
StatCrunch

Background Material and Guided Tours

Tom Trolen

Division of Business and Computer Information Systems
Scottsdale Community College

Table of Contents

INTRODUCTION.....	1
STATCRUNCH.....	1
OVERVIEW	1
PATIENCE, PATIENCE!	2
EXAM SCORE SCENARIO	3
GUIDED TOUR #1	4
<i>Tabular Methods</i>	5
<i>Graphical Methods</i>	9
<i>Numerical Summaries</i>	12
<i>Comparative Analysis</i>	12
GUIDED TOUR #2.....	14
REVIEW QUESTIONS.....	16
TURN IN.....	17
INDEX	18

Introduction

This document presents information regarding *StatCrunch*, the program you will use for your data analysis projects. *StatCrunch* relieves you of the computational burden associated with statistical analysis. *StatCrunch* does the number crunching, you interpret its output.

You will soon receive your first project, so you need to become proficient with this material right away. Over the next few days, read and study this document and work through the Guided Tours.

StatCrunch

StatCrunch is a fee-based web-delivered statistical analysis program that you can use on any computer that has an Internet connection and a Java-enabled browser, such as Internet Explorer or Firefox. You use *StatCrunch* by launching a web browser, then typing the address of the *StatCrunch* homepage, <http://www.StatCrunch.com>. If you've ever used a *spreadsheet* such as *Excel*, *StatCrunch* will seem familiar. Whereas a spreadsheet is a general-purpose tool, *StatCrunch* focuses on statistical data analysis. Imagine a large *worksheet* that has room for many rows and columns. You enter data for each variable into a worksheet *column* and use menu commands to tell *StatCrunch* how to manipulate the data. You then interpret the results.

Since no textbook covers *StatCrunch* features and commands, you'll use the background material presented in this and other handouts. You can also use *StatCrunch's* on-line *Help* system.

Overview

Figure 1 illustrates what you'll see when you view the *StatCrunch* homepage. It contains fields for existing users to login and link to subscribe as a new user.

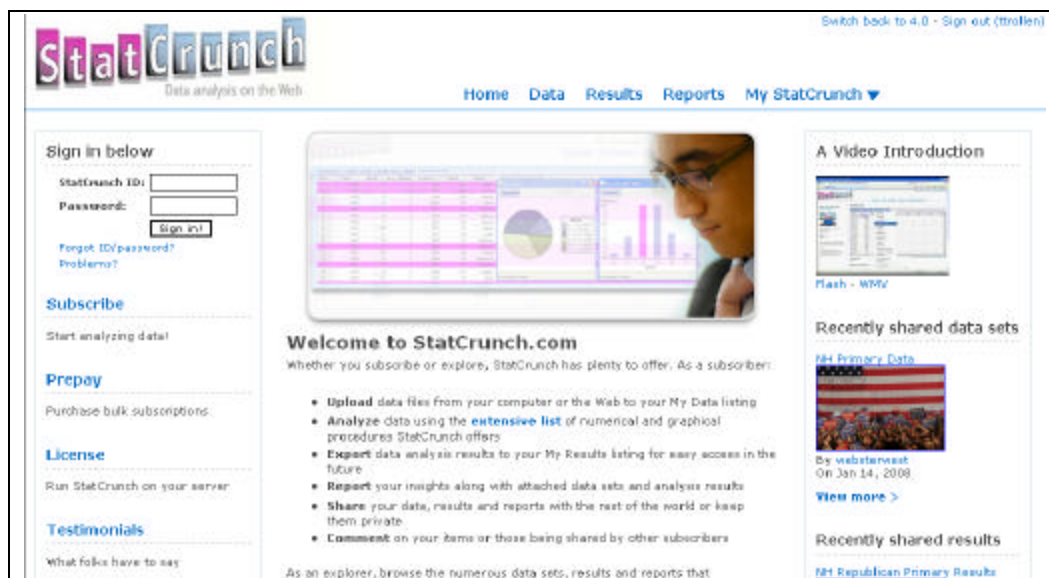


Figure 1

Figure 2 shows what the screen looks like after you've opened *StatCrunch*. To the right of the *StatCrunch* logo is the *Main Menu*, with links (such as *Home* and *MyStatCrunch*) that you can use to navigate the on the *StatCrunch* site.

The lower right portion of the window contains the *Data Table*, which has room for several columns of data.

Immediately above the *Data Table* is the *Data Table Menu bar* with menus named *StatCrunch*, *Data*, *Stat*, *Graphics*, and *Help*. You use these menus to issue *commands* to *StatCrunch*. For example, in Figure 3 the *Graphics* menu has been opened, revealing several types of graphs *StatCrunch* can produce.

As you work through the *Guided Tours* in this booklet, I'll need a way of indicating which menu you should open and which command you should choose. For example, if I want you to open the *Graphics* menu and then click the *Pie Chart* command, I'll put in the following step:

- 6) Click *Graphics*
> *Pie Chart*.

Notice this step is indented and is preceded by a number. I used the > symbol to indicate the sequence of first opening the *Graphics* menu and then selecting the *Pie Chart* command.

Patience, Patience!

