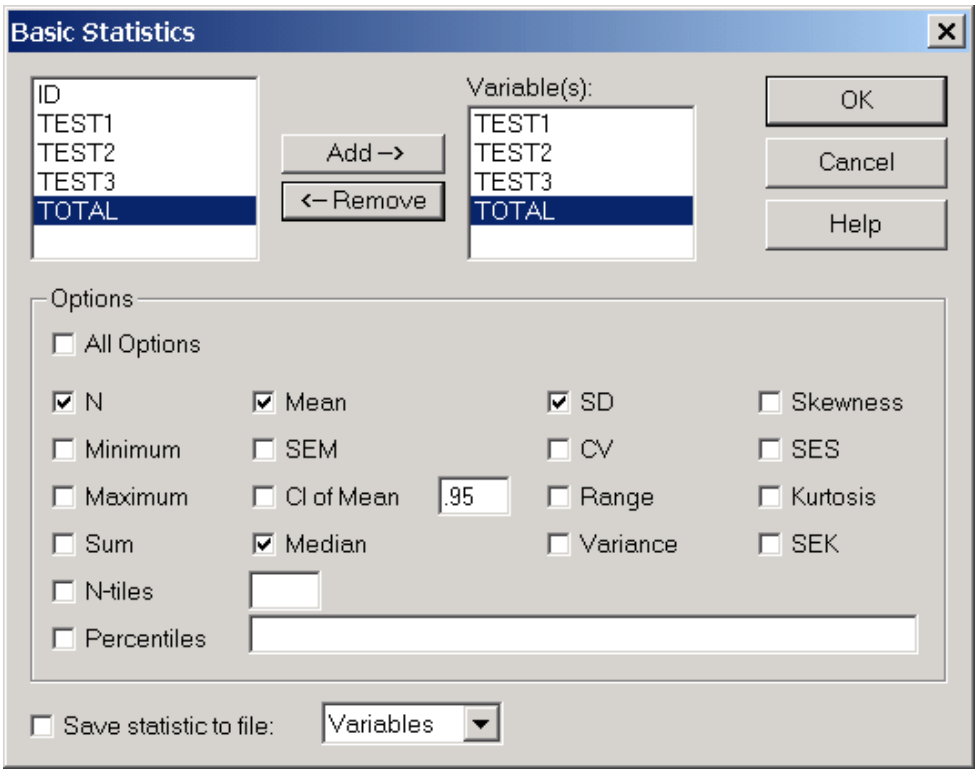
The output appears as shown below.

```
Frequencies
 Values for SEX$

          f     m    Total
      +-------------+
      |    5     5  |    10
      +-------------+
```

For a standard layout of one-way table select List layout from the One-way Frequency Tables dialog box.

To run the Basic Statistics procedure:

- From the Statistics menu select Descriptive Statistics
- From the Descriptive Statistics menu select Basic Statistics...



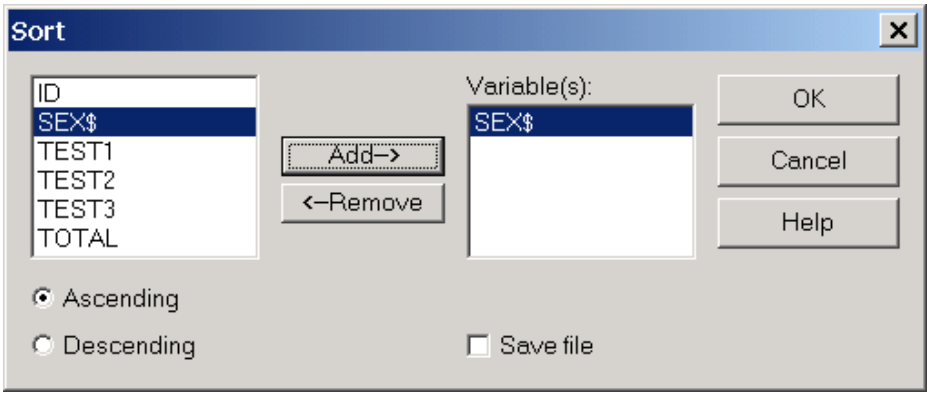
- Choose Mean, Median, SD, and N
- Select the variables TEST1, TEST2, TEST3, and TOTAL individually or collectively by holding down the control key.
- Click Add--> after selecting each variable or variables. The variables appear in the box on the right.
- Click OK

The output from the commands you just executed appears on the screen as shown below. If necessary, use the slide bar on the right side of the window to scroll down or up to view all of the output. For controlling the number of decimal points in the output you may select Options from the Edit menu (Main window) and choose the number of decimal places desired under the Data/Output Format, and Click OK. Now onwards the results will have the number of decimals you selected.

```
                    TEST1       TEST2       TEST3       TOTAL

N of cases             10          10          10          10
Median             83.500      82.500      82.000      82.333
Mean               80.100      82.600      81.200      81.300
Standard Dev       10.450       8.462      10.685       9.371
```

You may select a grouping variable to obtain separate analysis for individual groups. Suppose you want to obtain separate listing of the above analysis for males and females. The first task is to sort the file based on the grouping variable, e.g. sex, with all the females listed first followed by males (string variables are sorted in alphabetical order).

