# Table of Contents

**INTRODUCTION** ........................................................................................................... 5  
**GETTING HELP** ........................................................................................................... 5  
**INSTALLATION** ........................................................................................................... 5  
**UNINSTALLING CLARITY** .............................................................................................. 5  
**TERMINOLOGY USED IN THE MANUAL** ........................................................................ 5  
**CLARITY STARTUP SCREEN** ......................................................................................... 6  
**DATA SPREADSHEET** ................................................................................................... 7  
  - Rows ................................................................................................................................. 7  
  - Columns ............................................................................................................................ 7  
  - Cells .................................................................................................................................. 7  
  - Entering data .................................................................................................................... 7  
  - Editing data ...................................................................................................................... 7  
  - Spreadsheet and Data File Limits .................................................................................. 8  
  - Drag and Drop .................................................................................................................. 8  
**FILE MENU** .................................................................................................................... 8  
  - New .................................................................................................................................... 8  
  - Open AcaStat Data File .................................................................................................... 8  
  - Save Data File As .............................................................................................................. 8  
  - Import Tab Delimited Data ............................................................................................... 8  
  - Import Comma Delimited Data ........................................................................................ 8  
  - Tips on Importing Data .................................................................................................... 9  
  - Tips on using a spreadsheet application to create a csv or txt file. ............................. 9  
    - Pasting Data .................................................................................................................... 9  
    - Importing Text Data Files ............................................................................................. 9  
    - Example of Import Error .............................................................................................. 10  
  - Export Data (Tab Delimited) .......................................................................................... 10  
  - Export Data (Comma Delimited) .................................................................................... 10  
**EDIT MENU** .................................................................................................................. 11  
  - Delete Row ...................................................................................................................... 11  
  - Delete Column .................................................................................................................. 11  
  - Cut .................................................................................................................................... 11  
  - Copy .................................................................................................................................. 11  
  - Paste .................................................................................................................................. 11  
  - Clear Analysis List Boxes ............................................................................................... 11  
  - Clear All ............................................................................................................................ 11  
  - Select All ........................................................................................................................... 11  
  - Convert Output for Spreadsheet Pasting ........................................................................ 11  
**DATA MENU** .................................................................................................................. 12  
  - Copy Data from Clarity .................................................................................................... 12  
  - Paste Data into Clarity ...................................................................................................... 12  
  - Create Data Dictionary .................................................................................................... 12  
  - Filter Data .......................................................................................................................... 12  
**RECODE MENU** ............................................................................................................. 13  
  - Format Variable ............................................................................................................... 13

Clarity User Manual
 Variable Name .......................................................... 13
 Variable Label .......................................................... 13
 Change Decimals ...................................................... 14
 Value Labels ............................................................ 14
 Missing Values .......................................................... 14
 Recode Values .......................................................... 15
 Recode Strings .......................................................... 16
 Compute Variables .................................................... 17
 Combine Variables .................................................... 17
 Find and Replace ...................................................... 18
 Trim Strings ............................................................. 19
 DATA EXAMPLES MENU ............................................. 20
 CALCULATE MENU .................................................. 20
 Run Procedure .......................................................... 20
 DOCUMENTATION MENU .......................................... 21
 Student Workbook .................................................... 21
 Free Applied Statistics iBook ....................................... 21

 STATISTICAL PROCEDURES ........................................ 22

 ANALYSIS LIST BOXES .............................................. 22
 Variable List ............................................................ 22
 Dependent Variable ................................................. 22
 Independent Variable ............................................... 22
 Control Variable ...................................................... 23
 Variable Selection Indicators .................................... 23
 Tips for using the Control Box ................................... 23

 STATISTICAL PROCEDURES ........................................ 23
 Frequencies ............................................................. 23
 List Variables .......................................................... 23
 Descriptives ............................................................. 24
 Explore Means ........................................................ 24
 Crosstabsulation ..................................................... 24
 T-tests of means ..................................................... 24
 Paired t-test ............................................................ 25
 Wilcoxon Signed-Rank .............................................. 25
 Mann-Whitney U ..................................................... 25
 Analysis of Variance (ANOVA) .................................. 25
 Kruskal-Wallis ......................................................... 25
 Pearson Correlation ............................................... 26
 Spearman rho ........................................................ 26
 Point-Biserial Correlation ....................................... 26
 OLS Regression ....................................................... 26
 Logistic Regression ............................................... 26
 Diagnostic Accuracy .............................................. 27
 Appraisal Ratios Procedure .................................... 27
 Simulation .............................................................. 27
 Random Sample ..................................................... 27
 Repeated Sampling ................................................ 28

