

Chapter 1

An Introduction to JMP

An Introduction to the Introduction	1
Some Stylistic Conventions	3
Creating a New JMP Table	6
It's a Cheetah-Eat-Gazelle World Out There!	7
Numbers and Pictures—Univariate Data	16
Numbers and Pictures—Bivariate Data	25
What Have We Learned?	28
References	28

An Introduction to the Introduction

My goal in this chapter is to quickly get you up and running with JMP, by paying special attention to those features you will most probably wish to use and by drawing your attention to some of the important features of the JMP interface. As we work through the introduction, I encourage you to experiment with the features of JMP at your leisure. Interspersed with the discussion about JMP, I present some results from the research on teaching and learning statistics and relate those findings to the use of JMP in the

2 Teaching Elementary Statistics with JMP

classroom. My intent is to get you past the beginner stage and into the productive stage with JMP; as it happens, the ease of use of JMP will make this an easy task. To advance to the “guru” level, you should every now and then pick up the materials that come with JMP and leaf through them. They are well written, and you will undoubtedly discover new gems of JMP capability as you read. I know how little time you have in your teaching day, but perhaps an opportunity for casual JMP study will present itself at the next faculty meeting.

I am aware of many people’s aversion to reading software manuals. This is quite justifiable; manuals often seem to be written for people who (a) already know how to use the software or (b) wrote the software themselves. My fellow teachers and I are particularly adept at avoiding software manuals, a skill developed over many years of large class sizes and little time for professional reading. Only as a last resort do I reach for the software manual, and even then—I must admit—I frantically search for a quick fix. Manuals must communicate with users of all stripes, users with different levels of computer experience, a difficult task at best. Fortunately, my task here is much easier! This book is aimed at teachers of elementary statistics classes: a high-quality group with preexisting familiarity with computers.

The modern graphical user interface imposes a consistent look and feel across different software packages, and as a teacher with significant computer experience, I am sure you don’t need nor do you want a lot of basic hand-holding about your computer’s interface. That being said, there are features unique to statistical programs, and moreover features unique to JMP, that will extend your prior software experience. I will write in more detail when discussing these features.

As you know, statistical analysis is performed within the context of a problem, and the specific procedures used are defined by the nature of the statistical question and the available data. I realize that you may not have a ready stack of statistical problems complete with data sets, so I have included a few as examples. I include the problems and the accompanying data analyses to illustrate some of the capabilities of JMP, and especially its user interface. In addition, I hope that you will find the analyses interesting and usable in your classrooms.

At each step in the analyses to follow, JMP’s interface gives you a sense that you are in control of the process. You, not the software, are performing the analysis. If you have struggled with other statistical software in the past, you will better appreciate this aspect of JMP. This sense of conversation with data, a conversation facilitated by the software, is a design feature, not an accident, of JMP. A typical data analysis using JMP begins with an initial presentation of graphs and statistics. Notice I said that the analysis *begins* with an initial presentation. We seldom know in advance exactly how we will analyze data until we get an initial look at it. JMP encourages that exploration and offers help and guidance at each step. You will see how all this happens as we proceed.

Some Stylistic Conventions

My natural writing and teaching style is to engage others in a conversation, rather than lecture. I have a general aversion to using the “I” word, and the authorial “we” sounds to me too much like the “royal” we. When I see “we” in a software manual I imagine Queen Victoria, in her best teacher voice, suggesting, “We are not amused.” Generally, with first-person pronouns I will try to impart these meanings:

