

INTRODUCTORY LAB

INTRODUCTION TO STATCRUNCH 5.0

StatCrunch is a free web-based statistical software package containing all statistical features covered in introductory statistics courses. It is very easy to learn and produces web-compatible output. In this lab we will introduce you to the basic environment of StatCrunch. In particular, we will discuss the primary windows within StatCrunch, the menu items, and the format of dialog boxes. Moreover, we will discuss the basic data manipulation methods in StatCrunch. The lab is meant as an introduction to the statistical and graphical tools in StatCrunch discussed in detail in Lab 1- Lab 4 Instructions. No computer knowledge is assumed to follow the instructions below and the use of StatCrunch is explained with clear and easy to follow steps.

1. Getting Started

Before you start working with StatCrunch and its various features, you need to know a few things about Windows. Though StatCrunch can run on any of the three platforms (PC, MAC, and UNIX), almost all campus workstations use Windows XP or Windows 7 and also the vast majority of students use the system on their home computers. The instructions in this section is intended for those students who have had a very limited computer experience and are not familiar with the Windows operating system. All other students can skip the section and move to the next section.

Turn on the computer in your computer lab. The graphical user interface of the University of Alberta *Instructional Laboratories* is displayed.

Although you can communicate with your workstation with some combination of keystrokes, many pointing or choosing tasks are more easily done by using a pointer device such as a mouse, trackball, or touchpad. Moving a pointer device moves a mouse pointer, an onscreen graphic that in its most common form takes the shape of an arrow. Moving a mouse pointer over another object and pressing one of the buttons on the pointer device defines a mouse operation.

To select an on screen object (equivalently, to **click** on an object), move the mouse pointer (by moving the mouse) directly over the object and then quickly press and release the left mouse button. **To drag** or move an object, first move the mouse pointer over an object and then, while holding down the left mouse button, move the mouse. After the object has been dragged, release the mouse button. This technique is useful for moving windows and icons about on the screen. **To double-click on** an object, move the mouse pointer directly over an object and press the left mouse button twice in rapid succession.

If you are not familiar with the above mouse operations and Windows, double click *My Computer* icon on the desktop with the left mouse button, click *Help* once, and then on *Help and Support Center*. The *Help and Support Center* dialog window is displayed. Select *Windows Basics* in the left panel. The program covers the most important mouse and windows operations. Just follow the instructions.

There are several icons displayed on your desktop: *Computer*, *Recycle Bin*, *Authentication*, *My AFS Disk Space*, *Password Maintenance*, *Trouble Reporter* and *Google Apps*.

To see what any icon does, you double-click it. If that icon is a program, the program is started, and you see the program window. If that icon is a container icon like the *Recycle Bin* or *Computer*, you see the contents of the icon in a window. In other words, anything you do in Windows is displayed in a window.

The *Computer* program provides an easy access to computer resources and files. From this window, you can get to your computer drives, printers, network, and other resources. You can view what files, folders, and programs are stored on the hard and floppy disks. Files deleted from an application or from the *Computer* list end up in the *Recycle Bin*. You should click on this bin and empty it before signing off from the computer.

The great thing about windows is that they all work the same way; they all have the same set of controls. You open, close, move, resize, and scroll all windows the same way.

To resize a window, put the pointer on the window edge and drag. **To minimize** a window, click the *Minimize* button. The window is displayed as a taskbar button. **To maximize** a window, click the *Maximize* button. The window is enlarged to fill the entire screen. When a window is maximized, the maximize button changes to the *Restore* button. You can click this button to restore the window to its original size. **To move** a window, place your mouse pointer on the *Title bar* of the window and drag it to a new location.

The window with which you are working at any given time is called the *active* window. When more than one window is open, the windows are displayed so as to overlap. The active window appears in the foreground. You can change a window to active simply by clicking on the edge of the desired window.

2. **Launching StatCrunch.**

StatCrunch was designed and developed by a team of statisticians and programmers from the University of South Carolina. It differs significantly from other statistical packages. It is not a standalone application but a collection of Java applets embedded in HTML pages. The software can run on any of the three major platforms: PC, MAC, and UNIX. It only has the minimal requirement of a Java-capable Web browser (every computer shipped over the past few years has it). In particular, all campus workstations are equipped in Java-enabled browsers.

If you are working on your home computer and do not have a Java-enabled browser, you will probably see a gray box which may or may not have a red x in it after logging in to StatCrunch. You may also check whether your browser is Java-enabled by accessing the web site <http://www.javatester.org/enabled.html>.

StatCrunch can be accessed at <http://lucky.nic.ualberta.ca/statcrunch5.0/>. For your convenience, the above link is also provided at www.stat.ualberta.ca/statslabs/stat151/index.htm which is the address of *Stat 151 Laboratories* web site. You will be required to enter a password to access the software. The password will be provided by your course instructor.

In order to make sure that students will be able to use StatCrunch when needed, a backup server has also been provided and can be accessed at <http://leo.nic.ualberta.ca/statcrunch5.0/> Please use the backup server if you experience persistent problems with downloading StatCrunch from the primary server. The link to the backup server is also provided on Stat 151 Laboratories web site.

To launch StatCrunch, click the *StatCrunch* button located in the *Introductory Lab* box on *Stat 151 Laboratories* web site and enter the password. The program opens in a window called *data table* displayed below:



Row	var1	var2	var3	var4	
1					
2					
3					
4					
5					
6					
7					
8					
9					

3. Data Table

The data table consists of a grid of cells used for entering, editing and viewing data in StatCrunch.

Note that the top row in the data table is used to provide labels for each column. The default column names are var1, var2, and so on. To change the column name, put the cursor in the cell containing the column name and you may either use the DELETE key on the keyboard to delete the current name and enter the new name or use the BACKSPACE key to delete the characters of the current name in one by one fashion beginning with the last character and enter the new name. Try it and enter the following three column names (in this order): *Gender*, *Age*, and *Systolic*. The leftmost column assigns successive whole numbers to the rows in the data table and serves as row ids for the data table.