 OUTPUT LOG ............................................................ 29
FILE MENU.............................................................................................................................. 30
  New ........................................................................................................................................ 30
  Open Output File .................................................................................................................. 30
  Save Output As ..................................................................................................................... 30
  Print Output .......................................................................................................................... 30
EDIT MENU .................................................................................................................................. 30
  Cut.......................................................................................................................................... 30
  Copy ....................................................................................................................................... 30
  Paste ...................................................................................................................................... 30
  Select All ............................................................................................................................... 31
  Convert Output for Spreadsheet Pasting .............................................................................. 31
OTHER MENU ITEMS AND CLARITY FEATURES ..................................................................... 32
HELP MENU ................................................................................................................................ 32
  Clarity Help ............................................................................................................................ 32
  Statistics Glossary ................................................................................................................ 32
  Support Webpage .................................................................................................................. 32
  Check for Updates .................................................................................................................. 32
Introduction

Clarity provides students, teachers, researchers, and managers with an inexpensive and easy to use data analysis tool and instructional aid. Clarity includes a spreadsheet for entering and viewing data and an output log. A statistics glossary is also provided along with example data.

Getting Help

Application help can be found in this manual and in the Help module provided with Clarity.

In addition to application help, statistical help is provided in four formats:

✔ Glossary of statistical terms and definitions

✔ Application How To Guide

✔ Student Workbook in pdf format

✔ Applied Statistics Handbook

Installation

Download an evaluation version or purchase Clarity from http://www.acastat.com or the Mac App Store. The software is downloaded in a self-extracting file. Save the software and double-click on the file to begin the installation process.

Uninstalling Clarity

For Mac versions, drag the program file to the Trash Bin. For Windows versions, use the Windows Control Panel option ADD/REMOVE programs to uninstall Clarity. Files that were created or changed after installation of Clarity must be removed separately by the manual file deletion procedures in Mac and Windows.

Terminology Used in the Manual

When discussing statistical procedures, the manual will refer to two general types of data -- continuous and categorical. The following terms will be used for convenience and simplicity, but there are variations in these themes that will be apparent to those with statistics training. Please see the Statistics Glossary in Clarity for more information on these and related terms.

Continuous Data – This refers to data where it is reasonable to calculate an average. This type of data is also referred to in statistics as interval/ratio level of measurement or scale data.
Categorical Data – This refers to data where averages would not be meaningful. This type of data is also referred to in statistics as nominal/ordinal level of measurement or discrete data.

Clarity Startup Screen

Clarity starts by displaying the Analyze tab. A control panel along the left side of the application contains statistical procedure controls and a list of variables when a data file has been loaded. A data box is also displayed. Click one of the options or drag and drop compatible data files into the data box. The Output tab automatically displays the results for the current analysis and stores all analyses conducted after opening Clarity.

Startup View
Data Spreadsheet

The Analyze tab uses a spreadsheet to create, import, and export data files. Opening or importing a new file will replace the current contents.

Rows

Each row represents one observation (this is also known as the unit of analysis). If the data contain information on individuals, each row should contain the data for one person (observation). If the data contain information on cities, each row should contain the data relevant to one particular city.

Columns

Each column represents data for all observations relating to one variable or characteristic. As an example, if a column contains information on personal income in U.S. dollars, all data in that column should represent the income for each person in the data file. The spreadsheet will display up to 255 columns (variables) of data. Data files with more than 255 variables can be imported for analysis but only the first 255 variables will be visible in the spreadsheet.

Cells

Each cell represents a value for one observation for one variable. It can contain string data (words) or numerical data (numbers). Most research data files use numerical data to represent words. As an example, for the variable sex, it is common to code Male as "0" and Female as "1." This makes data entry much easier since one number is entered instead of typing a word. It also simplifies recoding variables, identifying controls, and running analyses.

Entering data

To enter data, point and click on a cell and begin typing. Pressing an arrow key will save the entry and move the data entry point to another cell. As an example, pressing the right arrow key will save the entry and move the data entry point to the right. Pressing the Escape (Esc) key will cancel data entry. Pressing the Enter key will move the data entry point down one row.

Editing data

To edit data, click on the cell to edit. Pressing any of the arrow keys will save the entry and move the data entry point to an adjacent cell. Do not enter characters such as a $, %, or comma (unless used as the locale decimal indicator) and avoid the use of strings (words) whenever possible.
Spreadsheet and Data File Limits

The spreadsheet will display up to 255 variables (columns of data). Data files with more than 255 variables can be imported for analysis but only the first 255 variables will be visible in the spreadsheet. The number of rows is controlled by system memory and speed of the computer. Files larger than 100,000 observations will slow loading and analysis.

Drag and Drop

The data spreadsheet accepts drag and drop for Clarity data files and delimited text data files (tab and csv). The Output tab accepts drag and drop for Clarity output text files.