Depending on your system's configuration and the speed of your Internet connection, you may need to wait for a command to complete. Before clicking a menu command or icon a second time, first look for signs that an action is already pending, such as the animated progress indicator in the upper right portion of the

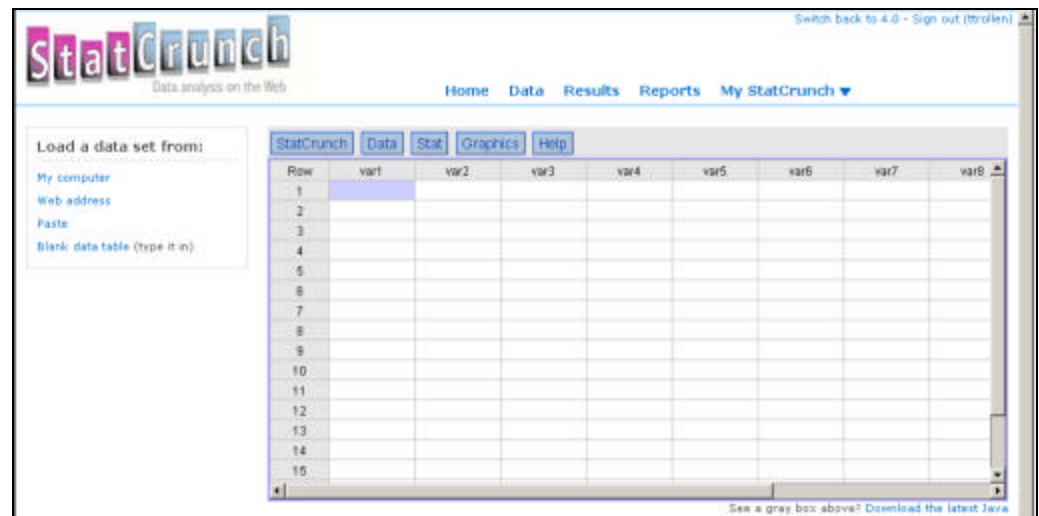


Figure 2

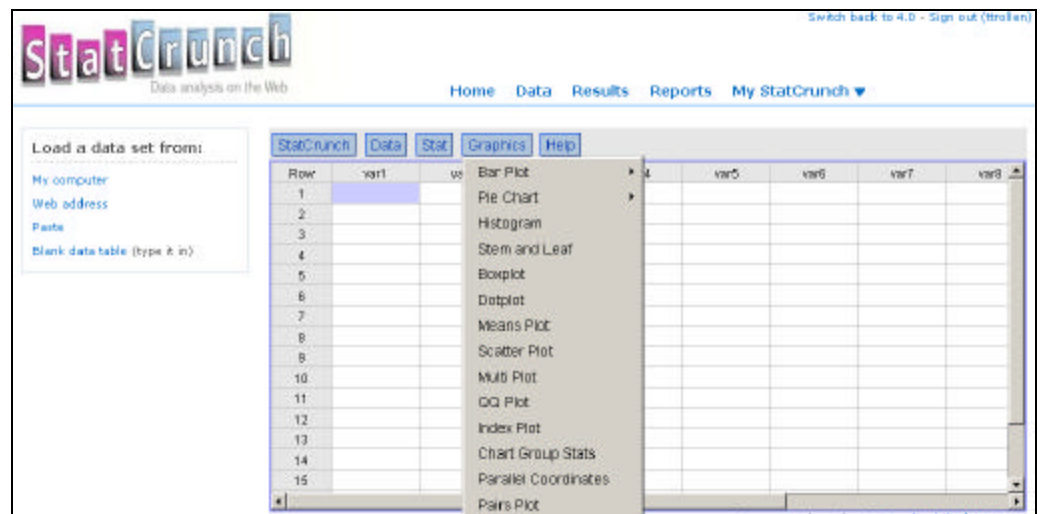


Figure 3

browser window, the *progress meter* in the Status Bar at the bottom of the browser window, or an hourglass pointer.

Exam Score Scenario

We will use the following (fictitious) data set to illustrate several *StatCrunch* commands in Guided Tour #1 below. Suppose a professor randomly select

a sample of eight students and ask each how many hours they spent preparing for a particular exam. The professor's records also provide each student's recent exam score, gender (using *codes* of 0=male and 1=female), GPA and absences.

Figure 4 illustrates the data as they've been entered into the *Data Table*. For example, the first row represents a student named Adams who

studied 20 hours, scored 64 on the exam, is male with a GPA of 3.1, and who had two (2) absences.

Notice that each *row* holds the data for a different student and that each *column* represents a different variable. Each column should be given a descriptive name up to 12 characters in length. The columns in Figure 4 bear the names *Student*, *Hours*, *Score*, *Gender*, *GPA* and *Absences*. Generally speaking, only *numeric* data may participate in calculations and graphs, which is why we used 0/1 codes for gender. What do we enter when a particular data value is missing or unknown? You may be tempted to simply leave the cell empty, but later you wouldn't know whether the cell was empty because a legitimate value had been deleted or because the cell never had a value entered. I recommend you type an asterisk (*) to indicate a missing or unknown value.