- From the Main window or Data window select Data menu
- From the Data menu select Sort



- From the variable list select SEX$ and click Add-->
- Click OK

The sorted file will be displayed on the screen. The sorted file can be saved by selecting Save file from the Sort dialog box and typing in a:\grade2.SYD as filename. (If you need to retrieve the original file, select Open from the File menu and type in a:\grade1.SYD).

The next step is to select By groups... from the Data menu for separate data analysis for males and females.

- Select By groups... from the Data menu from Main or Data window
- Select SEX$ from the variable list and click Add -->

(If you have missing values and exclude these values from analysis, you can check the box in front of Exclude missing. In our example there were no missing values.)

- Click OK

Now run the Statistics procedure.

- From the Statistics menu select Descriptive Statistics/Basic Statistics...
- Choose Mean, Median, SD, and N
- Highlight the variables TEST1, TEST2, TEST3, and TOTAL and click Add -->
- Click OK

The output from the commands you just executed will appear on the screen as shown below.

```
The following results are for:
              SEX$    = f


                        TEST1       TEST2       TEST3       TOTAL

  N of cases                 5           5           5           5
  Median                83.000      85.000      82.000      83.000
  Mean                  80.600      83.400      82.200      82.067
  Standard Dev           9.813       8.112       9.284       8.575


The following results are for:
              SEX$    = m


                        TEST1       TEST2       TEST3       TOTAL

  N of cases                 5           5           5           5
  Median                84.000      79.000      82.000      81.667
  Mean                  79.600      81.800      80.200      80.533
  Standard Dev          12.198       9.680      12.969      11.072
```

Printing Output

Once you are satisfied with your analysis you may want to obtain a hard copy of the output. You may print the entire output, or selectively delete the unwanted portion using the Edit menu and print the information you want, or highlight the part you want and then print it. You may save (File/Save as...) the output to a diskette for printing later. To print the output:

- From the File menu select Print...
- Click OK at the print menu.

The contents of the Window will now be printed.

Data Analysis

So far what we did was to look at SYSTAT to develop a basic idea on how SYSTAT for Windows works. The next step is to examine a few other data analysis procedures (e.g., correlation, regression, t-test) using SYSTAT for Windows. Only a limited number of procedures are discussed in this document. Refer to SYSTAT documents for further information.

Downloading Data

The data set we discussed in our earlier example was to get you started. Now we will examine another data set with more variables and cases, appropriate for the kind of analysis techniques we are examining.

In the example, you will import an ASCII data file, clas1.txt, created and saved in text format, into SYSTAT for Windows. The data collected from forty middle school students contains 28 variables. The first four variables (ID, SEX$, EXP, SCHOOL) are background variables. The Variable SEX$ has two levels (m=male, f=female). EXP (prior computer experience has three levels (1=less than one year, 2=1-2 years, 3=more than 2 years), SCHOOL (type of school system) has three levels (1=rural school, 2=suburban school, 3=urban school). The next 20 variables (C1...C10, M1...M10) are Likert type responses to a computer opinion survey, and mathematics anxiety survey. The last four variables (MATHSCOR, COMPSCOR, MANX, CANX) are scores on mathematics test, computer test, mathematics anxiety grouping, and computer opinion survey cumulative score. The variable MANX is a dichotomous variable created from low (coded as 0) and high (coded as 1) mathematics anxiety score.

To obtain a copy of this data file:

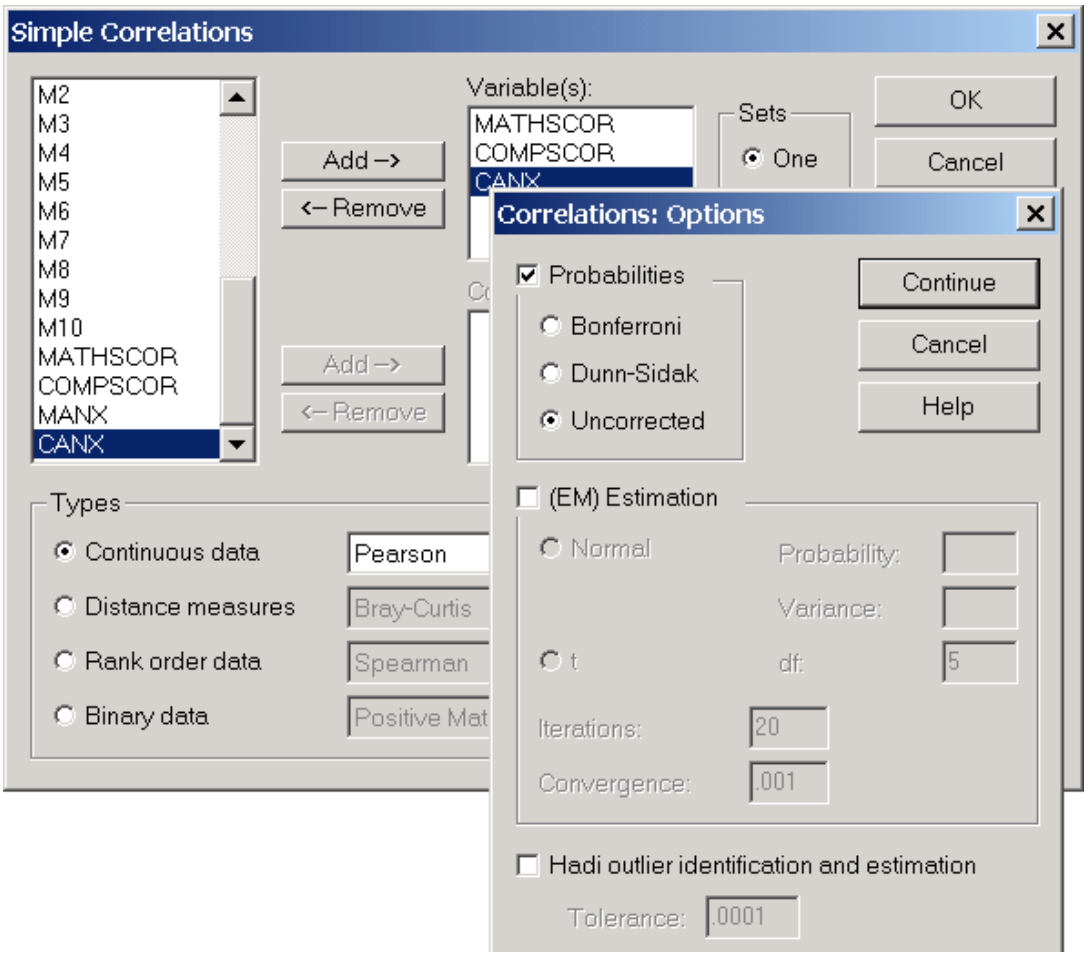
- Using a web browser (Netscape, Internate Explorer, lynx, etc.), download Sample SYSTAT Data.
- Save it to a file (for example, a:\clas1.txt).

Contact a STC consultant if you need assistance.

Import the file into SYSTAT using the method described earlier. The data will now be displayed on the Data window. Now you are ready for your data analysis.

Correlation Analysis

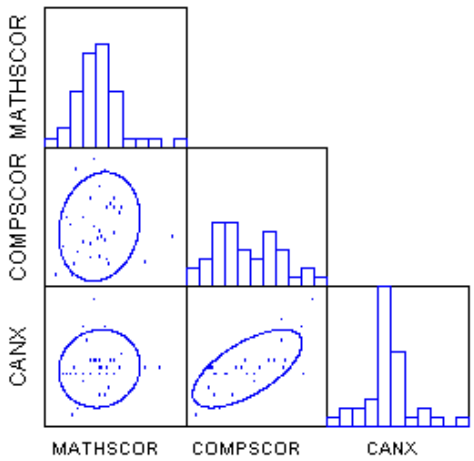
A correlation analysis is performed to quantify the strength of association between two numeric variables. In the following task we will perform a Pearson correlation analysis (SYSTAT can also perform a Spearman rank correlation). The variables used in the analysis are mathscor, compscor, and canx.



- From the Statistics menu select Correlations->Simple
- Highlight the variables, individually or collectively, MATHSCOR, COMPSCOR, and CANX, and click [Add-->]
- Selecting Options will open a Correlations: Options dialog box
- Check the Probabilities box and select Uncorrected
- Click Continue and finally OK

A symmetric matrix with the Pearson correlation as shown below will be displayed on the screen followed by another matrix with their probability values (p-values). The output (Quick Graph feature in Graph window) also includes a matrix of scatterplots (SPLOM) with one plot for each entry in the correlation matrix. Specify GRAPH=NONE in the command editor or select Options from the Edit menu and deselect Statistical Quickgraphs to suppress this feature.