File Menu

New

Clears all data from the spreadsheet and resets the variable lists.

Open AcaStat Data File

Use to open an existing AcaStat data file. AcaStat system data files have the extension ".dcs". AcaStat files contain data formatting and missing values.

Save Data File As

Saves the contents and any formatting of the data into an AcaStat data file.

Import Tab Delimited Data

Opens a text file that contains tabs between each variable so Clarity is able to import the data into a spreadsheet format. These files normally have a filename extension of ".txt" or ".dat". Clarity assumes the first row in an imported file contains the variable names.

Import Comma Delimited Data

Opens a text file that contains commas between each variable so Clarity is able to import the data into a spreadsheet format. These files normally have a filename extension of ".csv" (comma separated values). Clarity assumes the first row in an imported file contains the variable names. Countries that use comma decimal delimiters should use tab delimited instead of csv to import or export data.
Tips on Importing Data

Import data from text data files or by pasting data copied from a word processor, spreadsheet, or database table. Please see the section below on using a spreadsheet application to create compatible data files.

Tips on using a spreadsheet application to create a csv or txt file

The first row in the spreadsheet should contain a unique variable name that represents the data in that column (try to keep the names 8 characters or less). The variable names identify columns when imported into the Clarity spreadsheet and are used to select columns of data for analysis. To save a delimited text file, use File/Save As and select the file type as either tab delimited or comma separated values (CSV) format.

Pasting Data

To paste data from a table, select and copy data in a word processor, spreadsheet, or database table. Select "Data/Paste Data into Clarity" or use the paste data button on the toolbar. The paste procedure replaces all data in the current spreadsheet. Copied data must be consistent with the tips discussed above.

Importing Text Data Files

Clarity assumes the first row in an imported file contains the variable names. If the data file is large, it may take a few moments to read the data. Once read, the data will be displayed in the spreadsheet and variable names (column headings) will be visible in the variable list box. After importing data, save the file as an AcaStat system file. This will decrease the time it takes to load a file and allows the file to retain variable formatting.
Example of Import Error

When creating a comma separated value file, make sure all commas have been removed from the cell contents before creating the file. If they are not removed, as might be the case with currency data, Clarity’s import procedure will create an extra variable for each comma. This will cause the cell contents in the Clarity spreadsheet to be out of alignment with the column. In the example below, the data in rows two and three do not align properly with the appropriate columns.

<table>
<thead>
<tr>
<th>AGE</th>
<th>SEX</th>
<th>INCOME</th>
<th>V4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>65</td>
<td>M</td>
<td>33000</td>
</tr>
<tr>
<td>2</td>
<td>23</td>
<td>M</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>46000</td>
<td>45</td>
<td>F</td>
</tr>
</tbody>
</table>

Other import errors may occur if the text file does not exactly match the format specified above. The following list is not complete, but it contains common errors in importing data.

- No variable names in the first row
- Blank rows at the top of the data file
- Special embedded characters in the cells that are neither numbers nor words. This is especially common in data copied from an html table from the Internet.
- Embedded tabs in spreadsheet cells

Export Data (Tab Delimited)

Use Export Data to save the contents of the spreadsheet as a text file that contains tabs between each variable (column). These files are given the filename extension ".txt". This file can be viewed and edited in word processing or spreadsheet software and can be imported into other statistical software.

Export Data (Comma Delimited)

Use Export Data to save the contents of the spreadsheet as a text file that contains commas between each variable (column). These files are given the filename extension ".csv". This file can be viewed and edited in word processing or spreadsheet software and can be imported into other statistical software.
Edit Menu

Delete Row

Deletes the row for the currently selected cell. Delete one row or multiple selected rows. The change will be complete when the data file is saved.

Delete Column

Deletes the column for the currently selected cell. Delete one column at a time. The change will be complete when the data file is saved.

Cut

If the Output Log tab is selected, this cuts the selected text from the output screen.

Copy

If the Output tab is selected, this copies the selected text in the output screen.

Paste

If the Output Log tab is selected, this pastes the text into the output screen.

Clear Analysis List Boxes

If the Analyze tab is selected, this clears all variables displayed in the analysis list boxes and control. It does not clear the variable list box or spreadsheet. Use File/New to clear all list boxes and the spreadsheet.

Clear All

If the Output Log tab is selected, this clears all output log entries.

Select All

If the Output Log tab is selected, this selects all the text into the current output screen.

Convert Output for Spreadsheet Pasting

If the Output tab is selected, this formats the output with tabs to improve the alignment of output copied from Clarity for pasting into spreadsheet software. The tabs are used by
spreadsheet software to align the table data properly into columns for further editing and formatting.

Data Menu

Copy Data from Clarity

Copies all the data in the spreadsheet for pasting into other software.