Row	Student	Hours	Score	Gender	GPA	Absences	var7	var8
1	Adams	20	64	0	3.1	2		
2	Baker	16	61	0	2.9	1		
3	Clinton	34	84	1	3.3	1		
4	Dole	23	70	0	3	2		
5	Edwards	27	88	1	3.2	1		
6	Fox	32	92	1	3	0		
7	Gore	18	72	0	3.1	0		
8	Hale	22	77	0	3.1	3		
9								
10								
11								
12								
13								
14								
15								

Figure 4

Guided Tour #1

Estimated time: 45 minutes

Objectives: Retrieve an existing data worksheet; manually enter data; use menu commands to analyze and graph data; import *StatCrunch* results into a word processor document.

This *Guided Tour* can be done on any PC with an Internet connection. If you prefer, you may work in one of SCC's computer labs. As you work with *StatCrunch*, you'll produce results and graphs that you'll want to incorporate into a printed report. Before you start working with *StatCrunch* you'll launch your word processing program (e.g.: *Word* or *Works*) so you can copy and paste *StatCrunch* results into a document.

- 1) Begin by booting your system.
- 2) Choose *Start > Programs >* and launch your word processor.
- 3) Type your first and last name as the first line of the new document. Then press [Enter] twice.
- 4) Next, launch a web browser. Your *Taskbar* will now have two buttons, one for your word processor, and another for *StatCrunch*. You can navigate between these windows by clicking their respective *Taskbar* buttons. Warning: it can become confusing when you have more than a few windows open!

StatCrunch is a Java program that is delivered to your browser over the Internet. Before launching *StatCrunch* you must check your browser's **security settings** to allow the browser to create child windows (popups). Your computer may freeze or *StatCrunch* may not work correctly if you don't address these security settings. The next step shows how to set an appropriate security level for *Internet Explorer*.

- 5) Click *Tools >* click *Internet Options >* click the *Security* tab *>* click the *Internet globe >* drag the *Security* slider control to *Medium* (if the *Security* slider control is disabled, click the *Default Level* button) *>* click *OK*.
- 6) Go to the *StatCrunch* homepage by typing <http://www.StatCrunch.com> and pressing [Enter]. Your screen should resemble Figure 1.
- 7) Click the **Subscribe** link and complete the subscription registration process.


Now, you'll work with the data for the *Exam Score* data set. To save you from manually entering the columns of data, you'll load the data from a text file stored on the SCC web server.

- 8) Click *Open StatCrunch*. Your screen should resemble Figure 3.
- 9) Click the *Data* menu *>* *Load data >* *From file*.
- 10) Click *WWW*.
- 11) Type www.sc.maricopa.edu/trollen/gbs221/exams.txt and accept the default setting for the other controls. Scroll down and click the *Load File* button. Your screen should resemble Figure 4.

Suppose this professor assigns letter grades (A, B, C, etc.) based on a 90%, 80%, 70% scheme. So student Adams, who scored 64, would be assigned a grade of D. In the next few steps, you'll enter a new column of data for these letter grades.

- 12) Click the column heading cell that currently says *var7*, then press [Delete] and type Grade. Press [Enter].
- 13) Enter the following values in the column (from top to bottom, one letter in each cell): D, D, B, C, B, A, C, C.

Since you've changed the data, you'll save it for future use.

- 14) Click the *Data* menu > *Save data*. A  *Save Data* dialog appears.
- 15) Accept the default values and click *Save*. A message appears, confirming your data was saved. You should be returned to the *StatCrunch Data Table*.

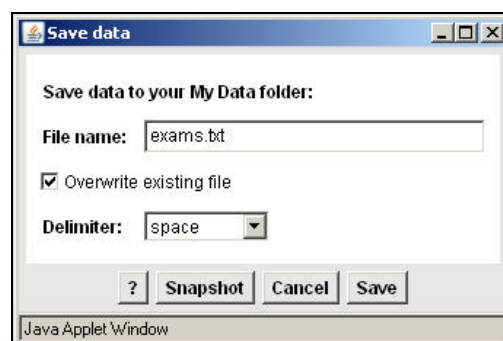










Figure 5

Tabular Methods

We will now use *StatCrunch* to work with two tabular tools presented in Chapter 2 of our text: frequency distributions and crosstabulations. A **frequency distribution** is constructed by grouping together observations with the same value, then reporting the frequency (and/or relative frequency) of each value. You'll produce a frequency distribution of the *Gender* values.

- 16) Click *Stat* > *Tables* > *Frequency*. A  *Frequency Table* dialog appears. Notice its *Taskbar* button has the Java logo .
- 17) Click *Gender* in the list of variables, and then click *Calculate*. The frequency table results are displayed in the  *Frequency Table* dialog. Can you interpret the results? Recall that a code of 0 was used for Male students, so this sample had 5 males and 3 females. The output also indicates that Males represented 62.5% of the students in this group. Of course, the remaining 37.5% were female (*Gender* code of 1).
- 18) In the upper part of the window, click the *Options* menu, then click *Edit*. This allows you to change command options, and then re-do the analysis.
- 19) If necessary, scroll down in the variable list, and then click *Grade*. Both variables are now listed in the right side list box. Click the *Calculate* button. When the results appear, what was the relative frequency (proportion) of students who got an A on the exam?