The cell that is highlighted receives keyboard input. To enter a value in a data table cell, just click the cell, type a value, and press ENTER. The ENTER key moves the highlighted cell one cell down.

The up, down, right and left keys shift the highlighted cell in the direction of the arrow. All of these keys can be used to increase the number of cells displayed in the table. For example, if the data table currently displays 6 columns and the user presses the right arrow key while in the highlighted cell in the 6th column, StatCrunch will then display the 7th column with the highlighted cell in the 7th column.

You may think of the cursor as occupying the position to the far right in the highlighted cell. The cursor may not be manually positioned inside the highlighted cell. Enter the number “123” into any cell and try to replace the second digit in it by “4” to demonstrate this.

To delete the old value in a cell, click the cell, press DELETE key on the keyboard, enter the new value and press ENTER key. In case, you wish to change only a portion of the cell contents, click the cell and use the BACKSPACE key to make the changes. To practice, replace the number “123” in a cell by “167”. To erase the last value entered, use UNDO CELL EDIT command in the EDIT menu. Apply UNDO command to the highlighted cell with the number “167” in it.

Although the data table has rows and columns, it is not a spreadsheet like Microsoft Excel. In StatCrunch, cells contain values that you type or generate with menu options. Cells do not contain formulas that update based on other cells.

Nevertheless, you may select a rectangular block of cells (highlight a rectangular area in the data table). In order to select a block, drag diagonally from the first cell (one corner) to the last cell (opposite corner of the block). Then release the mouse button. To practice, enter two columns of several numbers each and select the block.

The horizontal and vertical scroll bars at the top and bottom of the data table make it possible to move quickly through the table. Also the PAGE UP and PAGE DOWN keys can be used to quickly scroll through the table.

The columns will correspond to variables (height, weight, score, etc.) used in data file and rows will correspond to the cases (subjects or objects). We will demonstrate the basic StatCrunch features using the following example.

Example The Framingham Heart Study followed a cohort of 5209 men and women for over 25 years. The study has been important in identifying risk factors associated with cardiovascular disease. The following table contains data for a random sample of 14 subjects from the study. The following is a description of the variables we have selected from the study for our purpose:

<u>Column</u>	<u>Description of Variable</u>
1	Gender (M-Male, F-Female),
2	Age (30-64 years),
3	Systolic blood pressure (82-300 mm).

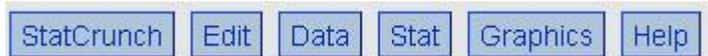
Gender	Age	Systolic
F	59	170
M	35	130
M	46	136
F	43	96
M	53	120
M	50	110
M	33	100
M	57	145
F	41	132
F	40	112
M	54	140
M	53	148
F	53	165
M	49	100

Click the link *Data in the Introductory Lab* panel on *Stat 151 Labs* website to download the data to your computer. The example will be used throughout the section to demonstrate the basic features of StatCrunch.

Warning: If you require to access another website during your StatCrunch session, open the site in a new browser window, otherwise your StatCrunch session will close without any warning and all your data and results will be lost. This is not any StatCrunch bug, this is how all web-based applications work. Please keep this in mind!

4. StatCrunch Menu

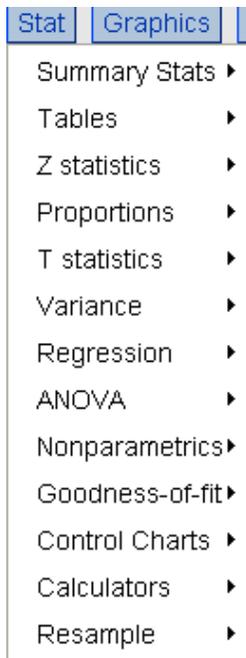
All commands in StatCrunch can be accessed with the menu bar. The StatCrunch menu bar consists of six buttons shown below: *StatCrunch*, *Edit*, *Data*, *Stat*, *Graphics*, and *Help*. The name of the file (*Untitled* in the window below) is always shown on the right side of the menu bar.



Selection of one of the above buttons results in the display of a pull-down menu, a list of more choices.

StatCrunch button allows you to organize your display so that data and results of statistical analysis occupy the specified and convenient for the user areas of the desktop. Use *Edit* button to undo the last change in the Data Table; cut, copy and paste data values including values in a blocks of cells; to delete or insert rows or columns. *Data* button allows you to choose the method to bring data into StatCrunch and also perform a variety of actions on cells, rows, and columns in the Data Table before you begin your analyses.

Stat button gives you an access to all statistical procedures available in StatCrunch. All the categories in *Stat* menu are followed by an arrow, which indicates that there are several analysis procedures available within the category; they will appear on a submenu when the category is selected.



Graphics button provides an access to many graphical techniques to display and explore your data. The graphical tools in StatCrunch will allow you to create graphs, and change almost every aspect of the graphs' appearance. Moreover, you will be able to brush the data points in graphs to see the corresponding values from data table.

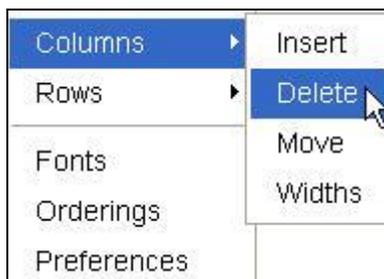
The last button in the menu, *Help* is linked to the StatCrunch help page and provides you with an overview of all StatCrunch features. The detailed discussion of the six buttons in the menu is given in Section 6.