Paste Data into Clarity

Paste data copied from a word processor, spreadsheet, or database table. The data must have variable names in the first row of the copied data. To paste data from a table, select and copy data in a word processor, spreadsheet, or database table and then paste the data into the Clarity spreadsheet. This procedure replaces all data in the current spreadsheet.

Create Data Dictionary

This option creates an item in the Output tab that displays data file variables, variable and value labels, and missing values formatting.

Filter Data

Select Filter Data to exclude observations from the analysis. This module can be used to limit analyses to a finely defined set of observations. Use up to two variables to exclude observations from subsequent analyses. The rows excluded from the analyses are marked by a bolded X. Resetting the filter removes the filter from future analyses. No data are permanently lost when using a filter.

A filter variable will be added to the spreadsheet. If the data file is saved after filtering, the filter variable will be saved and formatted with the filter specifications for later use as a control.
Recode Menu

Format Variable

Use the Format Variable module to view and edit variable formatting. In order to produce output that is easier to interpret, create variable names and assign short explanatory labels to variable names and values. Formatting information will be saved with the AcaStat data file and can be changed at any time. If data are exported, the formatting information will not be exported. Click the Copy Format button to copy value and missing settings for pasting into other variables with the same format, such as survey questions with identical response options. Click “Set” to record changes to formatting. The data file must be saved as an AcaStat data file to retain formatting changes.

Variable Name

Since variable names are used as column labels and are listed in the variable list boxes, it is important to create short but unique variable names (8 characters or less is recommended). As an example, level of education could be named DEGREE or EDUC or EDYRS.

Variable Label

Use variable labels to create a more detailed description of the variable. The variable label is used to create more meaningful output. Using the above example, DEGREE could be labeled
'Respondent Highest Degree.' Although variable labels can be longer, the Explore output display works best with variable labels that are 30 characters or less.

Change Decimals

Reduce the number of decimals displayed (and stored) for the selected variable. This option will permanently reduce the decimals displayed in a variable when the file is saved.

Value Labels

Use value labels to create a more detailed description of the values used in a categorical variable. The value label is used to create more meaningful output. If education is coded 0 through 4 to represent five general categories, the values would be more meaningful if labeled 0="< High School", 1="High school", 2="Junior college", 3="Bachelor", 4="Graduate". Without labels, the output will show the values 0 through 4. Limited to 12 value labels.

Missing Values

Identify up to four numerical or string values to be excluded from the statistical analysis (blank cells are always considered missing). Values that are less than (<) or greater than (>) a set value can also be set as missing.
Recode Values

Use the values of one variable to create a new variable. This is most useful when there are several categories in a variable that can be combined into fewer meaningful groups. As an example, recoding years of education into less than high school, high school, some college, and college graduate.

![Image of Recode Values]

Procedure

✓ Select the variable(s) to recode by clicking on the variable name in the list box.
✓ Select the operation to perform by clicking on a radio button. For coding less than 12 years of education, select the "<" button.
✓ Enter the value to use as the basis for the recode. For the above example, enter 12.
✓ Enter a variable name for the new variable (limit to 8 characters). Example "EduFmt" for education.
✓ Enter the new variable’s value to represent the old values. For this example, use 1 to represent anyone with less than 12 years of education.
✓ Click the Recode button.

A new column will be added to the spreadsheet that has the new variable name and the recoded value. Repeat the process until all values needed for the new variable have been recoded. A module is also available to conduct multiple recodes in a single operation (up 100). Click “Set Multiple Category Ranges,” select a variable, provide a name for the new variable, and then enter start and end values for each category range. The module automatically provides default values for the recodes but the values can be edited. Only use numerical data.

More than one variable can be recoded at the same time. Select and recode multiple variables by holding the Control button [Mac Command button] while selecting variables in the list box. Leave the variable name blank. Clarity will automatically name each new variable using the
original variable name and adding "-2" to the name. This option should only be used if the variables have the same range of potential values (e.g.; 1 to 10; 1,2,3,4; etc.).

Recode Strings

Create a new variable that recodes strings into 1 to 12 numerical categories. Recoding strings into values greatly simplifies analysis, control, and recode operations. Use the same recode value to combine multiple string categories into one meaningful category. Category values must be between 1 and 12. Enter a value label for each category. The module will display a list of strings in alphabetical order with category counts to help you decide which categories are most frequently used. Categories not coded will be treated as missing or use 0 to represent 'Missing' categories (recommended approach).

Procedure

- Open the Strings tab in the Recode module and select the variable you wish to change in the list box.
- Enter the values (1 to 12) to recode each string into a value.
- Enter a helpful value label.
- Enter a variable name for the new variable you are creating (limited to 8 characters).
- Enter a variable label for the new variable.
- Enter 0 for those strings you wish to make missing - do not enter a label.
- Click the Recode button.

