BASICS

1) To open MINITAB, you just need to click on the appropriate icon.
2) Once opened, you will have two environments in the software. The top part is called the "session window", the lower part is a spreadsheet (the grid) with columns (C1, C2, ...,) and rows (1, 2, ...). This part is called the "Worksheet". This is the place in which we enter the data.
3) Essentially that’s all we need to know to start with MINITAB!

DATA ENTRY

1) Place your cursor on the cell 1-C1 which is cell in the first row and the first column of the worksheet. Notice that the frame of the working cell is highlighted. Type the first number in your data-set and hit "Enter" key in your keyboard. This will take you to cell 2-C1 or the second row in the first column. In MINITAB, each column represents a set of data (as oppose to each row). When the data is large you might want to enter it row-wise (however, each column is still a different data-set.) To do this, click on the upper-left-corner square in your worksheet. The vertical arrow becomes horizontal and the data entry can be done row-by-row.

2) To edit a cell simply click on it. If you want to add a new data-point in the middle of a column, highlight the cell positioned after the desired row and right click on your mouse. By clicking on the "insert rows" option, you can add a new data point. You can delete a cell in the same fashion except you use "delete cells" command in the last stage. Alternatively, all the editing options can be performed using "edit" from the menu bar. A fine way of getting comfortable with the program is to create data and try editing it.

3) The best way to keep track of data-sets (or cells) is to name them. To name each column, you can click on the empty cell within the same column but above the first row. Later, when you apply various statistical methods, the software will recognize that column with the name you chose for it.

4) To obtain MINITAB data files from other sources, choose FILE from the menu bar, then click on "OpenWorkSheet ...". Finally, notice that you can switch between the two sessions by choosing "Window" option from the menu bar. At the very bottom of the "Window" menu, you can view the open worksheets and the window session.

5) We will discuss the importance and applications of window session later, after we introduce the statistical techniques. Briefly, the window session in the place at which the results of the statistical analysis will appear.
Creating Graphs

Pie Charts

Type the categories in the first and the frequencies in the second column of your spreadsheet. From the "Graph" menu, choose "Pie Chart ...". Click on Chart table. Place the corresponding columns in Categories in: and Frequencies in: respectively. Write an appropriate title for your graph. Finally, click on "OK".

Dot plots

Input the data to a column in your spreadsheet. From the "Graph" menu, choose "Dotplot ..."). Place the column of interest in the variables entry. Hit "OK".

Histograms

From the "Graph" menu, choose "Histogram ..."). The names of the columns that you have already created will appear on the box located at the left side of the "Histogram" application. Simply click on the column of interest and hit "select". The selected column will appear under the "X" mark. By hitting "OK", you will create your histogram. There are other options in this menu that we will learn about as we make progress into the quarter.

Simple Calculations

We can apply various simple operations on a given column. For example, let’s assume that we want to multiply each component of the first column (C1) by 6 and place the results in the fifth column (C5). Choose "Calc" from the menu bar and choose "Calculator". In the empty box in front of "Store result variable", type: C5. Click on the empty area under "Expression". There, type 6*C1. C1 can either be manually typed or can be selected using "Select" option. Hit "OK". You should have the result in the fifth-column.

Also, in "Calculator" and under the "Functions" menu, there is a whole list of functions you can use. For example, if you wanted to get the square root of column C1 and place it in C5, all you need to do is to type "Square root (C1)" in the expression box. But "Square root" is a provided function that is, you only need to double click on it from the "Functions" menu.
Problem 1. (The Marital Status Data) Let’s create a pie-chart for the marital status data we presented previously in the class. Remember that the data consist of 4 categories each representing the frequency associated with the marital status of female Americans 18+.

<table>
<thead>
<tr>
<th>Marital Status</th>
<th>Count (millions)</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never Married</td>
<td>43.9</td>
<td>22.9</td>
</tr>
<tr>
<td>Married</td>
<td>116.7</td>
<td>60.9</td>
</tr>
<tr>
<td>Widowed</td>
<td>13.4</td>
<td>7.0</td>
</tr>
<tr>
<td>Divorced</td>
<td>17.6</td>
<td>9.2</td>
</tr>
</tbody>
</table>

(a) Enter the name of the categories in column **C1** of your MINITAB software.

(b) Name this column ”categories”.

(c) Enter the frequency of the categories in column **C2** of your MINITAB software.

(d) Name this column ”counts”.

(e) Obtain a pie chart for the frequencies.

(f) Obtain a bar graph for the frequencies. Here is the procedure: From the ”Graph” menu, choose ”Chart...”. In that menu, put the ”frequencies” under **Y** and put the categories under **X** and hit ”OK”.

(g) Obtain a bar graph for percentages. First, let us obtain the relative-frequency of each category. To do this, we need to obtain the sum of all four categories. Go to **Calc** and choose **calculator .....** In the box **Store result in variable**, type **C3**. This tells MINITAB that you want the result of summation to appear in the third column. Next, by browsing the **Functions**: select the function called **Sum**. Then select the frequency column and place it between the parentheses of the function so that under the **Expression** box you will have: **Sum('counts')**. Then hit ”OK”. The sum (191.6) will appear in the first row at C3. Now, let’s get the relative-frequencies. Again, go to **calculator ....** This time store the results in **C4**. Inside the **Expression**: box, type: **C2/C3**. This tells MINITAB that you want the fourth column to be the relative frequency for each category. Hit ”OK”. Now if you want to explain the relative frequencies in terms of percentages, you should multiply the numbers in C4 by 100. This can be easily done using **calculator ....** Repeat the same procedure as for the previous section to get the bar graph expressed in percentages.
**Problem 2.** Below is a table of poverty population of ages 25 or older of the United States, listed according to the education level, from the census of 2001. The numbers are in thousands.