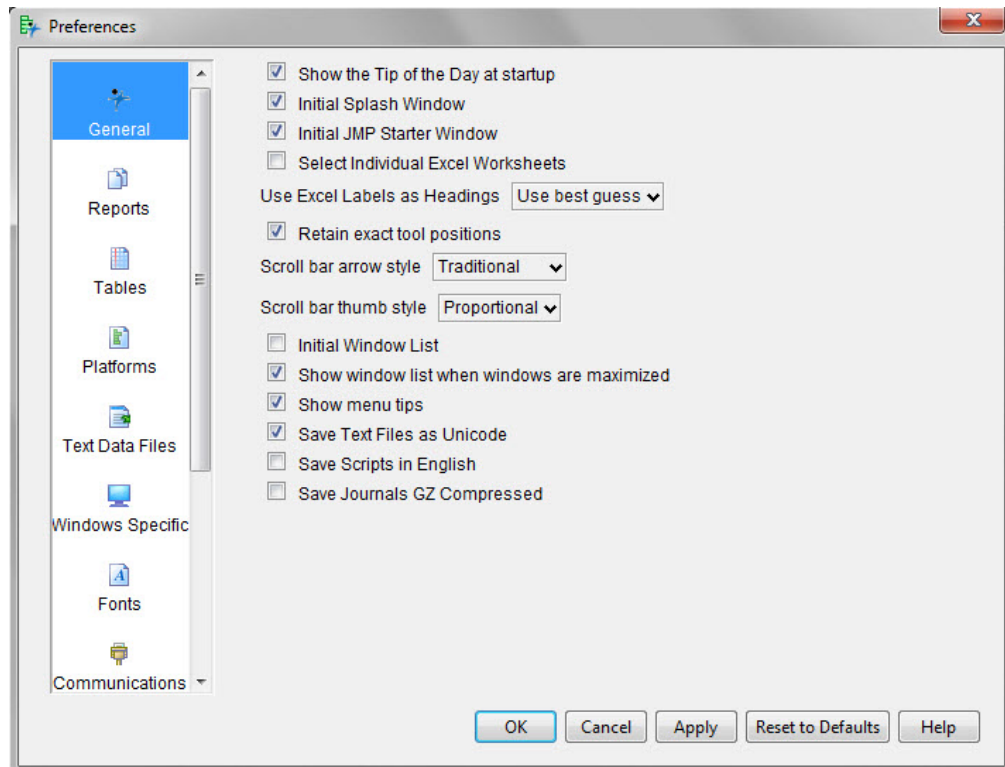
- “I” means I’m telling you what I am doing, what I’m thinking, or what I have experienced.
- “We” means that I’d like to guide you through something and am encouraging you to follow along. You are encouraged to experiment with JMP at any time, but if you do so when I am in the “we” mode your screens may differ from mine.

I will use basic terminology (click, drag, right-click, and so on) and am confident in your ability to master these directions. Frequently I will use the generic term “select,” which usually means left-click. I am writing on a PC system, and the screen shots will be from a Windows 7 operating system. Presentation of JMP in a Mac or Linux environment is very similar to what you will see here.

Some of the look, feel, and behavior of JMP can be adjusted through choices made in the Preferences menu, which is found by selecting the **File** menu and clicking **Preferences**. The **Preferences** window is shown in figure 1.1. Rather than impose my own sense of what the preferences should be, I will continue to capitalize on the judgment of the designers of JMP. (In the many years I have used JMP, the only preference I have ever changed is the font.) As you become experienced with JMP you may find yourself repeatedly changing some aspect or other of JMP. When that happens, I encourage you to check out the options embedded in the Preference capabilities.

4 Teaching Elementary Statistics with JMP

Figure 1.1 Preferences



As we progress through the chapters I will be less detailed in my directions. As an example, in later chapters I will assume you know how to open a JMP file and will skip that step in the sequence of operations.

As you proceed through this book, your attention will be moving back and forth from text to screen. The following typographical conventions will help you find your place in the text as you simultaneously attend to book, keyboard, mouse, and screen.