A new column will be added to the spreadsheet that has the assigned variable name and the recoded values. The values and value labels will be added to the data file. The data file must be saved to retain the new variable.
Compute Variables

Create a new variable by mathematically manipulating the values of the original variable. This is most useful when transforming a variable such as by squaring (\(^2\)) or taking the square root (\(^.5\)) of the original values. It is also useful for applying a constant to all observations in a variable to remove negative numbers or to develop an index. The Compute procedure can be used to create an exact duplicate of a variable, the natural log of a variable’s values, a mean centered variable, or a variable converted to standardized (Z) scores.

![Compute Variables GUI](image)

**Procedure**

- Select the variable to use for the compute from the list box.
- Enter a variable name for the new variable (limit to 8 characters). Example “AGESQ” for squaring age in years.
- Select the operation to perform by clicking on a radio button.
- Enter the value to use in the compute statement. Note that \(^2\) = square and \(^.5\) = square root.
- Click the Compute button.

A new column will be added to the spreadsheet with the variable name and the computed values for all observations. Select and compute multiple variables by holding the Control button [Mac Command button] while selecting variables in the list box. Leave the variable name blank. Clarity will automatically name each new variable using the original variable name and adding “-2” to the name.

Combine Variables
Use Combine Variables to create a third variable that is a mathematical combination of two other variables. This feature creates a new variable that contains the result of the computation. Repeat the operation to systematically combine several variables into one.

**Procedure**
- Select the first variable to use from the list box and click the Var1 button.
- Select the second variable from the list box and click the Var2 button.
- Enter a variable name for the new variable (limit to 8 characters).
- Select the operation to perform by clicking a radio button.
- Click the Combine button.

**Find and Replace**

Use Find and Replace to change cell contents within one or more variables. This feature is most useful for correcting data entry errors or removing unusual strings such as dollar notation ($). The Undo button will return the cell contents to the immediate proceeding value. It is recommended that you conduct find and replace operations on a duplicate variable so you do not disturb the original variable. Use the Compute procedure to duplicate a variable.

This procedure can be used on one or multiple variables by holding the Control button [Mac Command button] while selecting variables.

If you wish to delete the value or string in the Find box, leave the Replace box empty. Also note that this module will first search for an exact match of the entire cell contents and replace those contents when found. If there is not an exact match, the module will then look for a match embedded within the cell contents.
Procedure

Use Find and Replace to change cell contents of variables selected in the list box. This feature replaces the contents of the selected variable instead of creating a new variable. It is useful for correcting data entry errors or unusual entries such as dollar notation ($). Leave the Replace box blank to delete the Find criteria when there is a match.

1. Select the variable to search by clicking on the variable name in the list box.
2. Enter the value or string of the search parameter.
3. Enter the value to use to replace the search parameter. Leave blank if you want to delete the value or string (example above replaces 1 with 0).
4. Click the "Replace" button.

Note: It may be helpful to create and use a duplicate variable before conducting replace operations so the original variable is not disturbed. Use the Compute procedure to duplicate a variable. After ensuring the recoded variable is accurate, the original variable may be deleted by selecting the variable in the data spreadsheet and deleting that column.

Trim Strings

Use Trim Strings to create a variable that contains a truncated version of the original values. As an example, banking data may present activity with common leading text such as “Deposit” or “Transfer” that is followed by unique date or transaction numbers. These unique text entries prevent aggregation of common transactions.
Procedure

- Select the variable `trim` by clicking on the variable name in the list box.
- Select the location of the trimming as either starting from the left, starting from the far right or a range in between.
- Enter the number of characters to include in the trimming operation.
- Click the "Trim" button.

Data Examples Menu

Use the example files to view variable and value formatting and to practice using statistical procedures and the Output Log. The Student Workbook data file is a good practice file for experimenting with the statistical procedures and becoming comfortable with the Explore, Data, and Output Log tabs. Use the Diagnostic Tests data file to try the Diagnostic Accuracy procedure. Use the Appraisal Ratios data file to try the Appraisal module. Additional data files and codebooks are available at http://www.acastat.com/data/data.html.

Calculate Menu

Run Procedure

Select Run Procedure to start an analysis. This has the same function as the Calculate button on the toolbar.
Clarity includes links via an Internet connection to a User Manual, a Student Workbook, and an Applied Statistics Handbook. To read these files, Adobe Acrobat Reader version 6.0 or later must be installed on the host computer. Acrobat Reader is a free download from Adobe. For more information, visit http://www.adobe.com for download information. Internet access is required to view the documentation.