5. Dialog Boxes

The data to be analyzed is displayed in the data table located below the menu bar. After selecting a menu item, a drop-down menu will appear with a list of the available procedures for the menu item. When one of the procedures is selected, a dialog box for this particular procedure will appear and the user is prompted for inputs for the requested procedure. In general, dialog boxes are used to display status messages or prompt you to make choices or supply additional information. After providing the information and clicking *Okay* button in the dialog box, the procedure will be applied to the data in the current data table.

5.1 Select Column

For example, select *Columns* in *Edit* menu and *Delete* for the data table with the data from the Framingham study:



The resulting dialog box displayed below consists of two panes; the left pane displays all the variables in the current data table that are valid choices, the right pane will list the variables selected for the operation of interest (in this case the deletion of the selected variables).



To select a variable click the variable you want in the left pane. The variable selected will be displayed in the right pane and the operation of interest (in this case the deletion) will be applied to the variable only. The question mark button ? in the dialog box links the user directly to the relevant information for that procedure. Clicking the *Cancel* button closes the dialog box and cancels the operation.

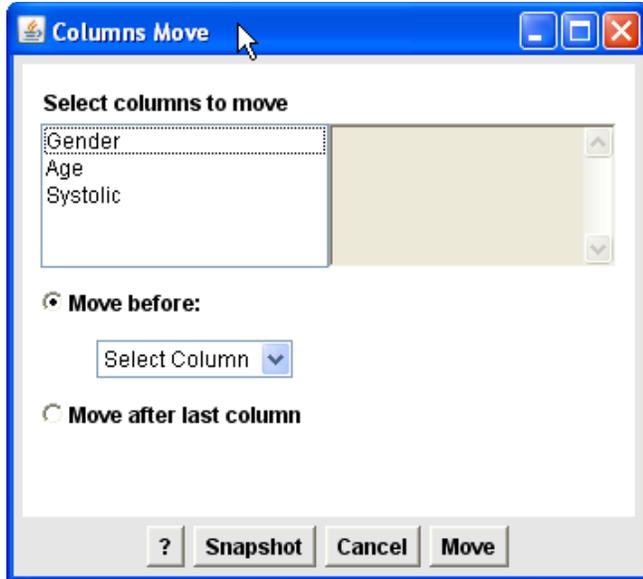


Clicking *Delete* button will remove the column *Age* in the data table. Please do not click the *Delete* button as the variable *age* will be still used in other examples.

5.2 Radio Buttons

Radio buttons are often used in StatCrunch dialog boxes. Radio buttons are displayed as a list of circular holes that can contain white space (for unselected) or a dot (for selected). Adjacent to each radio button is shown a caption describing the choice that this radio button represents. Only one button in a group of radio buttons can be selected at one time. Therefore, when the user selects a radio button, any previously selected radio button in the same group becomes deselected.

Selecting a radio button is done by clicking the mouse on the button. In order to illustrate this, click *Edit*, and then on *Columns*, and finally on *Move* submenu. The following dialog box containing radio buttons will be displayed:



The two radio buttons in the dialog box represent two mutually exclusive choices; the selected column can be moved before another column or after the last column. Only one of the buttons can be selected at one time to move the selected column to the new location.

5.3 Check Boxes

StatCrunch also uses check boxes in its dialog boxes. Check boxes are similar to radio buttons but unlike radio buttons they allow the selection of more than one optional action at a given time. For example, selecting *Load data* from file in the *Data* menu opens the following dialog box with two check boxes:



Clicking *Okay* or *Calculate* causes StatCrunch to execute an operation with the current values and choices as shown in the dialog box.

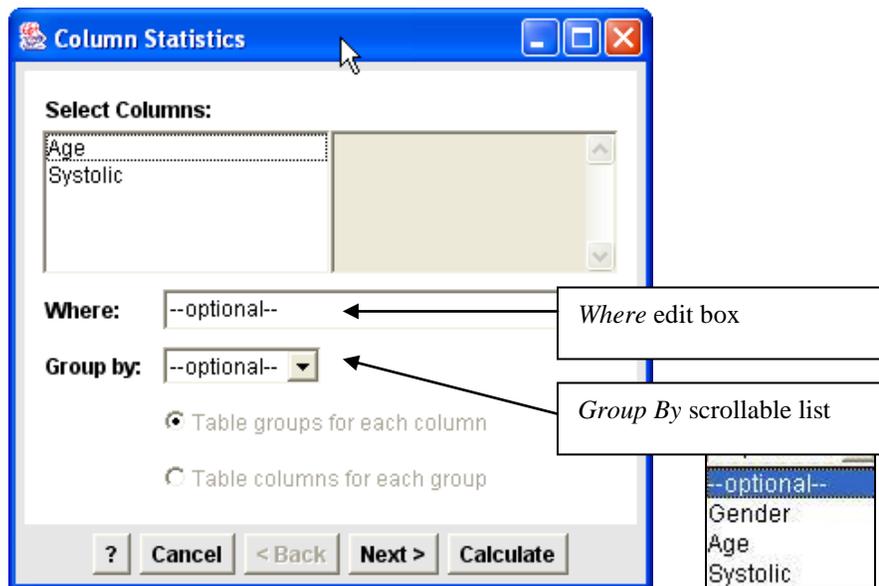
Other examples of dialog boxes will be provided while discussing particular StatCrunch features. Some of the dialog boxes will have scrollable list of items to choose from, edit boxes which provide an area into which a value can be typed (often combined with a scroll-down list to provide an alternative to the typing of a value)

5.4 Using Where and Group By

Sometimes we may be interested in obtaining the summary statistics or graphs for a subset of a data set. For example, we might be interested in obtaining the summaries or graphs of systolic blood pressure in the Framingham Heart Study, separately for males and females. Or we might be interested in obtaining the summaries and graphs for subjects satisfying some age criteria (for example older than 45). In either case, the statistical or graphical procedures available in StatCrunch would have to be applied to a subset of data. This does not require that the data be split into several subsets corresponding to the grouping variable (gender in our case), it is just the procedure that will be applied to the data subsets only.