Next, you'll copy and paste this frequency distribution into your word processing document. [*Note: some students have reported problems with the Export-based copy-and-paste method outlined in the next few steps. If you have problems, try the alternate method discussed in the Sidebar in Figure 7.*]

- 20) In the upper part of the  *Frequency Table* dialog, click the *Options* menu, and then click *Export to My Results (Save/Copy/Print)*. An  *Export Results to your My Results folder* dialog box appears.
- 21) Read the message in the  *Export Results* dialog, and then click the *Export* button. After a few moments, you'll see a message indicating your results were saved in your *My Results* area. Click the *Okay* button to close the message box.
- 22) You've probably returned to your  *Frequency Table* results window. Leave this results window open (for now).
- 23) Return to *StatCrunch* by clicking its *Taskbar* button. Notice the *My results for this data* pane in the left portion of the *StatCrunch* window. It should now show *Frequency table*, formatted as a blue hyperlink.
- 24) Click the *Frequency table* link. Your frequency table results will appear in a **new window**. Note that the new window's *Taskbar* button says *StatCrunch* and does **not** have the Java icon .

- 25) Click the *Copy* link. A *Copy Instructions* warning appears. Read the screen carefully, including the instructions on the left.



Figure 6


- 26) In the Security Warning dialog box, click the checkbox to the left of *Always trust content from this publisher*, then click *Run*. This creates a security certificate to allow *StatCrunch* to use resources on your PC and copies your frequency table results to the *Windows Clipboard*.
- 27) Click the *Taskbar* button for your word processor to return to your document.


- 28) Click a line below your name, then click *Edit > Paste*. Your frequency table results

should now appear in your document. Press the [Enter] key a few times to make room for additional pasted results later. [*Note: If you weren't able to paste the StatCrunch results, try the alternate method listed in the Sidebar.*]

- 29) **Do not continue until your document contains the pasted Frequency Table results.**

- 30) Save your document under the name *StatCrunch Tour #1*.

- 31) Return to the  *StatCrunch Frequency Table* results. (If necessary, click its *Taskbar* button.)

- 32) Close the window by clicking its *Close*  button.

- 33) Close the  *Frequency Table* results window by clicking its *Close*  button.

Sidebar: An Alternate Method



1. Click the *Taskbar* button for the  *Frequency Table* results window and resize (if necessary) so its contents are visible.
2. Hold the [Alt] key and press the [PrintScr] key. This places a snapshot of the *current window* on the Windows Clipboard.
3. Click the *Taskbar* button for your word processing document and use *Edit > Paste*.

Figure 7

A ***crosstabulation*** cross-classifies data based on two qualitative variables and can reveal relationships between the variables. Crosstabulations are also known as ***contingency tables***. Let's have *StatCrunch* create a crosstab for *Gender* and *Grade*.

- 34) Return to *StatCrunch* by clicking its *Taskbar* button. Your data should still be present in the *Data Table*.

- 35) Click *Stat > Tables > Contingency > With data*. A  *Contingency Table with data* dialog box appears.





- 36) Click the *Row variable* list box and select *Gender*. Click the *Column variable* list box and select *Grade*, then click *Calculate*. Your results appear.

- 37) Can you interpret this crosstabulation? (Ignore the two rows at the bottom.) The rows represent *Gender* (0=Male, 1=Female) and the columns are for *Grade*. Do you detect any *Gender-Grade* pattern? It appears that male students earned C's and D's, while the female students earned A's and B's! We'll revisit this apparent *Gender-Grade* pattern later.


- 38) Click *Options*, and then *Edit*. The previous settings are visible. Click *Next*. You can now add *relative frequencies* to the contingency table. To break the analysis down by *Gender* (the rows), click the check box to the left of *Row percent*. A check mark now appears within the box. Click *Calculate*.

- 39) Each cell in the contingency table now has two values. The upper value is the cell *frequency*, the lower value is the row-based *relative frequency* you just requested. Now you can see that 60% of the *males* got C's and the remaining 40% got D's. Among the *females*, 33.33% got A's and the remaining 67.67% got B's. In the Total row at the bottom of the table, you see the *overall* grade distribution once again (12.5% got A's, 25% got B's, etc.).

Copy this frequency distribution into your word processing document.

- 40) In the  *Contingency Table with data* results window, click *Options*, and then click *Export*. An  *Export Results to your My Results folder* dialog appears.
- 41) Click the *Export* button. Once you see a message that your results were saved, click the *Okay* button to close the message box.
- 42) You've probably returned to your  *Contingency Table* results. Leave this dialog open in case you need to use the *Sidebar* method of pasting your results into your document.
- 43) Return to *StatCrunch* by clicking its *Taskbar* button. Notice the *My results for this data* pane in the left portion of the *StatCrunch* window. It should now show *Contingency Table with data*.
- 44) Click the *Contingency Table with data* link. Your contingency table results will appear in a **new window**. Note that the new window's *Taskbar* button says *StatCrunch* and does **not** have the Java icon .
- 45) Click *Copy*. As before, this copies your results to the *Windows Clipboard*. (Note: you shouldn't see the Copy instructions security warning again because the security certificate is still in effect.)
- 46) Click the *Taskbar* button for your word processor to return to your document.
- 47) Click one line below your previous results, then click *Edit > Paste*. Your contingency table should appear in your document. Press [Enter] a few times to make room below for additional pasted results later. [**Note:** *If you weren't able to paste the Contingency Table results, use the alternate method listed in the Sidebar above.*]
- 48) ***Do not continue until your document contains the pasted Contingency Table results.***



In all future steps when pasting results into your report document, use your choice of either the ***Export*** method or the ***Sidebar ([PrintScrnl]) method***. Your report must contain all requested results.