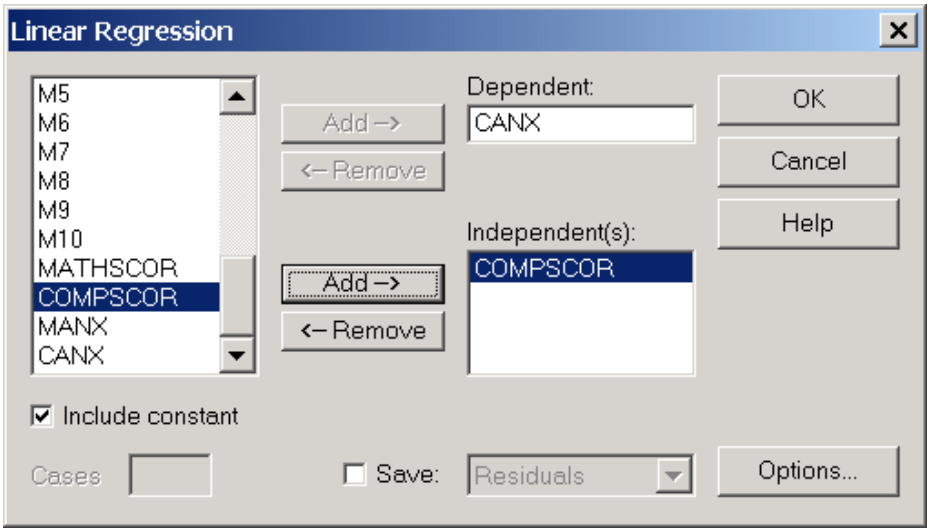
```
Pearson correlation matrix

                MATHSCOR     COMPSCOR         CANX
 MATHSCOR          1.000
 COMPSCOR          0.149        1.000
 CANX              0.068        0.657        1.000

Bartlett Chi-square statistic:    21.909 df=3 Prob= 0.000

Matrix of Probabilities

                MATHSCOR     COMPSCOR         CANX
 MATHSCOR          0.000
 COMPSCOR          0.359        0.000
 CANX              0.676        0.000        0.000

Number of observations: 40
```



Linear Regression

A correlation coefficient tells you that some sort of relation exists between the variables, but it does not tell you much more than that. For example, a correlation of 1.0 means that all points fall exactly on a straight line, but it says nothing about the form of the relation between the variables. When the observations are not perfectly correlated, many different lines may be drawn through the data. To select a line that describes the data, as close as possible to the points, you employ the regression analysis technique which is based on the least-squares principles. In the following task you will perform a simple regression analysis with 'canx' as the dependent variable, and 'compscor' as the independent variable.

- From the Statistics menu select Regression->Linear
- Highlight CANX as Dependent: variable and click on Add -->
- Highlight COMPSCOR as Independent(s): and click on Add -->



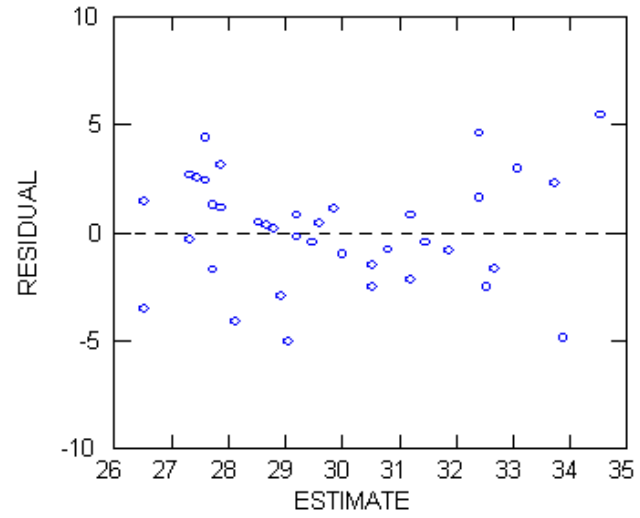
- Click OK

The output, as shown below, will be displayed on the screen with regression statistics including slope, intercept, and squared multiple R. Quick Graph feature appears in Graph window and it includes a plot of regression residuals against the predicted values. Use the same procedure as earlier to suppress this feature.

```
Dep Var: CANX   N: 40   Multiple R: 0.657   Squared multiple R: 0.432

Adjusted squared multiple R: 0.417   Standard error of estimate: 2.544

Effect       Coefficient    Std Error     Std Coef Tolerance     t   P(2 Tail)
```

```
CONSTANT          23.194        1.315        0.0          .       17.634    0.000
COMPSCOR           0.133        0.025        0.657      1.000      5.375    0.000

                              Analysis of Variance

Source              Sum-of-Squares    DF  Mean-Square      F-Ratio       P

Regression                186.922     1       186.922       28.891       0.000
Residual                  245.853    38         6.470
-------------------------------------------------------------------------------

Durbin-Watson D Statistic     1.103
First Order Autocorrelation   0.406
```