The *Where* and *Group By* statements in StatCrunch dialog boxes enable you to apply the selected statistical or graphical procedures to subsets of data set. We will illustrate the function of the two statements with an example.

Suppose we wish to obtain the summary statistics for the variable *Systolic* in the Framingham Heart Study. In order to obtain the summaries, choose *Summary Stats (Columns)* in the *Stat* menu. The corresponding dialog box will have the following form:



In order to obtain the summaries for the variable *Systolic* click the variable in the left pane to move it to the list of selected variables in the right pane. The *Where* edit box provides an area into which a logical expression can be entered that will restrict the analysis to only those rows that satisfy the condition.

For example, you may enter “Age>45” (the expression inside the quotation marks) to obtain the summary statistics only for subjects older than 45 years of age. In order to obtain the summary statistics for males older than 45 years old, you would have to enter the following expression:

Age>45 AND Gender=”M”

If you were to obtain the summaries for the subjects who are either males or older than 45, you would have to replace AND in the expression above with OR. Finally, if you wished to exclude the subjects 45 years old from the analysis, you would have to enter: “Age<>45”.

Notice that in order to obtain the summaries for males only, we could enter: Gender="M" in the *Where* edit box. Quotation marks are used on M because *Gender* is a categorical variable with labels "F" and "M" as possible values.

The *Where* box is optional, that means that if you do not enter anything into the box, the summary statistics will be obtained for all subjects (rows) in the data file.

The *Group by* scrollable list (also optional) displays a list of items for selection. The list allows the user to group results based on the column selected in the drop-down menu. Clicking the arrow in the box opens the list (displayed on the right of the dialog box). In our particular case, the list consists of the three variables *Gender*, *Age* and *Systolic* in the data file. Selecting *Gender* will produce the summary statistics for each value of the variable *Gender* i.e. for males and females.

Finally, if the *Group By* column is selected, the two radio buttons allow you to choose whether to display the output in separate tables for each column selected or in separate tables for each group. Selecting one radio button always deselects or clears the other option

6. Overview of the Menu Items

Now we will discuss in detail the six menu items in StatCrunch menu: *StatCrunch*, *Edit*, *Data*, *Stat*, *Graphics*, and *Help*.

6.1 StatCrunch Button

Click the first button *StatCrunch* in the menu bar. Several options to display your results are displayed. In the default option (*Session results...Display in... window*), each output obtained in your session (graph or result of a statistical analysis) opens in a separate window (one for each output). The results can then be copied and pasted into your report. The default option (displaying results in a window) is the most convenient and effective way to export the output to the Word processing application (for example, Microsoft Word) and produce lab reports.



Nevertheless, it is worthy to mention about other methods to display the output. For example, you may request that the StatCrunch window split into *Data* window (*Data Table*) and *Results* window, with the latter containing all the session results. In this case, the location of the *Results* window on your desktop can be specified by choosing one of the four available options *right half*, *left half*, *top half* and *bottom half*.

In case you request that either *Data* or *Results* windows be displayed, you should choose *full frame* option. In this case two buttons called *Data* and *Results* will be displayed at the bottom of each window that would allow you to switch to either window. The button that is highlighted indicates the window that is currently displayed.



In the above case, the *Data Table* is displayed and available to the user. In order to switch to the *Results* window, click the *Results* button.

Before, you explore some other features of StatCrunch, please reset your output settings so that your output is displayed in separate windows (default option).

6.2 Edit Button

Click the *Edit* button in the menu. The following window with several options to choose from opens:



If you select Undo or Redo (in the screenshot displayed above these options are faded because no data have been entered yet), the last change in Data Table is undone or redone, respectively. The *Cut* option allows you to delete the contents of a single cell or a block of cells. The options *Copy*, *Paste* allow you to copy and paste the contents of a single cell or a block of cells.

You can delete any column or several columns in your data table by selecting *Columns* and *Delete* option and selecting the columns to be deleted. You can also move any column in the data by clicking the *Move* option and specifying the new column location. Finally, you can clear the data table by selecting *Clear* option. You can insert or delete rows in your data table by selecting Rows option and clicking either *Insert* or *Delete* options.

The width of any column can be adjusted by clicking the right edge of the column and dragging the mouse. Alternatively, you may click *Columns* and then *Widths* option to specify the minimum column width (in pixels) and set other preferences for the default width for columns. By default, a column's width is set to 65 pixels. The column width will be automatically adjusted if data entered exceeds the default width.

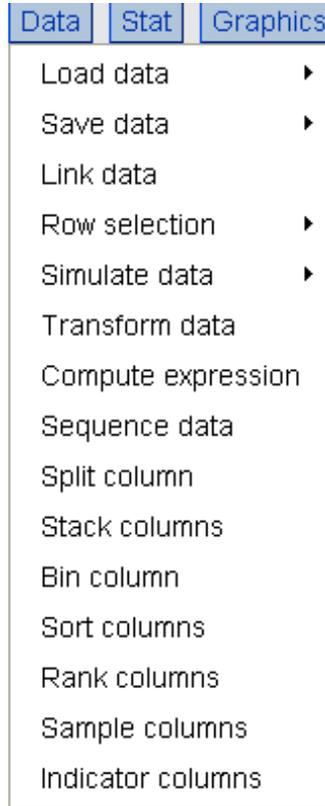
You can also change the font type and size. The width and height of cells in the data table will be automatically adjusted. By default, the font size is set to 12 and the default font type is "SansSerif".

6.3 Data Button

This item in the menu is used to load data into data table or to rearrange the data before you begin your analyses.

StatCrunch offers a variety of methods for loading data. In particular, data can be either loaded from your workstation (home computer) or network. Data can also be simulated in StatCrunch from one of several probability distributions.