- 49) Save your modified document.
- 50) Close the *StatCrunch-Contingency Table with data* window (probably the right-most button on the *Taskbar*), and also close the  *Contingency Table* dialog.




- 51) Return to *StatCrunch*. Check that the exams data is still present in the *Data Table*. (If not reopen the file as you did in step 9) and re-enter the letter grades.)

Graphical Methods

Next, you'll have *StatCrunch* create several types of graphs. A **pie chart** is constructed by grouping together observations with the same value, then scaling the slices of the pie based on the *relative frequency* for each class. We'll create a pie chart for *Gender*.


- 52) Click *Graphics > Pie Chart > with data*. A  *Pie Chart* dialog box appears.
- 53) Click *Gender* in the list of variables, then click *Next*.
- 54) Click the check box to the left of *Count* to *remove* the check mark. Each slice of your pie chart will be labeled with its *relative frequency*.
- 55) Click *Create Graph!* to accept the defaults and create the pie chart. Your pie chart appears. What percent of these students were female?
- 56) Close the  *Pie Chart* dialog and return to *StatCrunch*. (Don't include the pie chart in your report.)

A **histogram** groups together observations with similar values, then reports the frequency (or relative frequency) of observations within each class. Let's have *StatCrunch* create a histogram for hours studied.



- 57) Click *Graphics > Histogram*. A  *Histogram* dialog box appears.
- 58) Click *Hours*, and then click *Next*.
- 59) If necessary, change the *Type* to *Frequency*. Accept the default values for *Start Bins at* and *Binwidth*. We'll accept the remaining default settings so click *Create Graph!* The  *Histogram* dialog box now displays your histogram. What is plotted on the horizontal axis? What is plotted on the vertical axis? The chart shows a frequency of 2 for the third column, indicating that two students studied *at least 20 but fewer than 22.5* hours. Confirm that this is correct by clicking the *Taskbar* button for *StatCrunch* and scanning the *Hour* column in the *Data Table*. Do you see two students (Adams and Hale) who studied *at least 20 but fewer than 22.5* hours?
- 60) Return to the histogram by clicking the *Taskbar's*  *Histogram* button.

This histogram plots the *frequency* of each class on the vertical axis. If you wish, you can plot *relative frequencies instead*. You can also specify both the starting value and class width.



- 61) Click *Options* then click *Edit*. Change the *Type* to *Relative Frequency*. Change *Start Bins at* to 15. Change *Binwidth* to 5. Click *Create Graph!*
- 62) Look at your modified histogram. The first column starts at 15 hours and uses columns that are each 5-units wide. Approximately what relative frequency (proportion) of students studied at least 20 but fewer than 25 hours?

- 63) Paste your *relative* frequency histogram into your report document using your choice of method. Your histogram must appear in your document. Press [Enter] a few times to make room for additional pasted results later.
- 64) Save your modified document.
- 65) Close the  *Histogram* window, and return to *StatCrunch*.


What was the distribution of *Scores* for these students? We could look in the *Data Table* to see each student's individual score. Instead, we'll construct a ***stem-and-leaf display***, which combines the graphical nature of a histogram with the detail of a data array (sorted column).

- 66) Click *Graphics > Stem and Leaf Plot*. A  *Stem-and-Leaf* dialog appears.
- 67) Click *Score*.
- 68) Click *Create Graph!* The stem-and-leaf plot appears.
- 69) Inspect your stem-and-leaf plot. Each observation was split into a *stem* portion and a *leaf* portion. For example, the lowest score, 61, was split into a stem of 6 and a leaf of 1. Thus the split occurred between the one's and 10's digits. Can you combine each leaf with its stem to re-create the original values (61, 64, 70, 72, 79, 84, 88, 92)?
- 70) Close the  *Stem-and-Leaf* window. (Don't paste it into your report.)


A ***scatter plot*** simultaneously plots two variables and can reveal interactions between them. Let's use a *scatter plot* to investigate whether there is a relation between *Score* and *Hours* studied for these eight students.

- 71) Choose *Graphics > Scatter Plot*. A  *Scatter Plot* dialog appears.
- 72) Select *Hours* as the X variable and *Score* as the Y variable, then click *Create Graph!* What does this plot indicate about a *relationship* between exam score and hours studied?
- 73) Close the  *Scatter Plot* dialog. (Don't include it in your report.)


Earlier, when we produced a contingency table, we detected a consistent *Gender-Grade* pattern. Let's see if a *Gender vs. exam Score* scatter plot reveals the same pattern. Even though *Gender* is nominal-level (categorical) data, our use of ***numeric codes*** (0=male, 1=female) enables us to make a scatter plot involving *Gender*.

- 74) Choose *Graphics > Scatter Plot*.
- 75) Select *Gender* as the X variable and *Score* as the Y variable, then *Create Graph!* Does there appear to be a relation between gender and exam score?
- 76) Close the  *Scatter Plot* dialog. (Don't include it in the report document.)


Before we conclude that this professor favors certain students based on *Gender*, let's also prepare a scatter plot of *Gender vs. Hours*.