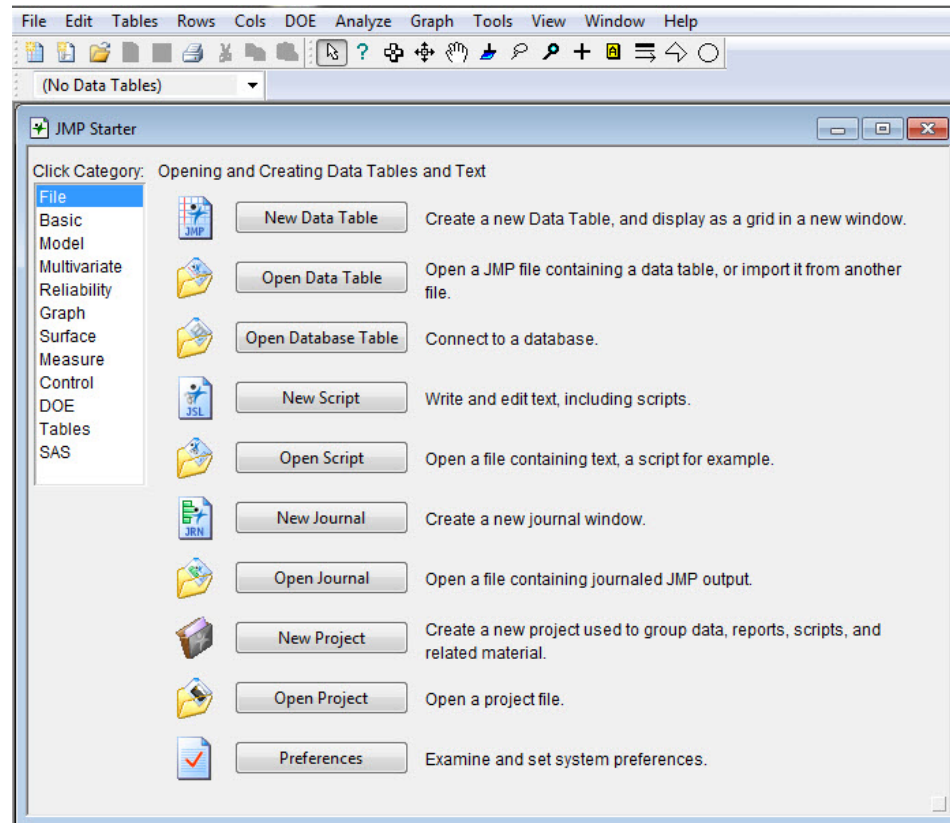
- Reference to JMP entities will appear in the **Arial bold** font to cue you to look for those words in the presentation of JMP (for example, “Select the **File** menu, then the **Save** command” will be presented in syntactical English with the JMP words or phrases in Arial bold).
- When I ask you to select a sequence of commands, the sequence will appear as phrases separated by an arrow, such as **File → Save As**.
- I will present references to variable names, data table names, and some other items in **Arial bold**, but they may appear in illustrations in either a plain or boldface font. These items may also display differently on your screen if you have set your JMP Preferences different from the defaults.
- When there is a call for action on your part, please follow along with the example by following the instructions. “Hands-on” is an effective learning technique. Following the instructions will keep you from that quick-and-dirty strategy of imagining you know what would happen without actually doing it. (Been there, done that!)

One final ground rule: Please use the JMP Help features if you sense you might be bordering on being confused. JMP has an excellent Help system, and I will not feel slighted if you use Help to get a second opinion.

OK, let’s begin. To start JMP, double-click the JMP application icon.

When JMP starts up you will see the JMP menu bar and the **JMP Starter** window (see figure 1.2). The **JMP Starter** window (the smaller one) is designed for the more experienced software user and presents functionality equivalent to the options presented above it. For our purposes we will close the **JMP Starter** window. (Closing the larger enveloping window will cause you to exit JMP.)

Figure 1.2 JMP Starter menu



Consistent with other Windows applications, the **File** menu contains all the strategic commands, such as opening and saving data tables and exiting JMP. You may also see one of the “Tips of the Day,” a set of nifty advisories about the capabilities of JMP. You can read them as they appear, run through the list of them, or even turn them off if you wish.

Creating a New JMP Table

This initial exercise will be the first and last time in this book that you will create a JMP table by entering data from scratch. After this data entry exercise you will open existing JMP data files. Statistical analysis of data is a context-driven activity, so for best results please read the context of the data as they are presented in sections and chapters to come.

It's a Cheetah-Eat-Gazelle World Out There!

For our first data analysis experience with JMP I will transport you to the grassland plains and savanna of the Serengeti in Tanzania, Africa. In the Serengeti National Park, cheetahs are predators of Thomson's gazelles. As a result, gazelles need to be attentive to the surrounding area, thus decreasing their vulnerability to attack, since one never knows what might be lurking nearby. However, gazelles feed on Serengeti grass and must therefore lower their heads to graze, thus increasing their vulnerability to attack. Cheetahs initially stalk their prey, generally moving to within 30 meters before they initiate the final chase. Upon reflection, it might seem reasonable that a fast cheetah could simply flush a herd of grazing gazelle and with a burst of speed pick off a slow gazelle for a fast lunch. Evolutionary development encourages energy efficient behavior, and one would think that a cheetah on stalk would be able to put off choosing its prey until the last moment and capitalize on a slow-moving element in a target-rich environment as the chase develops. However, once the cheetah initiates a chase it appears not to be dissuaded from an initially chosen target by the emergence of slower moving gazelles. This seems like a strategy that would burn excess energy. Is there an efficient predatory strategy that might account for this behavior?