Student Workbook

The Student Workbook contains over 20 lessons designed to introduce basic statistical concepts while also demonstrating the functionality of Clarity. The Appendix includes instructions for using SPSS.

File Name: Workbook.pdf
Availability: Website http://www.acastat.com/support/support.html

Free Applied Statistics iBook

The Applied Statistics iBook contains formulas, definitions, and examples along with annotated output from statistical software. It is designed to be a quick reference resource (not a complete statistics text). The iBook is available from Apple or as a pdf from http://www.acastat.com.

Availability: Website http://acastat.com/statbook/statbook.html
Statistical Procedures

Use the Procedures list to select the type of analysis. An options box is displayed for each procedure and analysis list boxes display depending on the type of procedure selected.

Analysis List Boxes

Variable List

Use the Variable List to select the variables for analysis. As an example, to produce frequencies tables for the variables sex and race, select the sex variable in the variable list box. A [+] indicator will appear in the Analyze header. To add the selected variable to the Analysis Variables list box, click the list box header or drag the variable to the Analyze list box. Use the same procedure to select and place the race variable in the list box. Clicking the Calculate toolbar button will produce frequencies for both variables. Output from the analyses will be displayed automatically and added to the Output Log.

Dependent Variable

The topmost variable list box is used for all univariate analyses (Frequencies and Descriptives) and one bivariate analysis (Correlation). At least one analysis or dependent variable must be placed into the Analyze list box to run the procedure.

Independent Variable

With the exception of Frequencies, Descriptives, and Correlation procedures, an independent variable must also be selected (another list box will appear below the Analyze list box). In the case of t-tests, subgroup values may be set in the options box for the independent variable so Clarity compares the correct groups; or, when left blank, a t-test comparison is conducted on all possible group comparisons.
Control Variable

A control can be established for most procedures. This limits the analysis to a specific subgroup. As an example, to compare mean education level by race for single adults locate the Adults variable name from the variable list box and drag it into the control box, click the "=" button and enter 1. This would result in an Explore Means procedure that excludes all family units with more than one adult. The control should be a value (number) not a string (character). If using a string control, only use the "=" control operator.

Variable Selection Indicators

Clarity uses three color indicators to identify selected variables and their placement in the list boxes. Blue for the top analysis list box, green for the second list box, and red for the control variable.

Tips for using the Control Box

If the “All Categories” option is available and selected, Clarity produces statistics for each category in the control variable. Some procedures will not allow or display the “All Categories” option. In this case, the specific control value must be set. The All Categories feature is handy for categorical controls such as sex and race but is not recommend for continuous controls such as age, body weight, and income unless they are first recoded into fewer groups as categorical values.

Statistical Procedures

Frequencies

This produces a listing of all the values in a variable, the number of times the values occur in the variable, and the percent of cases represented by each value. Frequencies can be produced on numerical and string data.

List Variables

This produces a listing of observations for up to ten variables. Observations are listed in the order they appear in the data file.
Descriptives

Calculates measures of central tendency (mean, median) and variation (variance, standard deviation).

Explore Means

Explore Means produces descriptive statistics (sum, count, mean, sample standard deviation, standard error, 95% confidence interval) for each subgroup in the selected categorical variable(s). Explore requires continuous data for the analysis variable. The categorical variable can be a numeric or string variable.

Crosstabulation

This is a useful procedure for representing the association between two categorical variables in a contingency table. Each cell in a contingency table represents a subgroup. For each cell, the count, row percent, column percent, and total percent are reported (default). This procedure also automatically produces a chi-square statistic. Crosstabulations are easiest to interpret when the number of rows and columns in the contingency table are 5 or less.

Column Variable Label: Respondent Sex
Row Variable Label: Race of Respondent

<table>
<thead>
<tr>
<th>Count</th>
<th>Col %</th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td></td>
<td>689</td>
<td>861</td>
<td>1550</td>
</tr>
<tr>
<td></td>
<td></td>
<td>77.33</td>
<td>74.67</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td></td>
<td>118</td>
<td>193</td>
<td>311</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13.24</td>
<td>16.74</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td>84</td>
<td>99</td>
<td>183</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9.43</td>
<td>8.59</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>891</td>
<td>1153</td>
<td>2044</td>
</tr>
</tbody>
</table>

T-tests of means

Compare the difference between two sample means. For t-tests, the dependent variable must be a continuous variable to produce the means for analysis. The independent variable must be a categorical variable (often it will be a dichotomous variable such as sex: 0=male, 1=female). Values representing the comparison groups for the independent variable can be set in the option frame. These values are used to separate the dependent variable into two groups. Be careful to use the correct values or there may be an error reported that indicates there are not enough cases to compute the t-test. If there are more than two values in the independent variable, the other subgroups will be excluded from the analysis. If the group values are left blank, a t-test comparison is conducted on all possible group comparisons.
Paired t-test

Evaluates the significance of the mean difference between matched pairs (e.g., a pre-post experimental study design). Paired t-test requires numerical data. Select the variable(s) in the variable list box and drag one variable to the Sample A list box and one paired variable to the Sample B list box. You can select multiple variables by holding the Control button [Command button on Macs] and clicking variables in the list box.