- 77) Back in *StatCrunch*, choose *Graphics > Scatter Plot*.
- 78) Select *Gender* as the X variable and *Hours* as the Y variable, then click *Create Graph!* Does there appear to be a pattern between Hours studied and Gender? It seems female students (*Gender* code of 1) tended to study more than the male students, which may explain why female exam scores were higher than those of the males!
- 79) Paste your Gender vs Hours scatter plot into your report document using your choice of method. Press [Enter] a few times to make room for additional pasted results later.
- 80) Save your modified document.
- 81) Close the  *Scatter Plot* dialog, and return to *StatCrunch*.

Although not presented in our textbook, a **dotplot** graphically portrays each observation using a dot. You now produce a dotplot of the *GPA*'s.


- 82) Click *Graphics > Dotplot*. The  *Dotplot* dialog re-appears.
- 83) Select *GPA* as the variable, and then click *Create Graph!* Your *dotplot* appears. This dotplot is easy to interpret since there are only eight values in this data set. Each student's *GPA* is represented by a dot. Students with the same *GPA* are stacked on top of one another. You'll use dotplots in conjunction with another data set in *Guided Tour #2*.
- 84) *StatCrunch* can plot multiple variables on the same chart. Click *Options* then click *Edit*. In the variables list, *GPA* is still selected. Now click *Absences*. Two variables, *GPA* and *Absences*, are now selected.
- 85) Click *Create Graph!* The graph contains a dotplot for each variable. The horizontal axis applies to both dotplots. You can see that *Absences* varied between 0 and 3 and that *GPA*'s were centered near 3. You could determine *frequency* by counting the stacked dots.

One advantage of **dotplots** over histograms is that no decisions need to be made regarding the number of classes, the class limits, or the interval widths. Another advantage is that you see where each data value actually lies. Recall that a histogram merely plots a bar whose *width* represents several data points falling *somewhere* within its span and whose *height* indicates the frequency (count) of values appearing within the span.


- 86) Close the  *Dotplot* dialog, and then return to *StatCrunch*. (Don't include it in the report.)

Numerical Summaries

Now that we've seen several graphs, we'll have *StatCrunch* calculate many of the numerical *summary measures* from Chapter 3. We can get these summary measures with a command on the *Stat* menu.

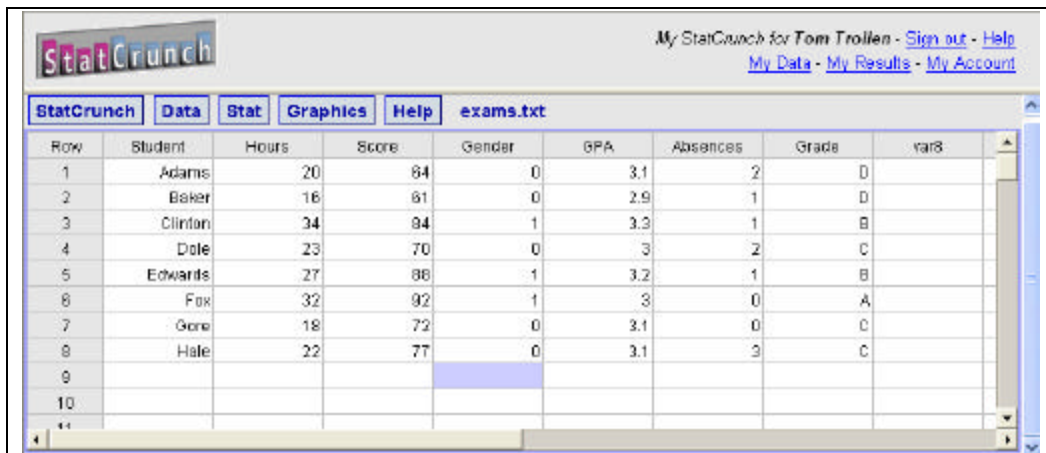
- 87) Click *Stat* > *Summary Stats* > *Columns*. A  *Column Statistics* dialog appears.
- 88) Click *Score*, and then click *Next*. You're presented with a list of several numerical summary measures and each is pre-selected (appears in the right list box). One of these, *Std. Err.*, isn't covered until Chapter 7 of our text so we may as well deselect it here. In the left list box, click *Std. Err.*
- 89) Click *Calculate*. The results appear. What was the highest exam score? The lowest score? The mean score? The other results will make more sense as we go through Chapter 3.

You just had *StatCrunch* calculate summary measures for the exam scores. You can easily have it compute summary measures for multiple variables simultaneously, and will do so next.

- 90) Click *Options*, then click *Edit*.
- 91) Click *Back* to return to the variables list. Click *Hours*, *Gender*, *GPA*, and *Absences*. Click *Calculate*. Do you now see a numerical summary of *each* of the five variables?
- 92) Notice the mean for variable *Gender* (.375). An interesting thing happens when one uses *0/1 codes* for *categorical* data, as we did for *Gender*. Earlier, when we made a pie chart of *Gender*, we saw that 37.5% of the students in the sample were female. When 0/1 codes are used to represent categorical data, the mean (.375) corresponds to the relative frequency (proportion) for the value coded as 1.
- 93) Close the  *Column Statistics* window. (Don't include it in your report.)