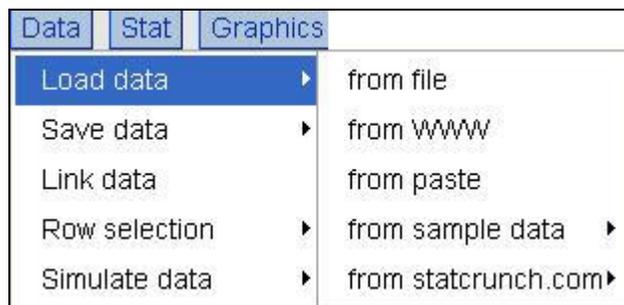
Click the *Data* button in the menu bar. The following submenu opens:



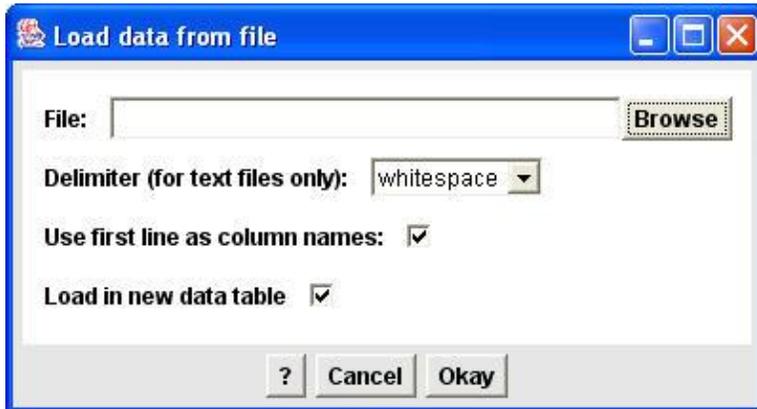
Below we will discuss some of the most important features in the *Data* menu.

(a) Load Data

Data can be loaded into data table in one of the following four ways: by importing a file from a local machine or network, by typing in the URL of a file on WWW, by pasting data directly into data table, and by loading the data from StatCrunch database. The above options are displayed when you click *Load data* in the *Data* menu:



StatCrunch can load text and Excel files from a local computer or from a WWW address. To load data from a file, select this option from the *Load Data* submenu. The following dialog box opens:



Now you can either click *Browse* button and select your file from your local directory or type the full path to the file. If the first line of the file contains column names (*Gender*, *Age*, and *Systolic*), make sure that the *Use first line as variable names* option is checked.

In case of text files you will have to specify delimiter; this step can be ignored if Excel file is to be loaded. Delimiters are commonly used in computer files to separate data values. For example, whitespace, tab, comma, or semicolon can be used in StatCrunch as possible delimiters between fields. For example, the text file saved in Notepad can have a form

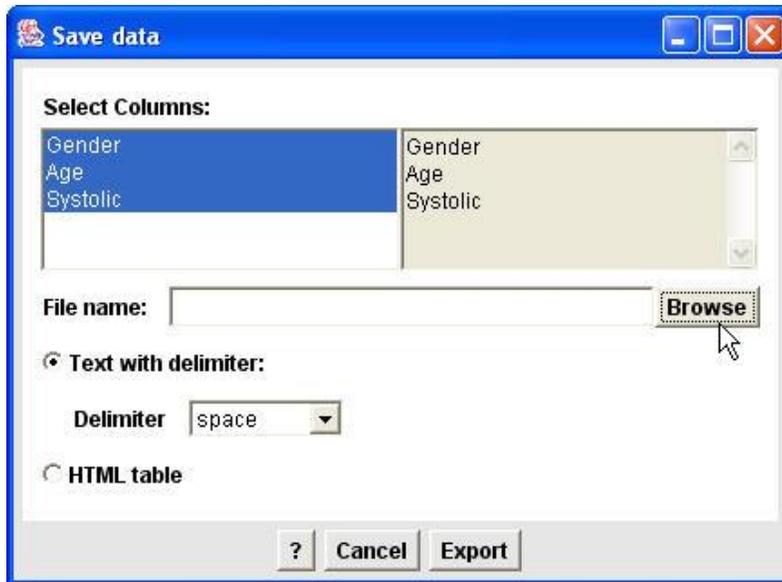
Gender	Age	Systolic
F	59	170
M	35	130
M	46	136

The file can be loaded into StatCrunch with the column names, *Gender*, *Age* and *Systolic* respectively by filling the above dialog box as shown above.

You may also paste data into StatCrunch with “from paste” option. However, as this option is not very consistent across platforms, loading data from a file is recommended. There is a database of interesting real-world data sets in StatCrunch that can be accessed by clicking *Sample Data*.

(b) Save Data

The data in the data table can be saved as a file in the user’s local machine by selecting the *Save Data* option under the *Data Table* submenu of the *Data* menu. You can save the data either as a text file or as an HTML file. For the text option, you will have to specify the delimiter to be used. The text option should be used when the data may be reused in the future. The HTML option should be used to export the data for publication on the Web or to paste the data into a word processing application.



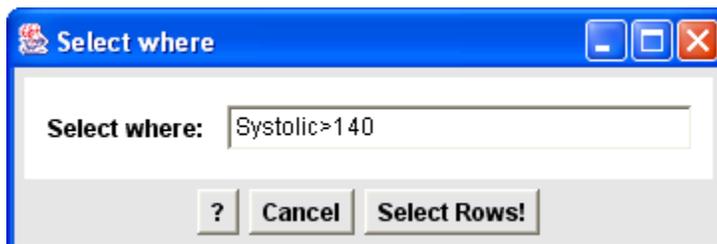
Click *Browse* to specify the desired directory and enter a name for the file. Click *Export* to save the file in the directory.