<table>
<thead>
<tr>
<th>Education Level</th>
<th>Total Population</th>
<th>Number Below Poverty Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>No High School Diploma</td>
<td>28,948</td>
<td>6,447</td>
</tr>
<tr>
<td>High School Diploma, No College</td>
<td>58,456</td>
<td>5,628</td>
</tr>
<tr>
<td>Some College</td>
<td>46,042</td>
<td>3,025</td>
</tr>
<tr>
<td>Bachelor’s Degree or More</td>
<td>48,696</td>
<td>1,625</td>
</tr>
</tbody>
</table>

Use MINITAB to find out the following information:

(a) Enter the ”total population column” to ”C1”. Call that column pop.

(b) Enter the ”below poverty level” in ”C2”. Call that column pov.

(c) Calculate the population of the United States of age 25 or above and the total poverty population of age 25 or above.

(d) Calculate the percentage of population below poverty level in each of the education levels. (create a new column ”C3” which and call it ”percent”.)

(e) Is ”pov” a quantitative or a qualitative random variable?

(f) Draw a pie chart of poverty population where each slice of the pie represents one education level. Can a pie chart illustrate the relationship between poverty population and education level? (Interpret your finding).

To draw the pie chart correctly, you need to follow these steps:

**Step 1:** Create a new column ”C4” and call it ”category”.

**Step 2:** Input numbers 1, 2, 3 and 4 in these categories. Alternatively, you can put four letters that represent the four educational levels. For example, NH (for no high school), HNC (high school no college), SC (some college) , and BD (bachelor’s degree or more).

**Step 3:** In MINITAB, choose ”Graph” option. Next, choose ”Pie Chart ...”. Click on ”chart table”. In the box called Categories in, input the ”category” column. In the box called Frequencies in, input the ”pov” column you just created.

**Step 4** Click ”OK”.

extra Create a pie chart for column C3 or ”percent”. The result looks interesting. Don’t you think? Can you explain.
Here are the number of home runs that Babe Ruth hit in each of his 15 years with New York Yankees, 1920 to 1934. 54, 59, 35, 41, 46, 25, 47, 60, 54, 46, 49, 46, 41, 34, 22

(a) Make a stem and leaf plot for these data. To do this, after typing the number of home runs in a column, go to Graph and choose Stem-and-Leaf option. Select the column of interest and place it in the variables: box. Hit ”OK”. Interpret the output.

(b) Let’s put the numbers in stems of 10 instead of stems of 5. This time while repeating the same procedure as above, you need to type 10 inside the box Increment:. Click ”OK”. Interpret the output.

Problem 4. GPA, Math and Verbal scores
From the File menu, click on Open Worksheet ..., select the file called Ga.MTW, and hit ”OK”. You should have a new worksheet open with 100 rows and three columns named ”Verbal”, ”Math”, and ”GPA”. These columns reflect the sampled verbal and math scores of 100 students along with their GPA. Alternatively, this file is also available on the class website under the name Ga.MTW at: http://www.csub.edu/~sbbehseta/Labs140spring.htm

Answer the following questions:

(a) Are these three random variables continuous or discrete. Why?

(b) Let’s assume ”you are interested in studying the effect of Math-score on overall GPA of those students. Then which variable is considered dependent (response) and which is considered independent? why?

(c) Draw a histogram for the ”Verbal” scores.

(d) Draw a histogram for the ”Math” scores.

(e) Can you compare the performance of the 100 students in terms of their ”math” and ”verbal” scores simply by looking at the histograms of the two variables?

(f) WE WILL REVISIT THIS PROBLEM LATER AS WE MAKE PROGRESS IN THE QUARTER.
Problem 5. Open Netscape or Internet Explorer, choose your favorite news website. Choose three pieces of news article that illustrated the use of statistics. Briefly (in not more than 5 sentences) quote and state how statistics is used in each of the articles you picked. Here are some of the suggestions of web sites:

<table>
<thead>
<tr>
<th>News Organization</th>
<th>Site Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bakersfield Californian</td>
<td><a href="http://www.bakersfield.com">http://www.bakersfield.com</a></td>
</tr>
<tr>
<td>Los Angeles Times</td>
<td><a href="http://www.latimes.com">http://www.latimes.com</a></td>
</tr>
<tr>
<td>CNN</td>
<td><a href="http://www.cnn.com">http://www.cnn.com</a></td>
</tr>
<tr>
<td>Reuters</td>
<td><a href="http://www.reuters.com">http://www.reuters.com</a></td>
</tr>
<tr>
<td>San Francisco Chronicle</td>
<td><a href="http://www.sfgate.com">http://www.sfgate.com</a></td>
</tr>
<tr>
<td>USA Today</td>
<td><a href="http://www.usatoday.com">http://www.usatoday.com</a></td>
</tr>
<tr>
<td>Washington Post</td>
<td><a href="http://www.washingtonpost.com">http://www.washingtonpost.com</a></td>
</tr>
<tr>
<td>Sports Illustrated</td>
<td><a href="http://www.si.com">http://www.si.com</a></td>
</tr>
<tr>
<td>Science Journal</td>
<td><a href="http://www.science.com">http://www.science.com</a></td>
</tr>
</tbody>
</table>

EXIT MINITAB!