Comparative Analysis



At this point we've learned several *StatCrunch* features. Along the way, a crosstabulation (contingency table) and a scatter plot each portrayed an apparent pattern that females tended to score higher than males. *StatCrunch* has a feature that makes it easy to compare different *subgroups*, provided the data set includes a column which contains values that indicate subgroup membership. For example, look at the *Gender* column in the *Data Table*. These values indicate which students are male (0) and which are female (1). *StatCrunch* can use these codes to split the students into two subgroups: males and




Row	Student	Hours	Score	Gender	GPA	Absences	Grade	var8
1	Adams	20	84	0	3.1	2	D	
2	Baker	16	61	0	2.9	1	D	
3	Clinton	34	84	1	3.3	1	B	
4	Dale	23	70	0	3	2	C	
5	Edwards	27	88	1	3.2	1	B	
6	Fox	32	92	1	3	0	A	
7	Gene	18	72	0	3.1	0	C	
8	Hale	22	77	0	3.1	3	C	
9								
10								

Figure 8

females. You'll take advantage of this feature by making a new, comparative dotplot of *Score*.

- 94) Click *Graph > Dotplot*. The  *Dotplot* dialog appears.
- 95) In the *Select columns* list box, click *Score*.
- 96) Click the *Group by* arrow, then click *Gender*. This tells *StatCrunch* to use the 0/1 codes in the *Gender* column to split the students into two subgroups, and then make a score dotplot for each subgroup.
- 97) Click *Create Graph!* Your dotplot should appear. Since you asked *StatCrunch* to perform a *subgroup analysis* based on the gender codes, it places both dotplots on a common X axis, which makes it easy to compare the male scores (lower section) with the female scores (upper section). The set of female scores is obviously shifted to the right compared with the set of male scores. It appears the females did much better!
- 98) Paste your comparative dotplot into your report document using your choice of method. Press [Enter] a few times to make room for additional pasted results later.
- 99) Save your modified document.
- 100) Close the  *Dotplot* dialog, and return to *StatCrunch*.

StatCrunch's Group by option is also available in several other commands. Let's use it to obtain separate numerical summaries for males and females.

- 101) Click *Stat > Summary Stats > Columns*.
- 102) In the *Select columns* list box, click *Score*.
- 103) Click the *Group by* arrow, and then click *Gender*.
- 104) Click *Calculate*. *StatCrunch* uses the gender codes to split the students into two subgroups, and then calculates separate results for each gender. The comparative dotplots above indicated that female students tended to have higher scores. Do these *numerical summaries* confirm that? Notice that the female values for mean, median, min, max, Q1, and Q3 are each about 20 points higher than the corresponding value for male students.
- 105) Include the comparative numerical summary measures in your report. Paste them into your report using your choice of method. Press [Enter] a few times to make room for additional pasted results later.
- 106) Save your modified document.
- 107) Close the  *Column Statistics* dialog, and return to *StatCrunch*.

If this were an actual data analysis project, you'd go into your document and type answers/interpretations in the appropriate places between the pasted results and graphs. Since this is a *Guided Tour*, it will be sufficient to merely print your pasted results, without interpretations. The steps below conclude *Guided Tour #1*.

- 108) Close the *StatCrunch* window.
- 109) If necessary, click the *Taskbar* button for your word processor.
- 110) Print the document and turn it in when requested.
- 111) Close your word processor.

Guided Tour #2

Estimated time: 30 minutes

Objectives: Retrieve an existing data worksheet; produce graphs and summary measures for comparative subgroup analysis; paste *StatCrunch* results into a word processor and provide an interpretation of the results.

As in *Guided Tour #1*, the actions you take appear in indented, numbered paragraphs; however, in this guided tour you'll include a few short interpretations into your report document where indicated.

- 1) Choose *Start > Programs >* and launch your word processor.
- 2) Type your first and last name in as the first line of a new document. Then press the [Enter] key twice.
- 3) Save the new document as *StatCrunch Tour #2*.
- 4) Launch a web browser and, if necessary, adjust its security settings.
- 5) Go to the *StatCrunch* homepage (<http://www.StatCrunch.com>).
- 6) Login to *StatCrunch*.
- 7) Click the *Data* menu > *Load data > From file*.
- 8) Click *WWW*.
- 9) Enter www.sc.maricopa.edu/trollen/gbs221/reading.txt and accept the default setting for the other controls. Scroll down and click *Load File*. The *Data Table* should resemble Figure 9.

These columns represent scores from a reading test administered to two different groups of third-graders. The researcher provided the following description:



These are the results of an experiment to test whether directed reading activities in the classroom help elementary school students improve aspects of their reading ability. A *treatment* class of 21 third-grade students participated in these activities for eight weeks, and a *control* class of 23 third-graders followed the same curriculum without the activities. After the eight-week period, students in both classes took a Degree of Reading Power (DRP) test which measures

StatCrunch				
Data				
Stat				
Graphics				
Help				
Row	Score	Group	var3	var4
1	26	Control		
2	56	Treatment		
3	54	Treatment		
4	57	Treatment		
5	60	Control		
6	58	Treatment		
7	37	Control		
8	42	Control		
9	37	Control		

Figure 9


the aspects of reading ability that the treatment is designed to improve.

Notice the reading scores appear in the first column. The second column indicates whether the student participated in extra reading activities (Treatment) or not (Control). This is similar to the exam score data set where male and female students were mixed together and where a column of 0/1 codes indicated each student's gender.