Fitzgibbon (1989) has performed some research that may provide a possible answer. She reported on a possible relation between gazelle "vigilance" and the cheetah's choice of prey. In her study, she defined *gazelle vigilance* as the percentage of time during a stalk a gazelle had its head up scanning for predators. From a safe distance, she taped a series of stalks and kills by cheetahs and reviewed the footage to measure gazelle vigilance levels during the stalk. Fitzgibbon matched adult gazelles selected by the cheetah with the nearest actively feeding, same-gendered adult within five meters of them. She reasoned that if vigilance is a factor in prey selection, the vigilance of the selected victim should be less than the vigilance of the nearest neighbor who, after all, could have been chased just as easily. If vigilance is a factor, that could indicate the cheetah is "studying" the herd and identifying the less vigilant gazelles before settling on the specific prey.

Sixteen stalks' worth of data on the vigilance of the chosen gazelles and their nearest neighbors are presented in table 1.1. The data is of two types, quantitative (numeric) and qualitative (categorical). As you know, different statistical procedures are used with qualitative versus quantitative data. Mean- and median-based procedures are used on quantitative data, while frequencies and proportions are the usual fare for qualitative data. We will see as we proceed that JMP keeps these data types straight and will offer you appropriate statistical techniques during the analysis of data.

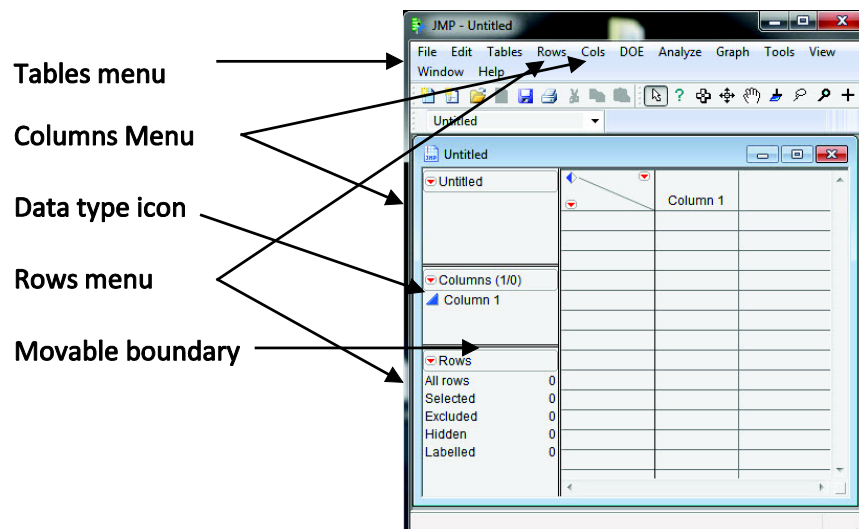
Table 1.1 Gazelle vigilance levels

Gazelle chased	Gazelle ignored	Gazelle gender
8.0	17.5	Male
31.4	63.0	Male
40.0	70.2	Male
15.7	38.8	Male
40.0	45.0	Male
35.3	39.1	Male
31.7	65.2	Male
68.5	72.5	Male
10.0	65.0	Female
78.7	84.2	Female
52.0	50.0	Female
23.9	20.5	Female
81.0	96.3	Male
49.7	90.1	Male
62.0	88.3	Male
72.6	89.9	Male
8.0	17.5	Male
31.4	63.0	Male
40.0	70.2	Male
15.7	38.8	Male
40.0	45.0	Male
35.3	39.1	Male
31.7	65.2	Male
68.5	72.5	Male
10.0	65.0	Female
78.7	84.2	Female
52.0	50.0	Female
23.9	20.5	Female
81.0	96.3	Male
49.7	90.1	Male
62.0	88.3	Male
72.6	89.9	Male
8.0	17.5	Male

1. If you have not already done so, close the **JMP Starter** window and select **File → New → Data Table**.

JMP initially displays data in a window in the form of a table with rows and columns known as a **Data Grid**. In this window, you can perform a variety of table management tasks: editing cells; creating, rearranging, or deleting rows and columns of data; and sorting. Figure 1.3 identifies some of the commonly used parts of the window, including the JMP data table. If you move the mouse over the table you will see the cursor change shape now and then, alerting you to different functions that are potentially active when the mouse is in a particular position. For example, if you position the mouse over one of the contextual pop-up menus (the red triangles) the cursor changes to a finger-pointer, indicating that you are over a clickable item. When you are at a boundary of a window the cursor will change to a vertical, horizontal, or diagonal double-pointed arrow, alerting you to the option of resizing the window.