Plot of Residuals against Predicted Values

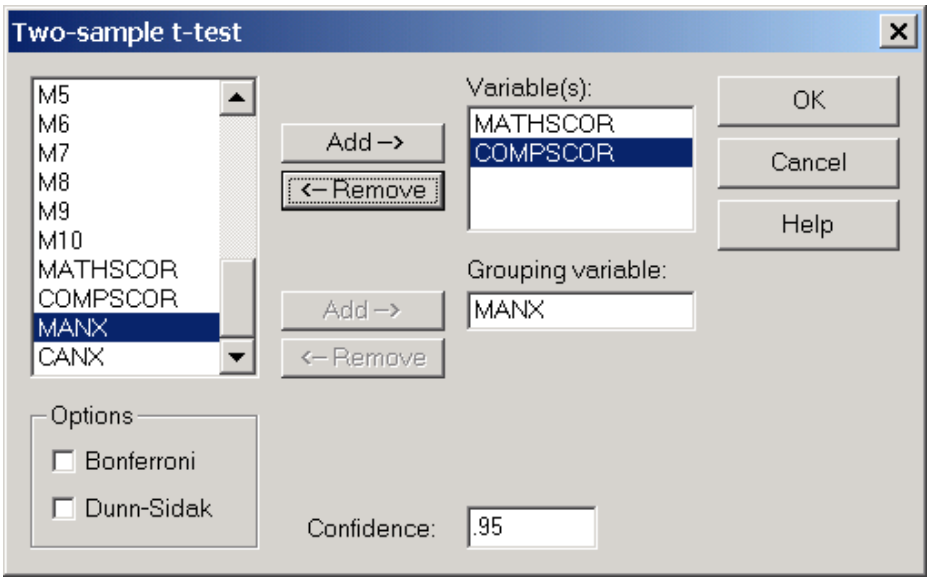


T-test

T-test is a data analysis procedure to test the hypothesis that two population means are equal. SYSTAT can compute both independent (unrelated groups) and dependent (related groups) t-tests. For independent t-tests, your grouping variable should have exactly two values (e.g., male/female, pass/fail). The grouping variable may either be numeric or character. If a grouping variable has more than two categories then you can use the Data/Select cases... menu to select the two values you want to perform t-test with. Once you select cases make sure you deselect it to restore the data set if you plan to use all the cases for subsequent data analysis.

In the following task we will perform an independent t-test. The dependent variables are mathscore, and compscor, and the independent (grouping) variable is manx.(If you do not select a grouping variable by default a paired t-test will be performed.)

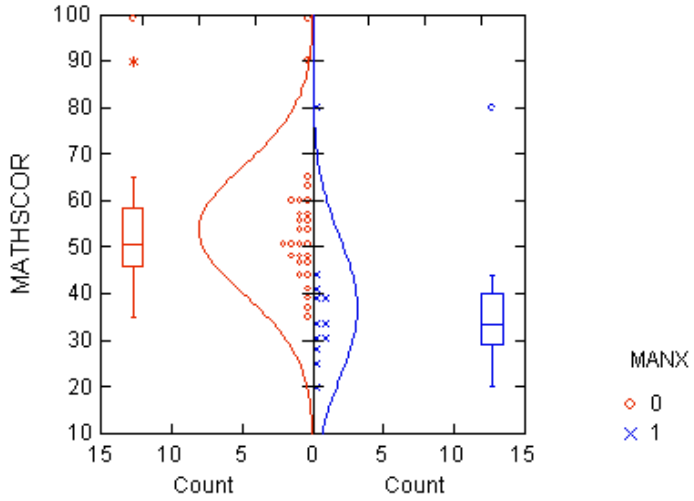
- From the Statistics menu select t-test->Two-Groups



- Highlight the dependent variables MATHSCOR, and COMPSCOR, and click Add-->
- Highlight MANX for grouping variable and click Add-->
- Click OK

The output from the run will be displayed on the screen as shown below. Quick Graph feature in Graph window includes a combined display of three graphs (a boxplot, a normal curve and a dit-plot) for each group. Use the same procedure as earlier to suppress this feature.

```
Two-sample t test on MATHSCOR grouped by MANX

  Group                N         Mean           SD
            0         28       53.750       13.845
            1         12       37.000       15.214


     Separate Variance t =        3.277 df =   19.2    Prob =        0.004
     Difference in Means =       16.750   95.00% CI =      6.058 to     27.442

       Pooled Variance t =        3.406 df =   38      Prob =        0.002
     Difference in Means =       16.750   95.00% CI =      6.793 to     26.707
```

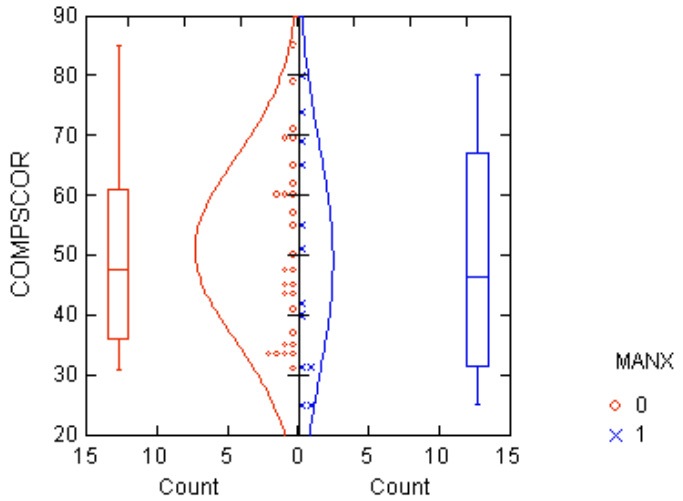


Two-sample t test on COMPSCOR grouped by MANX

Group	N	Mean	SD
0	28	51.000	15.253
1	12	49.083	19.486

Separate Variance t = 0.303 df = 17.1 Prob = 0.765
Difference in Means = 1.917 95.00% CI = -11.416 to 15.249