- 10) Click *Graphics > Dotplot*. The  *Dotplot* dialog appears.
- 11) In the *Select columns* list box, click *Score*.
- 12) In the *Group by* list box, click *Group*, then click *Create Graph!* The  *Dotplot* dialog now contains two dotplots on the same set of axes.
- 13) Paste this comparative dotplot into your report document using your choice of either the Export method or the Sidebar ([PrintScrn]) method.
- 14) If a pasted graph seems too large, you may *resize* it. To do so, click anywhere on the graph. You should now see a border around the graph that also contains small black squares (sizing handles) on each corner and at the middle of each side. Drag the lower-left hand corner upward and to the right to make the dotplot shorter and narrower.
- 15) Take a few moments to compare the dotplots. Since they share a common X axis, it will be easy to compare these two sets of reading scores. The *upper* dotplot shows the distribution of reading scores for students who participated in extra reading activities (i.e., the *Treatment* group). The *lower* dotplot shows students who did not participate in the extra reading activities (i.e., the *Control* group).


Note that, as a group, the upper distribution is shifted somewhat to the *right* compared with the lower distribution. **What does this right shift suggest about the effectiveness of the extra reading activities?** In your report document, below the pasted dotplots, type in your answer to the bold question above. A sentence or two should be sufficient.

Also, note that the values in the lower distribution are *more scattered* than the values in the upper distribution. **What does this difference in scatter suggest about the effectiveness of the extra reading activities?** In your report document, type in your answer to the bold question above. Once again, a sentence or two should be sufficient.


- 16) Save your modified document, which now includes your answers to step 15).
- 17) Close the  *Dotplot* dialog.
- 18) Return to *StatCrunch*.

Now, have *StatCrunch* compute a separate numerical summary for each subgroup.

- 19) Click *Stat > Summary Stats > Columns*. A  *Column Statistics* dialog appears.
- 20) In the *Select columns* list box, click *Score*.

- 21) In the *Group by* list box, click *Group*, then click *Calculate!* The  *Column Statistics* dialog now shows two rows of numerical summary results.
- 22) Paste this comparative numerical summary into your report document using your choice of either the *Export* method or the *Sidebar ([PrintScrn])* method.
- 23) Take a few moments to study the comparative numerical summaries. Specifically, note that the Treatment group's mean, median, minimum, maximum, Q1 and Q3 are each 10-20 points higher than the corresponding value for the Control group. **What does this suggest about the effectiveness of the extra reading activities?** Type your brief answer in your report document, below the pasted numerical summaries.
- 24) Save your modified document, which now includes your answers to step 23).

Perform the remaining steps to wrap up *Guided Tour #2*.

- 25) Close the  *Column Statistics* window.
- 26) Save your modified document.
- 27) Print the document.
- 28) Close your word processor.
- 29) Close the *StatCrunch* window.

Review Questions

[**Hint:** use the index to locate topics you don't remember.]

- a) What is a *histogram*? How do you tell *StatCrunch* to make one?
- b) What is a *stem-and-leaf plot*? How do you tell *StatCrunch* to make one? What do you have after you combine each leaf with its stem?
- c) What is a *dotplot*? How do you tell *StatCrunch* to make one?
- d) Which numerical summary measures can *StatCrunch* calculate? What menu command generates them?
- e) Rather than leaving a cell blank, what can be entered to indicate the value is missing or unknown?
- f) What does the *mean* of a categorical variable represent when 0/1 codes are used?
- g) Briefly describe *comparative subgroup analysis*. Which *StatCrunch* option do you use to perform one?

Turn In

- 1) Printouts of the documents produced in *Guided Tour #1 and #2*.
- 2) Your answers to the Review Questions.

Index

<i>0</i>	
0/1 Codes	12
C	
Categorical data	12
Codes	
0/1	3, 10, 12
missing value (*)	3
Comparative Analysis	12
Contingency Table	
overview	7
relative frequencies	7
Crosstabulation.....	7
D	
Data Table	
overview	2
screen shot	3, 14
Dotplot	
compared with histogram.....	11
overview	11
E	
Exam Score Scenario	
overview	3
Exam Score vs. Hours Studied Scenario	
data	4
<i>Export Results</i>	5
F	
Frequency	
histogram.....	9
Frequency Distribution.....	5
G	
Graphs	
dotplot.....	11
histogram.....	9
pie chart	5, 8
resizing	15
stem-and-leaf display.....	9
Group By option	12, 13
H	
Histogram	
frequency.....	9
overview.....	9
relative frequency	9
M	
Menu bar.....	2
Missing value code (*).....	3
N	
Nominal-level data	10, 12
Numerical Summaries.....	11
P	
<i>Paste Results</i>	
<i>Export Method</i>	5
Sidebar [PrintScreen] Method	6
Pie chart	8
PrintScreen Method.....	6
Proportion	12
R	
Relative frequency	
histogram.....	9
pie chart.....	8
Resizing Graphs.....	15
S	
Scatter plot	
overview.....	10
Security Settings	4
Sidebar.....	6
<i>StatCrunch</i>	
home page.....	1
menus	2
overview.....	1
Stem-and-leaf display	9
Subgroup Analysis	12
Summary Measures	11
T	
Tables	
contingency table	7
crosstabulation	7
frequency distribution	5
U	
Unknown value code (*).....	3