Figure 1.3 JMP data table



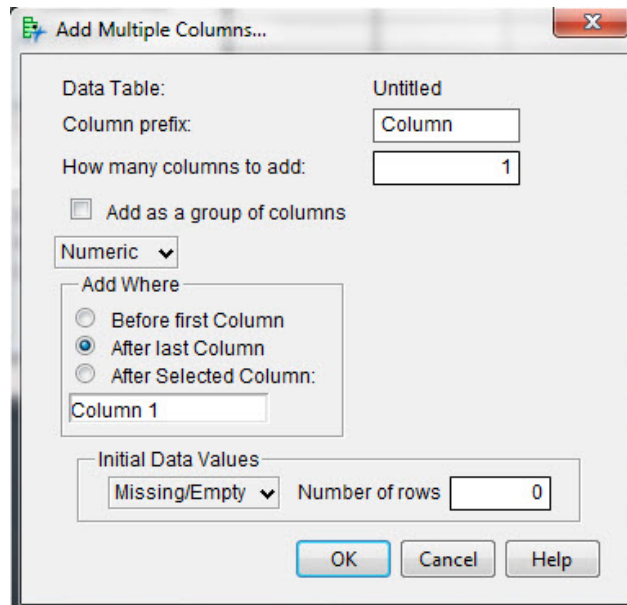
JMP data entry is similar to data entry in common spreadsheet programs. One significant difference is that whole rows and whole columns, not individual cells, are conceptual units of data entry. JMP tables consist of rows and columns, and JMP will interpret these rows and columns of data, not individual isolated storage cells for numbers. An isolated cell or group of cells can have user-defined meaning in a spreadsheet, but not in JMP. Columns represent *variables* and rows represent *observations* or data values for the variables in question.

10 Teaching Elementary Statistics with JMP

We will need to create sixteen sets of three observations on these variables, and so we will need to create a data table consisting of sixteen rows and three columns. Variables can be added one at a time, or en masse. My own personal preference is to set up the table in advance before beginning data entry. Once the table is set up I enter the data a row at a time. It is not always possible to specify the number of rows in advance, since some data entry might take place before all the data are gathered. JMP allows rows to be added later (using **Rows** → **Add Rows**). Some may prefer to enter data in a manner similar to their method with calculators: enter observations one variable at a time. Still others just start entering, and the rows are updated on the fly. The “right” way to enter data for an individual is most likely whichever way results in fewer data entry errors. One method may take more or less time, use greater or fewer keystrokes, or perhaps even be more or less psychologically uplifting; the important value is getting the data entered correctly, not how it is entered. You should be guided by your sense of what is comfortable and natural for you. To initialize the data table for our cheetah data

2. Select **Cols** → **Add Multiple Columns**. JMP will display the **Add Multiple Columns** panel, as shown in figure 1.4.

Figure 1.4 Add multiple columns



3. Respond to the **How many columns to add** question with “3”.
4. Select **Rows → Add Rows**. JMP will display the **Add Rows** panel.
5. Respond to the **How many rows to add** question with “16”.

Other options are available in the **Add Multiple Columns** panel, and you may wish to explore these. Our data consists of two quantitative variables and one qualitative variable, and we will specify the types of variables individually rather than **Add as a group of columns**.

6. Click **OK** and resize the windows so that you can see all the rows and columns.

Now that we have set up the data for table entry, let’s actually enter the data. First, we will give the variables more informative names than **Column 1**. Figure 1.5 shows what the table looks like before we begin data entry. It is considered good form to enter descriptive variable names, and I will use the following: **Chosen**, **Ignored**, and **Gender**.