(c) Row Selection

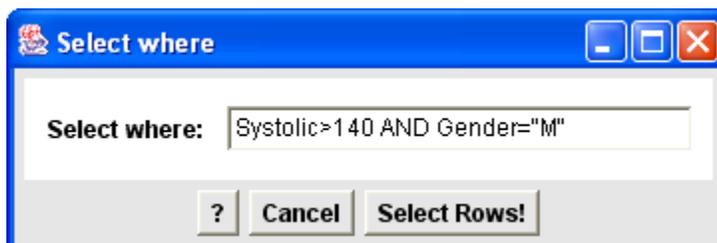
Rows in the data table may be highlighted by clicking on the appropriate cell in the column containing the row numbers (leftmost column in the data table). Clicking on the cell inside a highlighted row will deselect the row. All rows in the data table may be deselected by clicking on the Row label in the top row of the leftmost id column.

Select Where

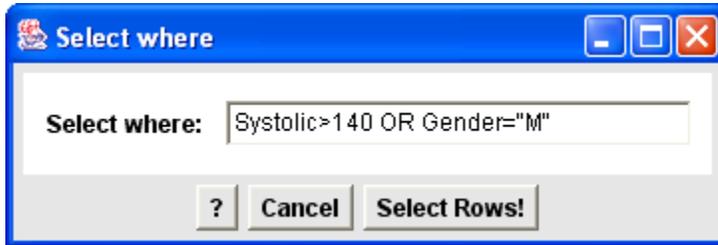
This feature will highlight rows in the data table that satisfy a logical expression specified by the user. For example, to highlight all rows that have *Systolic* value exceeding 140, we should enter the following expression into the dialog box:



To highlight all rows that correspond to all males with the *Systolic* blood pressure exceeding 140, we should fill in the entry box as follows:



In order to exclude all rows that produce the *Systolic* value of 100, we enter the following expression: `Systolic<>100`. In order to highlight all rows that correspond to either males or have *Systolic* variable value exceeding 140, we should enter the following expression:

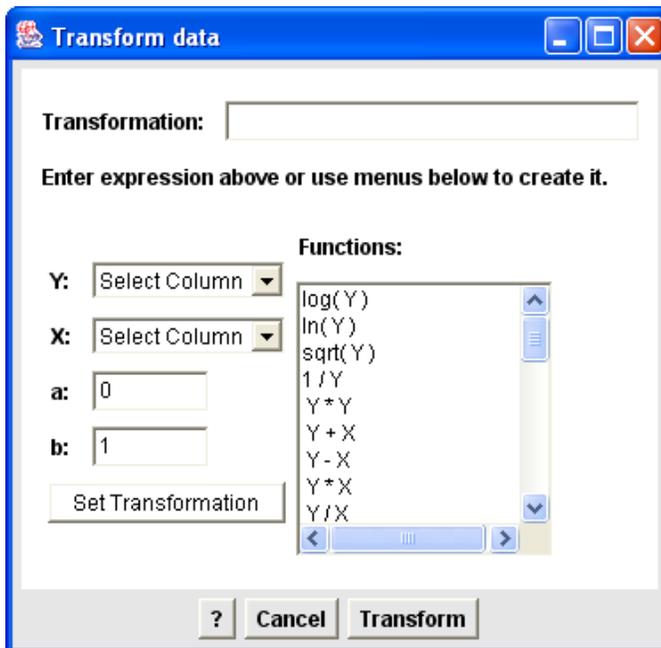


Deselect all

This option allows the user to deselect all rows highlighted in the data table. All rows in the data table may also be deselected by clicking on the Row label in the top row of the leftmost id column in the data table.

(d) Transform Data

Occasionally it is useful to apply transformations to the data to answer some questions of interest. StatCrunch offers a variety of techniques to transform your data. Click the *Transform Data* option in the *Data* menu. The following dialog box opens:



As the systolic blood pressure of 135 is considered to be normal, we can calculate the excess systolic blood pressure as the difference between the systolic blood pressure for each subject and 135. Thus, enter in the *Transformation* entry box the following expression: `"Systolic -135"`. A new column *Systolic-135* will be added in the data table.

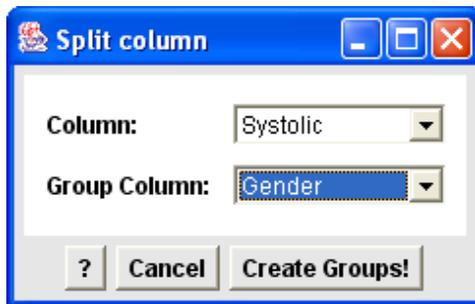
You may also construct some other transformations from the menu of common transformations in the *Functions* right pane. For example, to obtain a new variable that is the sum of two columns, you may select the columns Y and X and in the left pane of the dialog box and then click "Y+X"

expression in the *Functions* pane. Next click *Set Transformation* to have the proper expression placed in the *Transformation* input box and finally click the *Transform* button to compute the values for the new column. The new column will be added to the right of any existing columns in the data table.

You can also use the transformations like “a+bY” or “a+b/Y” specified by the two parameters a and b to be entered in the left pane entry box.

(e) Split Column

This option splits values in a column into separate columns, one column for each value in a grouping column. For example, suppose that we wish to obtain the systolic blood pressure readings for males and females separately (in two different columns, one for females, and the other for males). In this case, the dialog box for the *Split column* feature should be filled in as follows:



Click *Create Groups* button to create the two new columns: “F Systolic” containing the observations for females and “M Systolic” containing the observations for males.

F Systolic	M Systolic
170	130
96	136
132	120
112	110
165	100
	145
	140
	148
	100

Remark: Notice that some items in the *Data* menu have not been discussed above. Some of them will be discussed later in Lab 1-5 Instructions (i.e. *Sample Columns* and *Simulate data*) that accompany lab assignments that use the features. Others are not critical for the labs in the course.