Wilcoxon Signed-Rank

Wilcoxon Signed-Rank is a nonparametric version of the paired t-test used to compare two related (correlated) populations, as is the case in pre-post treatment studies. In contrast to the paired t-test, the Wilcoxon Signed-Rank Test does not require interval-level (continuous) data. Select the variable(s) in the variable list box and drag one variable to the Sample A list box and one paired variable to the Sample B list box. You can select multiple variables by holding the Control button [Command button on Macs] and clicking variables in the list box.

Mann-Whitney U

Mann-Whitney U is a nonparametric test that evaluates the significance of the difference between two distributions of ordinal samples. This procedure is similar to the two-sample t-test but it does not require the assumption of continuous data in the dependent variable. It requires numerical data in the ordinal list box. Select the variable(s) in the variable list box and drag the ordinal variable to the Ordinal list box and one categorical variable to the Categorical list box. Select multiple variables by holding the Control button [Command button on Macs] and clicking variables in the list box.

Analysis of Variance (ANOVA)

Compare the difference between two or more sample means. Like t-tests, the dependent variable must be a continuous variable to produce the means for analysis. The independent variable must be a nominal level variable (e.g., race: 1=white, 2=black, 3=Asian, 4=Hispanic. or ordinal (e.g., political ideology: 1=conservative, 2= moderate, 3=liberal). ANOVA optional output includes summary statistics for each subgroup and the Bonferroni post hoc test for multiple comparisons.

Kruskal-Wallis

Kruskal-Wallis is a nonparametric test that evaluates the significance of differences between the distributions of two or more samples. This procedure is similar to one-way ANOVA but it does not require the assumption of continuous data in the dependent variable. It requires numerical data in the ordinal list box. Select the variable(s) in the variable list box and drag the ordinal variable to the Ordinal list box and one categorical variable to the Categorical list box. Select multiple variables by holding the Control button [Command button on Macs] and clicking variables in the list box.
Pearson Correlation

Calculates the Pearson correlation coefficient to represent the association between two continuous variables. There are two options: create matrix or a scatterplot. The scatterplot option provides basic statistics and displays a visual representation of the association between two variables. For multiple comparisons, a correlation matrix is often the best choice.

Spearman rho

Evaluates the linear association between two ordinal variables. This procedure is similar to Pearson Correlation but it does not require the assumption of continuous data in both variables. It requires ordinal numerical data. Select the variable(s) in the variable list box and drag the ordinal variables to the Analyze list box. Select multiple variables by holding the Control button [Command button on Macs] and clicking variables in the list box.

Point-Biserial Correlation

Evaluates association between one continuous variable and one dichotomous variable. This procedure provides results equivalent to Pearson Correlation but it includes additional options for descriptive statistics and a test for homogeneity of variance. It requires numerical data in the Continuous list box. Select the variable(s) in the variable list box and drag the continuous variable to the Continuous list box and one categorical dichotomous variable to the Categorical list box. Select multiple variables by holding the Control button [Command button on Macs] and clicking variables in the list box.

OLS Regression

Calculates simple regression (one independent variable) or multiple regression (two or more independent variables). Ordinary least squares (OLS) regression requires a continuous dependent variable. As an option, ordinary least squares regression will add new variables to the spreadsheet representing residuals or predicted values for each observation. OLS also allows models with no constant (regression through the origin). Summary statistics for the variables included in the model are automatically included in the output.

Logistic Regression

Logistic regression requires a dichotomous (two values) dependent variable. The two values in the dependent variable must be “0” and “1”. Logistic regression will automatically produce summary statistics and, as an option, will add a new variable to the spreadsheet representing the predicted values for each observation.
Diagnostic Accuracy

Diagnostic accuracy compares a test result to a dichotomous reference standard. It is used to evaluate the sensitivity and specificity of a diagnostic test. The reference variable can only be two values (usually 0 for non-disease and 1 for disease). This represents the gold standard to compare to the test results.

Each cell in the table holds counts. The test variable can be nominal, ordinal, or interval/ratio. The lowest value for the reference standard must always represent non-disease. The cutpoint (T+ >) represents the point beyond which the test is positive (disease indicated).