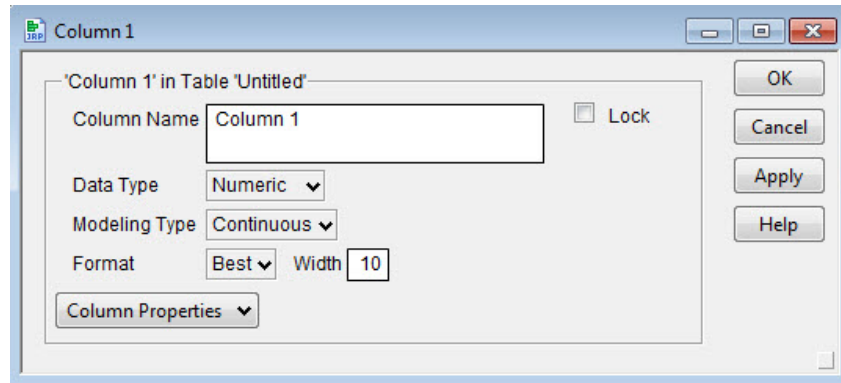
Figure 1.5 Initial columns and rows

	Column 1	Column 2	Column 3
1	.	.	.
2	.	.	.
3	.	.	.
4	.	.	.
5	.	.	.
6	.	.	.
7	.	.	.
8	.	.	.
9	.	.	.
10	.	.	.
11	.	.	.
12	.	.	.
13	.	.	.
14	.	.	.
15	.	.	.
16	.	.	.

To edit the information in column 1:

1. Double-click on the column 1 heading to display a **Column 1** window, as shown in figure 1.6.
2. Enter the variable name (**Chosen**) in the **Column Name** field.
3. Click **OK**.

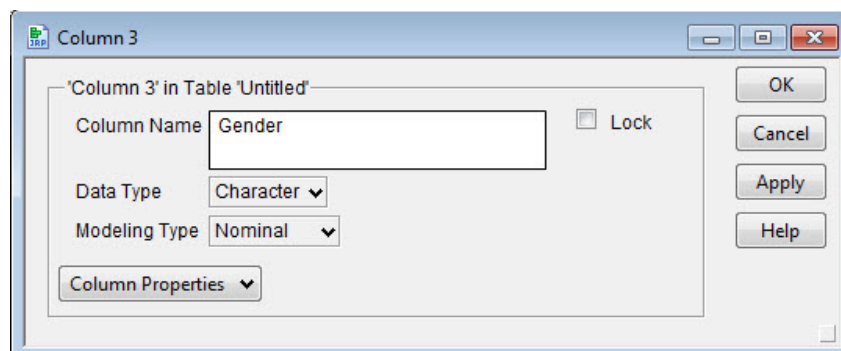
Figure 1.6 Set up a single column



Now repeat this process and enter the information for column 2, using **Ignored** as the variable name. Column 3, **Gender**, will require a little more attention because it is a categorical variable, different from the default numeric data type in JMP.

1. Double-click on the column 3 heading to display a **Column 3** window, as shown in figure 1.7.
2. Enter the variable name (**Gender**) in the **Column Name** field.
3. Click the **Data Type** menu and select **Character**. (Notice that the **Modeling Type** changes to **Nominal**. JMP now knows this is a categorical variable.)
4. Click **OK**.

Figure 1.7 Set-up for a categorical variable



Now enter the first row of data. Be aware that on a PC with a numeric keypad, the Enter keys operate differently in JMP. The usual Enter key, when pressed, will send the cursor down to the next row. However, the numeric keypad Enter key, when pressed, will move the cursor across the columns before descending to the next row. If you wish to use the numeric keypad Enter key, you can change its behavior with the following keystroke sequence:

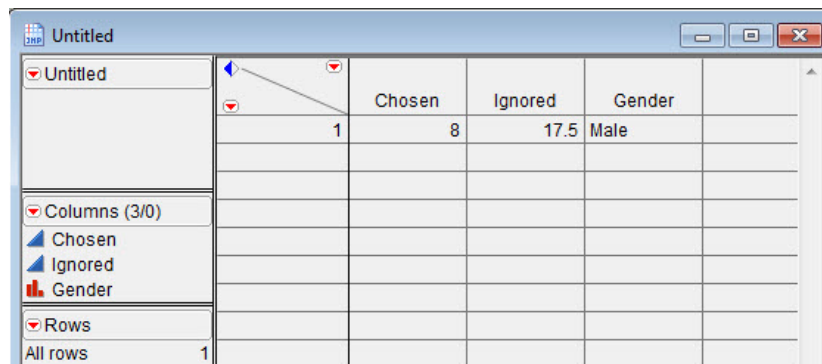
5. **File → Preferences → Tables → Numeric keypad Enter key moves down → OK.**

By whatever path and Enter key you choose, your first line of data should appear as shown in figure 1.8. Complete the data entry, check the numbers to make sure they are correct, and save the file with a name and in a location of your choice using the sequence:

6. **File → Save As.**

I will use “Gazelles” for the file name.