(f) Bin Column

This option assigns a class interval to each observation in a column specified by the user. The class intervals are non-overlapping and cover the whole range of the data for the column. The endpoints of the class intervals are called bins. The bins can be either automatically generated by StatCrunch or specified directly by the user. The following convention is being used while constructing the class intervals: the left-endpoint is included, the right endpoint is excluded. As a consequence, each observation in the column falls into exactly one class interval.

The dialog box for this feature is shown below. First you will have to select the column to be binned from the drop-down menu. Then you will have to choose the method to define the bins. Three radio buttons provided in the dialog box allow you to choose the appropriate binning scheme.

First bins can be defined automatically by StatCrunch, in this case the class intervals determined by the bins have equal width. You may also use fixed width class intervals, with this option, the user specifies a starting point for the sequence of bins and defines the class interval width (bin width). Finally, bins can be defined by cut points, in this case the user specifies the sequence of comma delimited values for class interval endpoints (bins). This option can be used to create class intervals of different widths.

The dialog box 'Bin column' has the following settings:

- Bin column:** Systolic
- Define bins:**
 - Automatically
 - Use fixed width bins
 - Start bins at:** 96
 - Binwidth:** 20
 - Use cut points
 - Cut points:** (empty field)
- Bin edges:**
 - Include left endpoint
 - Include right endpoint

Buttons: ? Snapshot Cancel Calculate

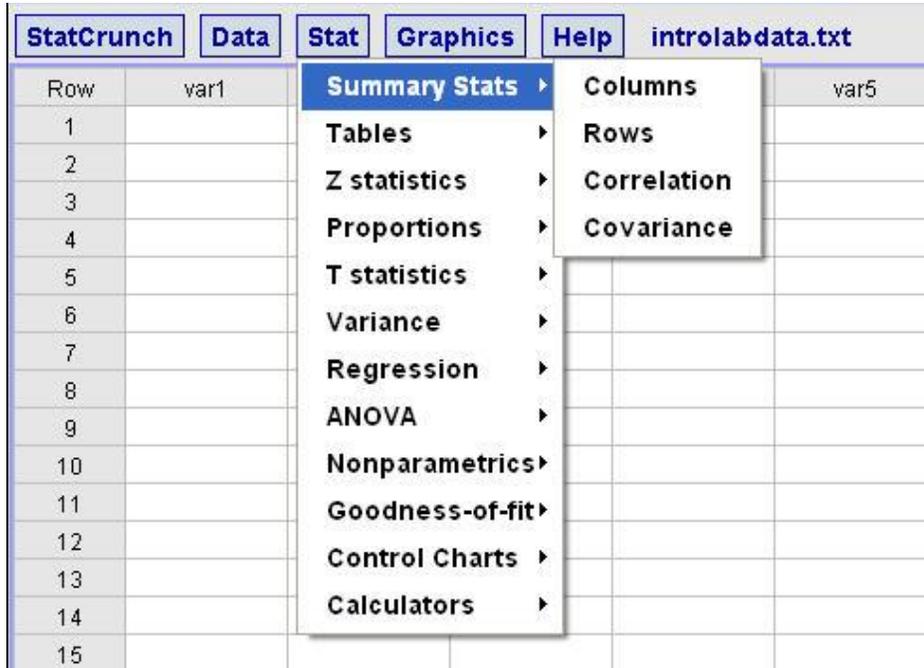
Systolic	Bin(Systolic)
170	156 to 176
130	116 to 136
136	136 to 156
96	96 to 116
120	116 to 136
110	96 to 116
100	96 to 116
145	136 to 156
132	116 to 136
112	96 to 116

The above Bin(Systolic) column was obtained for the automatically generated bins in StatCrunch.

6.4 Stat Button

Stat button gives you an access to all statistical procedures available in StatCrunch. Each of the twelve categories listed in the Stat menu is followed by an arrow, which leads to several analysis procedures available for the category.

The procedures under the *Stat* menu allow the user to obtain the numerical summaries of the categorical and numerical data (in particular, frequency tables and measures of center and spread for numerical data), explore the relationships between two categorical variables (contingency tables) and the relationship between numerical variables (regression and correlation), and make statistical inferences.



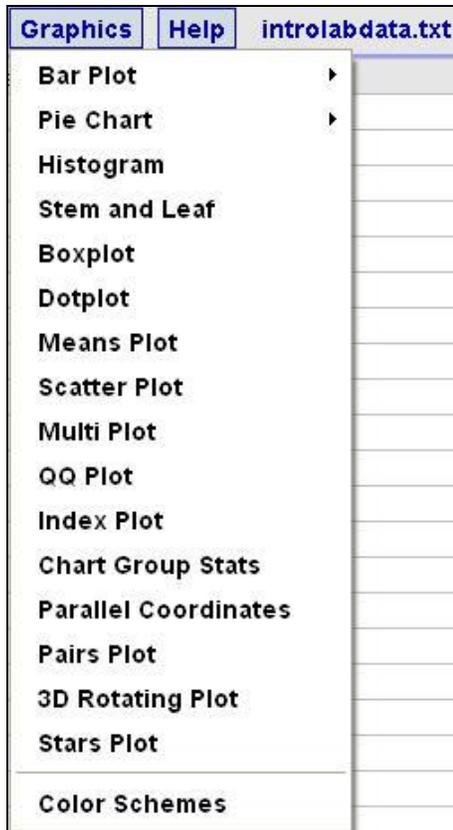
The statistical procedures in the *Stat* menu will be discussed in *Lab 1-5 Instructions* when the corresponding topics will be discussed in the *Labs 1-5*.

6.5 Graphics Button

StatCrunch offers a variety of graphical tools to display and analyze data. The vast majority of the graphs are interactive. In particular, StatCrunch supports brushing so you will be able to brush the data points in graphs to see the corresponding values from the data table. Moreover, you will be able to change almost every aspect of the graphs' appearance.