As an example, if possible test results range from 1 to 4, a cutpoint of 1 means a negative test result and a positive test is indicated by a value of 2, 3, or 4. A practice data file “Diagnostic Tests” is included with Clarity. The data represent the use of Power Doppler imaging to estimate the level of a patient’s carotid stenosis (blockage of the carotid artery).

Appraisal Ratios Procedure

The Appraisal Ratios procedure is used to evaluate tax assessment values of real estate. Use of this procedure requires data on the market (selling) price and the assessed value of homes to compute ratios and other statistics. A practice data file “Appraisal Ratios” is included with Clarity.

Simulation

Two simulation procedures can be used for manipulating data in the spreadsheet. They are designed to help students develop a better understanding of random sampling. The Random Sample option creates a random sample from a larger data file. Assuming the larger data file represents the entire population, a smaller random sample helps students discover how summary sample measures such as means and proportions will vary from population parameters.

Random Sample

This procedure is used to create a data file that is a randomly selected subset of the current spreadsheet. Enter the sample size desired. The sample size must be less than the total number of observations in the current data file. Click the Run button to replace the current data with a random sample.
Repeated Sampling

The Repeated Sampling procedure conducts multiple random samples on one continuous variable to build a data set that reflects summary measures of that variable for each random sample. This procedure helps students test the Central Limit Theorem:

**Central Limit Theorem**

As sample size increases, the sampling distribution of means approximates a normal distribution and is close to normal at a sample size of 30.

This procedure creates repeated random samples for one continuous variable in the current spreadsheet. For each sample (iteration), the mean and standard deviation are computed and placed into the spreadsheet. This results in a new data file where each row represents the summary statistics for one random sample.

Enter the size of the samples (n) and the number of random samples (Iterations) desired. Select a continuous variable to conduct the random sampling on and place it in the list box. Only one continuous variable will be used for the random sampling. Please note that the sample size must be less than the total number of observations in the current data. Click the Run button to replace the current data with a sampling distribution. When the sampling is complete, Clarity will run the Descriptives procedure on the mean variable.
Output Log

The Output tab displays the results of the analysis in a text window for editing, printing, and saving as a text file. Each time a statistical procedure is run, the results are automatically sent to the output log. Notes can be added to the output before printing or saving.

The Output text box uses Courier font. This is a fixed font that ensures proper alignment of the tables. When copying output to a word processor, the default font may not be Courier in the word processor. This will result in tables that are out of alignment. To correct this, highlight the output in the word processor and change the font to Courier. The font size may also need to be reduced to ensure the output fits widthwise on one page.
File Menu

New

Click the "New" button to erase all output.

Open Output File

This feature will load output saved from AcaStat Software statistical software products (AcaStat and StatCalc).

Save Output As

This menu item saves output as a text file with the file extension ".txt". This file can be opened in any word processor or spreadsheet application. Save all output or selected items in the output log. To select multiple items in the Mac version, press and hold the Command key while selecting items in the output log. For the Windows version, use the Control key for the same function.

Print Output

This menu item prints all output or selected items in the output log. To select multiple items in the Mac version, press and hold the Command key while selecting items in the output log. For the Windows version, use the Control key for the same function. By default, Clarity prints a date and time header and a Notes section on each page printed.

Edit Menu

Cut

Use this menu item to remove selected output text and save the text to the system clipboard.

Copy

Use this menu item to copy selected output text to the system clipboard.

Paste

Use this menu item to paste text from the system clipboard into the output screen.
Select All

Selects all the text in the output screen.

Convert Output for Spreadsheet Pasting

Use this option to convert output to improve copying and pasting tables into a spreadsheet. This option inserts tabs into the output so that pasted output aligns properly in the columns of a spreadsheet.

Use the Edit menu item Convert Output Spreadsheet Pasting to convert the text. The operation will automatically select the converted text.

Click the copy button and open a spreadsheet program, select a cell in the spreadsheet, and click the paste button in the spreadsheet. Check the results carefully to ensure the table is formatted and aligned properly.
Other Menu Items and Clarity Features

Help Menu

Clarity Help

Click Help to open the help module. The contents can be sorted or searched by topic.

Statistics Glossary

The glossary contains over 250 terms and definitions. The search button will find matching records and sort the list box so all topics matching the search are listed at the top in bold. Pressing the Enter key after entering a search term will also initiate the search. Press Reset before conducting another search.

Support Webpage

This links to the AcaStat Software support web page.

Check for Updates

The Help menu contains an option to check for updates. Clarity connects to the Internet and automatically compares the installed version with the version of the update. If a newer update is available, the appropriate installation file may be downloaded to the computer desktop.

The update option is only available for licensed users. It is not available for versions purchase from Mac App Store. Updates for Mac App Store purchases are automatically pushed to licensed users.

System Requirements

Mac OS X 10.7 or later
Windows Vista or later