Figure 1.8 One line of data entered



	Chosen	Ignored	Gender
1	8	17.5	Male

The JMP table should now look like figure 1.9 and reflect our progress so far. There are three columns (complete with variable names) and sixteen rows for data. The file name is Gazelles. JMP knows the **Chosen** and **Ignored** variables are quantitative, as evidenced by the small triangles showing in the **Columns** panel. Likewise, JMP knows that **Gender** is a character/nominal variable—as evidenced by the small bar-chart icon in the **Columns** panel.

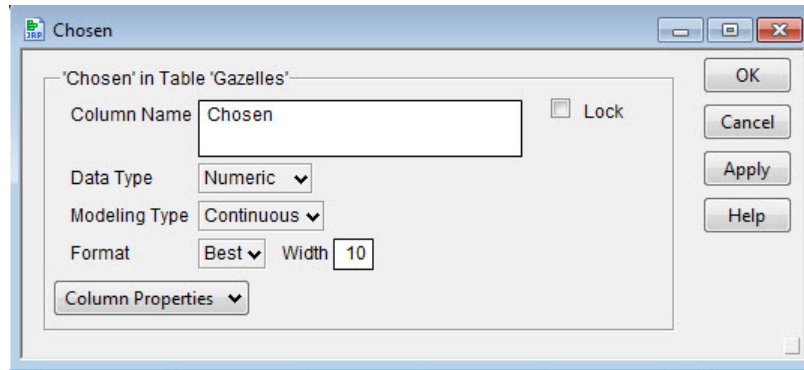
Figure 1.9 Gazelles data

	Chosen	Ignored	Gender
1	8	17.5	Male
2	31.4	63	Male
3	40	70.2	Male
4	15.7	38.8	Male
5	40	45	Male
6	35.3	39.1	Male
7	31.7	65.2	Male
8	68.5	72.5	Male
9	10	65	Female
10	78.7	84.2	Female
11	52	50	Female
12	23.9	20.5	Female
13	81	96.3	Male
14	49.7	90.1	Male
15	62	88.3	Male
16	72.6	89.9	Male

Maybe it is just me, but after entering my data I believe it is easier to read if it is in a consistent decimal form, so I will change the format so that all numbers are expressed to the tenths decimal place.

1. Double-click on the **Chosen** column heading to display a **Chosen** window, as shown in figure 1.10.
2. Select **Format** → **Fixed Dec.**
3. Select the **Dec** panel and enter “1”.
4. Click **OK**.

Figure 1.10 Chosen window



Repeat these steps for the **Ignored** column. Your completely entered data should now be displayed as shown in figure 1.11.

5. Select **File** → **Save** to make these changes permanent.

Figure 1.11 Completed data entry

	Chosen	Ignored	Gender
1	8.0	17.5	Male
2	31.4	63.0	Male
3	40.0	70.2	Male
4	15.7	38.8	Male
5	40.0	45.0	Male
6	35.3	39.1	Male
7	31.7	65.2	Male
8	68.5	72.5	Male
9	10.0	65.0	Female
10	78.7	84.2	Female
11	52.0	50.0	Female
12	23.9	20.5	Female
13	81.0	96.3	Male
14	49.7	90.1	Male
15	62.0	88.3	Male
16	72.6	89.9	Male

Numbers and Pictures—Univariate Data

Now we get to the fun part of statistics: interpreting the statistics and graphs. We carefully entered the data in the table, but psychologists (and experience!) tell us that when we check as we type, we sometimes see what we intended to enter, not what we actually entered. If this is your first experience with JMP, take extra care to verify not only that we are in the same place visually, but that we also have the same numbers and pictures.

While being in the same place is valuable, please feel free to explore different JMP paths at any point—not only will JMP gracefully allow you to experiment, but it is also designed to facilitate experimenting with different representations and summaries of your data. The JMP folks know that data analysis is not a rigid lockstep process, but one of moving x steps in one direction and y steps in a different direction, depending on the unfolding story the data is telling. The data analyst ponders now and then, and sometimes takes paths less traveled by, depending on the data. If you go off on a tangent of exploration, it is easy to return to the main path we are traversing at the time.

