JMP® Introductory Lab Activities

Activity 3: Describing Numerical Data

Data Set: Big Class.jmp

Summary

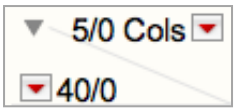
In this lab, you will graph and summarize continuous data in JMP using **Analyze** > **Distribution**.

You'll explore the data and create a report with JMP results, your commentary and your interpretation (required output and discussion is in italics).

Opening the Big Class Data

Open the file **Big Class.jmp** from the JMP Sample Data directory (go to **Help > Sample Data**, then click **See an Alphabetical List of all Sample Data Files**).

The data represents 40 students who are enrolled in a martial arts class. Their teacher has to buy new uniforms for an upcoming competition and has taken measurements to help her. For each student, four variables are shown: age, sex, height and weight.



Note: You may need to click on the gray triangle in the upper left-hand corner of the data grid to open the Data Table Panel. This panel provides additional information on the data table and easy access to menu options.

Graphing and Summarizing Continuous Data

We'll start by exploring the **height** data using **Analyze > Distribution**. Select **height** for the **Y, Columns** variable and click **OK**. The output window will display a histogram, a boxplot and some numerical summaries.

To change the histogram to a horizontal layout, click the red triangle next to **height**, and select **Display Options > Horizontal Layout**.

Using the **height** red triangle menu again:

- Create a **Stem and Leaf** plot for the heights.
- Use **Histogram Options** to modify the histogram to include both a **Count Axis** and probability **(Prob) Axis.**

Interacting with Your Data

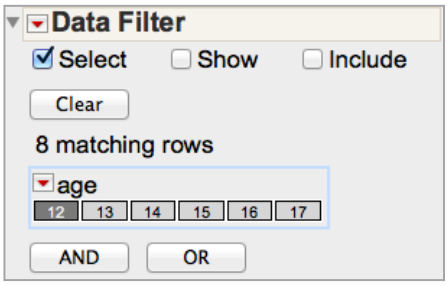
One important feature of JMP is its interactive nature. Here are some examples for you to try. (Note: In general, you can undo a mistake you make in the data table by using **Edit > Undo**):

- Arrange the windows on your screen so that both the **Big Class** data table and **Distributions** windows are visible. (You may need to resize or tile the windows.)
- Click on a row of the data table, then look at the histogram: Its location is highlighted.
- Look at the stem and leaf plot: This same value highlighted in the histogram is shown in bold type.
- Select all the 12-year-olds from the data table. There are many ways to do this. Explore these two methods:

  (1) From the **Rows** menu, select **Row Selection > Select Where**. Select the variable **age**. Make sure **equals** is chosen in the box below the variable selection, and enter **12** in the box to the right of this. Click **OK**.

  (2) From the **Rows** menu, select **Data Filter**. Select the variable age, and click **Add**. In the resulting window, click on the **12** age group to select.

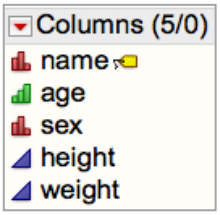
  (Hint: To deselect rows or columns, click anywhere in the data table, or click **Clear** in the **Data Filter**.)



  Where do the 12-year-olds fall in the histogram and the stem-and-leaf plot for **height**?

- Click on a bar of the histogram. The corresponding rows of the data table are selected and stem and leaf plot values are bolded.
- The boxplot shows two potential outliers (individual points). Move your cursor over one of the points. The name of this student should appear as the pointer moves over the dot.

  The variable **name** has been assigned the role of label. This is indicated in the columns panel (to the left of the data table) by the yellow tag icon.



-

- Add the label role to the variable **age**. (Hint: Right-click on the **age** column, then select **Label/Unlabel**). The yellow tag icon will appear next to the newly labeled variable in the columns panel.

- If you have both **name** and **age** as labels, determine what happens when the pointer moves over one of the outliers. Click on the outlier point. Note that the row in the table is highlighted, as is its location in the histogram and stem and leaf plot.

Create a Report

Create a new report including the following:

- *A histogram for height that includes a count axis.*
- *A boxplot for height that appears by itself.*
- *A stem and leaf plot for height.*
- *A five-number summary (minimum, lower quartile, median, upper quartile, and maximum) for height from the **Quantiles** display. Note that JMP gives you additional quantiles beyond the five-number summary. You can edit the JMP table in your report to remove extraneous values.*
- *A table from the **Summary Statistics** text box that includes the mean, standard deviation, and number of students. Note: You can click on the red triangle next to **Summary Statistics** and use the **Customize Summary Statistics** option to remove unwanted statistics.*
- *A one-paragraph description of the heights of students from **Big Class**. This should include a description of the shape of the distribution of heights. Refer to specific aspects of the graphs and numerical displays to support your answer.*

Note: Refer to Lab Activity 1 for a refresher on selecting output using the **selection** tool (), and copying and pasting your output into another program.