Pooled Variance t = 0.335 df = 38 Prob = 0.740
Difference in Means = 1.917 95.00% CI = -9.671 to 13.505



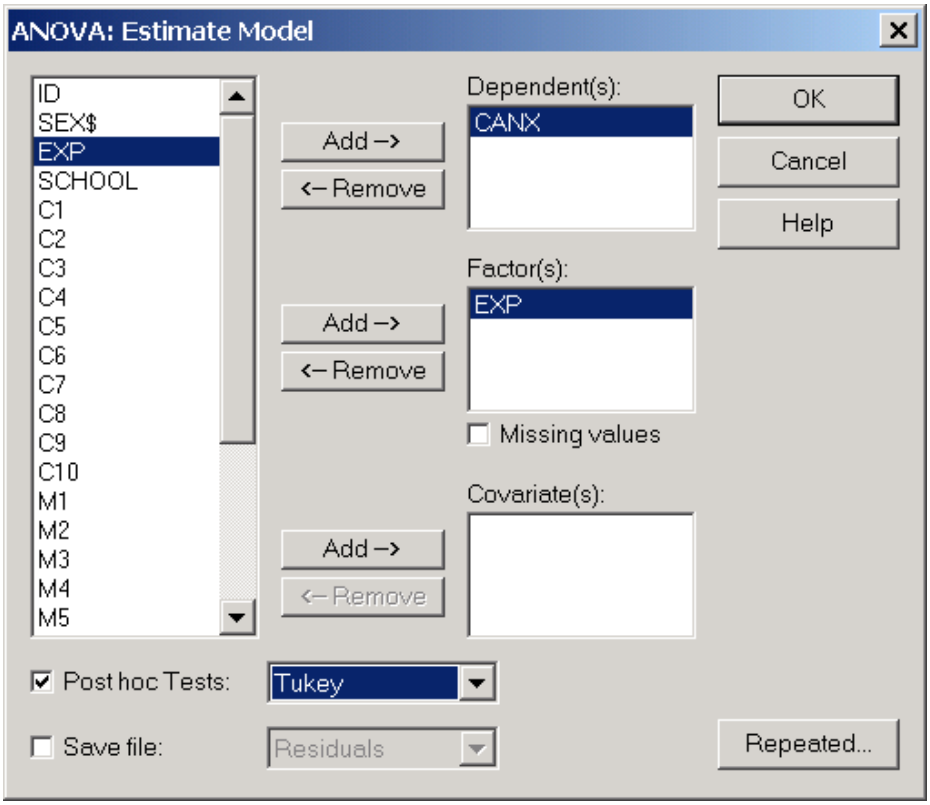
Analysis of Variance

The statistical technique used to test the null hypothesis that several means are equal is called analysis of variance. It is called that because it examines the variability in the sample and, based on the variability, it determines whether there is reason to believe the population means are not equal. In analysis of variance, the observed variability in the sample is divided, or partitioned, into two parts: the variability of observations within a group (around the group mean), and variability between the group means. If the two estimates are substantially different, you can reject the null hypothesis that the population means are equal. The statistical test for null hypothesis that all of the groups have the same mean in the population is based on computing the ratio of the two estimates, called an F statistic. The observed significance level is obtained by comparing the calculated F value to the F distribution (the distribution of the F statistic when the null hypothesis is true).

A significant F value only tells you that the means are probably not all equal. It does not tell you which pairs of groups appear to have different means. To pinpoint exactly where the differences are, multiple comparisons may be performed.

In the following exercise you will perform an ANOVA with canx as the dependent variable and 'exp' as the factor variable. To perform a pairwise mean comparisons to identify which means differ from others a Tukey HSD test has been employed.

- From the Statistics menu select Analysis of Variance (ANOVA)/Estimate Model(Selecting General Linear Model/Estimate Model from Statistics menu will result in the same procedure.)