(a) Graphs in StatCrunch

Click the *Graphics* button. The following submenu will be displayed.



Some of the graphs allow the user to display and summarize the categorical data (*Bar Plot* and *Pie Chart*) and numerical data (*Histogram*, *Stem and Leaf*, *Boxplot*, *Dotplot*). Other graphs allow the user to examine the relationship between two numerical variables (*Scatter Plot*, *Pairs Plot*) and among three numerical variables (*3D Rotating Plot*). Some of the graphs will be discussed in detail in *Lab 1-5 Instructions*.

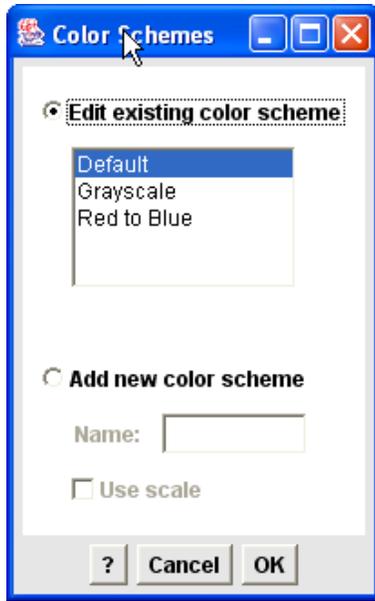
(b) Color Schemes

The last item in the Graphics menu called “Color Schemes” allows the user to specify the colours that will be used in StatCrunch graphs. As this is a general StatCrunch feature we will discuss it here in more detail. You may omit the material covered in this section unless you are planning to edit existing colours or add new colours in StatCrunch.

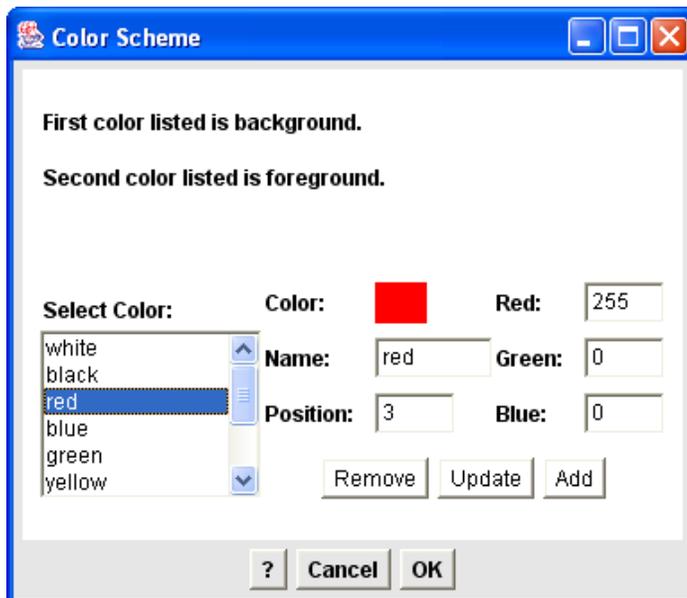
A color scheme is defined by an ordered sequence of colors that are accessed in succession when StatCrunch produces a graphic. One color from the sequence is defined as the background color for the graphic, and one color is defined as the foreground color which is the default color of axes and other standard graphical elements. The background and foreground colours are excluded from the list of colours used by StatCrunch in multiple-colour graphs.

The default color scheme consists of the sequence: black, white, red, blue, green, yellow, cyan, orange, dark green, and gray with black being the foreground and white being the background. StatCrunch also offers two standard color schemes which consist of a range of colors between two primary colors: *Grayscale (white to black)* and *Red to Blue*. The background and foreground for these color schemes are white and black, respectively, and these definitions may not be changed.

In order to edit an existing color scheme or add a new color scheme, select the *Color Schemes* option in the *Graphics* menu. The following dialog box will be displayed:



With the “Default” selected, click OK. The following dialog window for editing the “Default” color scheme will be displayed:



Options are provided for selecting the background and foreground colors from the list. StatCrunch defines colors using RGB values.

A brief introduction to RGB colour scheme follows. In the RGB colour scheme the three colours **R**ed, **G**reen, and **B**lue are combined in various ways to reproduce other colours. A color in the RGB color model can be described by indicating how much of each of the red, green and blue colour is included. Each can vary between the minimum of zero (no color) and maximum of 255 (full intensity). If all the colors are at minimum of zero the result is black. If all the colors are at maximum of 255, the result is white. In other words, each colour (RGB value) is specified by the ordered sequence of three integers between 0 and 255, each representing red, green and blue intensities, in that order.

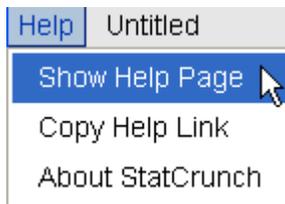
Let's examine the above dialog box in detail. When the color is selected from the color list (in our case "red"), the RGB attributes of the color as well as its position in the sequence are displayed. These properties may be changed by clicking the *Update* button.

New colors may also be added to the list by clicking the *Add* button after specifying a name for the color, its RGB attributes, and its position in the sequence. A selected color may be removed from the list by clicking the *Remove* button. When adding/editing a scale color scheme, the RGB values for the starting and ending colors of the color sequence must be specified.

When you finish editing the colour scheme, click the *Save* button.

6.6 Help Button

StatCrunch comes with an extensive online Help system. Click the *Help* button. The following sub-menu is displayed:



The first item "Show Help Page" provides links to several sources of information about StatCrunch.



The first item on the above list allows the user to access StatCrunch Help containing complete documentation on each feature. StatCrunch also has contextual help which can be obtained by clicking on the question mark button ? in the dialog box for a particular feature being used. In order to see it, select *Stat* in the menu, *Summary Stat*, *Columns* and click on the question mark.