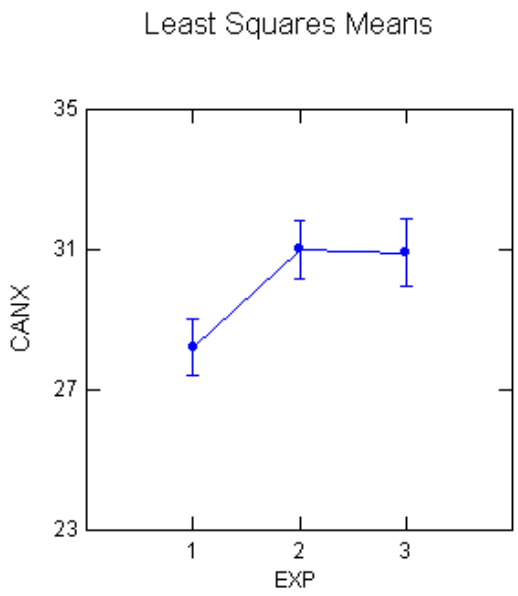


- Highlight CANX as Dependent(s): and click [Add-->]
- Highlight EXP as Factor(s): and click [Add-->]
- Select Post hoc Tests and choose Tukey from drop-down list. (Selecting General Linear Model/Pairwise Comparisons from Statistics menu will result in the same procedure plus Dunnett's test. But this option becomes active only after you run your ANOVA.)
- Click OK

The output as shown below will be displayed on the Main window. In the Graph window, Quick Graph includes a plot of residuals from each estimated cell mean versus the estimated cell mean.

```
Effects coding used for categorical variables in model.

Categorical values encountered during processing are:
EXP (3 levels)
          1,        2,        3

Dep Var: CANX   N: 40   Multiple R: 0.406   Squared multiple R: 0.165


                            Analysis of Variance

Source             Sum-of-Squares   df  Mean-Square     F-ratio       P

EXP                        71.466     2       35.733       3.659       0.035

Error                     361.309    37        9.765
```

```
-------------------------------------------------------------------------------
```




```
-------------------------------------------------------------------------------

*** WARNING ***
Case           39 is an outlier        (Studentized Residual =        3.385)

Durbin-Watson D Statistic     1.769
First Order Autocorrelation   0.043
COL/
ROW EXP
  1  1
  2  2
  3  3
Using least squares means.
Post Hoc test of CANX
-------------------------------------------------------------------------------

Using model MSE of 9.765 with 37 df.
Matrix of pairwise mean differences:

                         1           2           3
             1          0.000
             2          2.800       0.000
             3          2.709      -0.091       0.000

Tukey HSD Multiple Comparisons.
Matrix of pairwise comparison probabilities:

                         1           2           3
             1          1.000
             2          0.054       1.000
             3          0.087       0.997       1.000
-------------------------------------------------------------------------------
```

The output shows that there is a significant difference among groups with different levels of computer experience at least at .05 probability level.

The output for pairwise comparisons include a table of mean differences and another table of probabilities. To determine significant differences, examine the pairs and their probability level. From the output it is evident that there is a marginally significant difference between group 1 (exp=1) and group 2 (exp=2). None of the other combinations produced a significant difference at least at the 0.05 level.

Using SYSTAT's Graph Menu

SYSTAT provides a wide selection of graphics for every stage of your project: exploration, research, and presentation. The graphics capabilities of SYSTAT include:

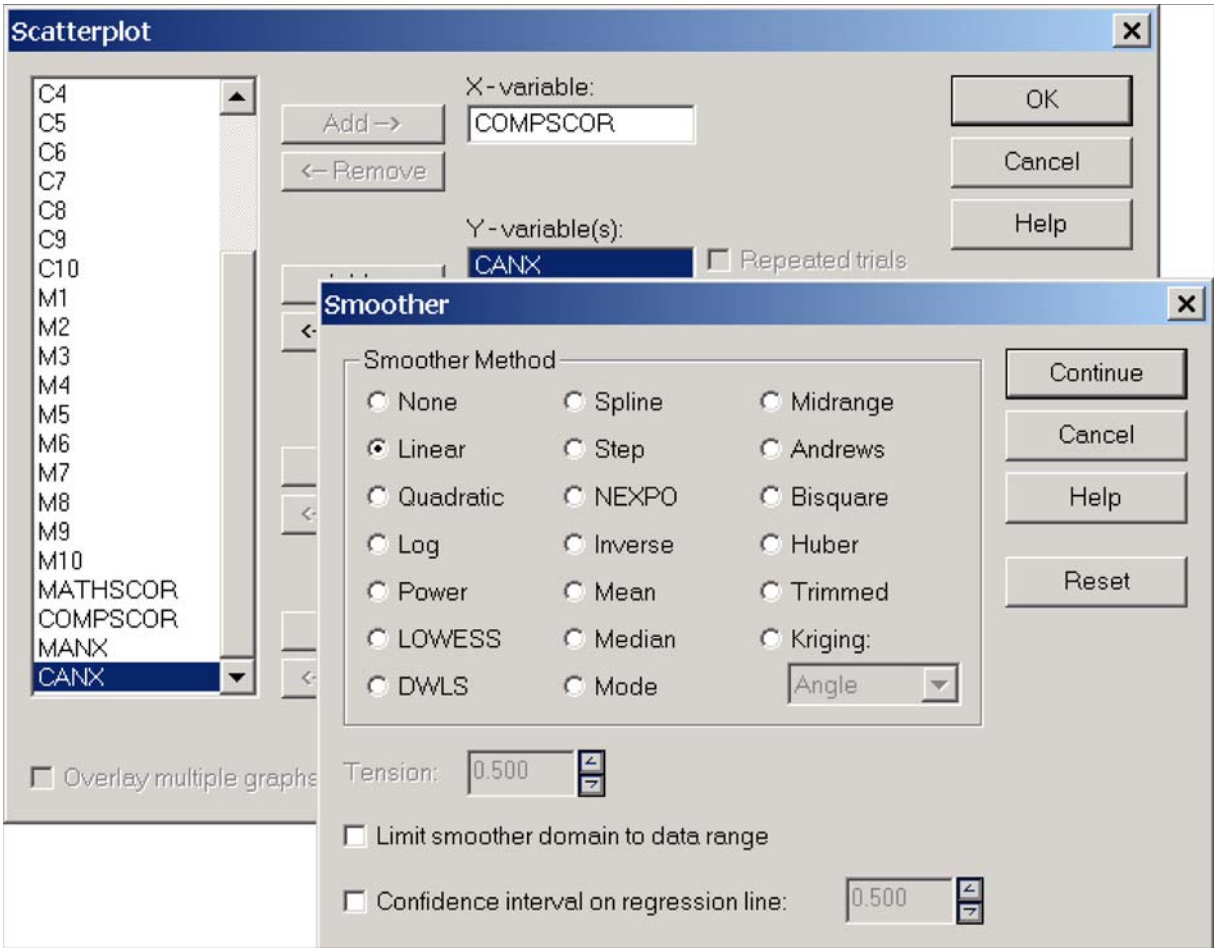
- histograms with curve fitting
- bar graphs, box plots, stem-and-leaf diagrams, pie charts
- 3-D rotation, maps with geographic projections
- mathematical function plots, log and power scales
- confidence intervals, ellipses, and centroids
- contour plots, control charts
- case coding of labels and symbols
- linear, quadratic, step, spline, polynomial, LOWESS, exponential, and DWLS smoothing in two and three dimensions
- rectangular, spherical, polar, cylindrical, and triangular coordinates, perspective depth and projections.

Plotting Two Variables with SYSTAT

Looking at a plot is one of the best ways to examine relationships and patterns. For example, a scatterplot allows the visual representation of two separate distributions on a single diagram.

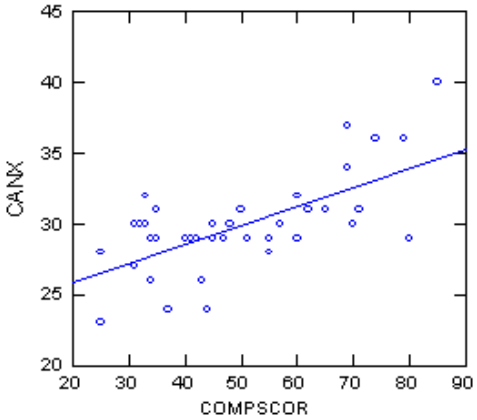
In the following task you will plot the variables CANX (dependent variable) by COMPSCOR (independent variable). We will also fit the data points on the scatterplot based on the least-squares principle.

- From Graph menu (Main window) select Plots and then choose Scatterplot



- Highlight COMPSCOR as X-variable: and click Add -->
- Highlight CANX as Y-variable(s): and click Add -->
- Choose Linear as smoother method from Options/Smoother and click Continue
- Click OK

The plot, as shown below, will be displayed on the screen



To print the graphics output, select File/Print..., and respond to the queries appear in subsequent dialog boxes. You may save your graphics streams to a file using File/Save as... To remove the Graphics window select File/Close Window.

Detailed discussion of all the graphics capabilities of SYSTAT is not possible through this document. You may refer to SYSTAT's Graphics document for learning more about the graphics capability of SYSTAT.

Further Help

The material covered in this document illustrates some of the basic capabilities of SYSTAT for Windows. Examining additional features of SYSTAT for Windows is beyond the scope of this class. However, with a little effort it is possible to gain mastery on SYSTAT.

For further information refer to SYSTAT documents at the Software Manual Locations mentioned at the beginning of this document and technical support in using SYSYAT contact the Stat/Math Center (812/855-4724 or 317/278-4740; email statmath@iu.edu M-F 9.00 a.m - 5.00 p.m.).

The documents currently available for SYSTAT for Windows are:

1. SYSTAT 10.2 Getting Started
2. SYSTAT 10.2 Data
3. SYSTAT 10.2 Statistics I
4. SYSTAT 10.2 Statistics II
5. SYSTAT 10.2 Graphics
6. SYSTAT 10.2 Language Reference

These documents may be purchased through the IU Bookstore or directly ordered from SYSTAT Software, Inc. (http://www.systat.com).