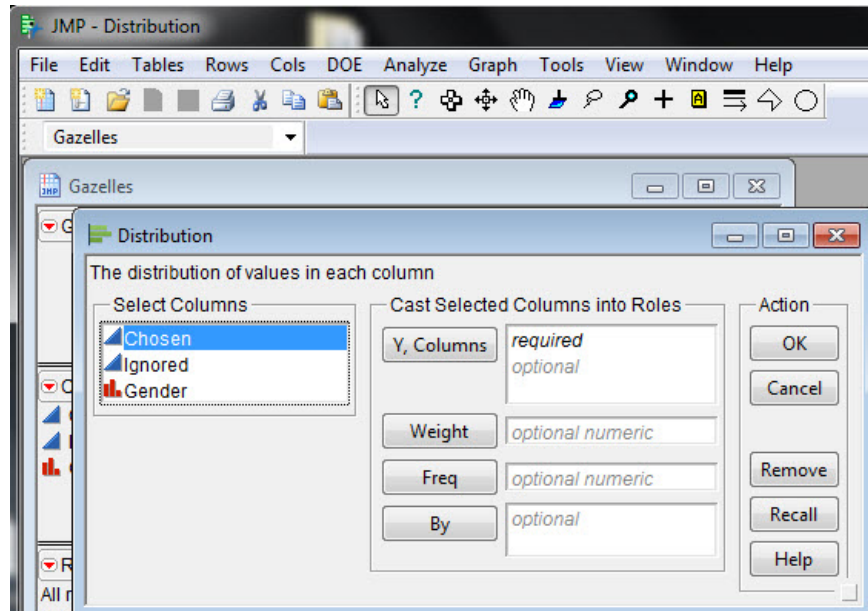
I present a more detailed, more microstepped discussion in this chapter than in discussions that follow. You will see many important aspects of the look and feel of JMP in this example, and I will tend toward pointing them out in this chapter rather than in later chapters, when you will have more experience.

OK, let's analyze!

1. Select **Analyze** → **Distribution**.
2. In the **Select Columns** panel, click the **Chosen** variable.

You should see something similar to figure 1.12. (Your screen may vary a bit from mine depending on your preference for organizing windows.) You should not fret if you do not completely understand everything in this window. JMP is presenting many options to you because it does not yet know how simple or complex you intend your analysis to be.

Figure 1.12 Choosing the variable

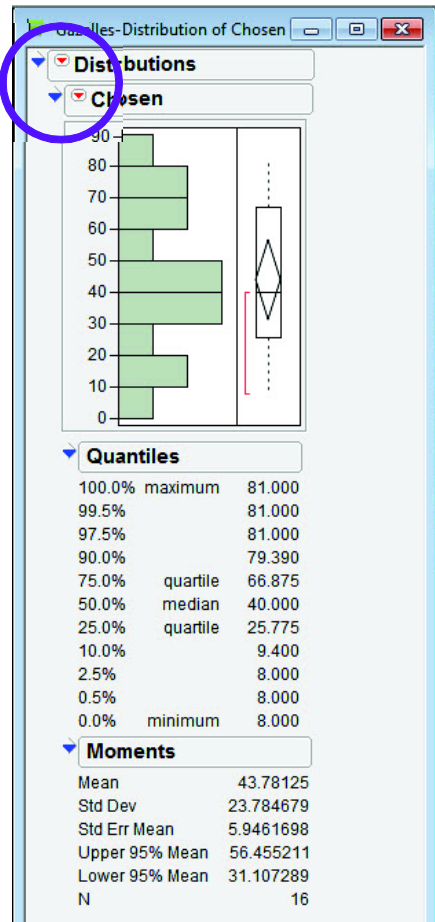


You will become very familiar with the **Select Columns** panel. It contains a list of the variables in the Gazelles file. The idea is that you will assign various “roles” to each variable. Many of the roles are for more advanced statistical analyses than you may be performing. For the most part the role selection in elementary statistical analysis will be choosing the variable(s) to analyze in a univariate setting, and choosing which variable(s) to be explanatory and response variable(s) in a bivariate setting.

- Continuing on, key the sequence, **Chosen** → **Y, Columns** → **OK**.

Do you see those little icons I have circled in figure 1.13? They should appear as half-blue diamonds and downward-pointing red triangles. These are common and useful features in the various windows of JMP. The half-blue diamonds are “disclosure diamonds.” Clicking on them opens or closes the associated item. You may choose to “disclose” or not the information associated with the icon. The red triangles, which JMP refers to as “hot-spot icons,” present context-specific options for selection. These options are ones that are meaningful where you are in your analysis, that is, the “current context.”

Figure 1.13 Distributions



4. Click the disclosure icon beside **Distributions** to see its effect on the window.

See? The entire display was concealed. Now click on the same disclosure icon to expose it.

5. Click the disclosure icon beside **Quantiles**.

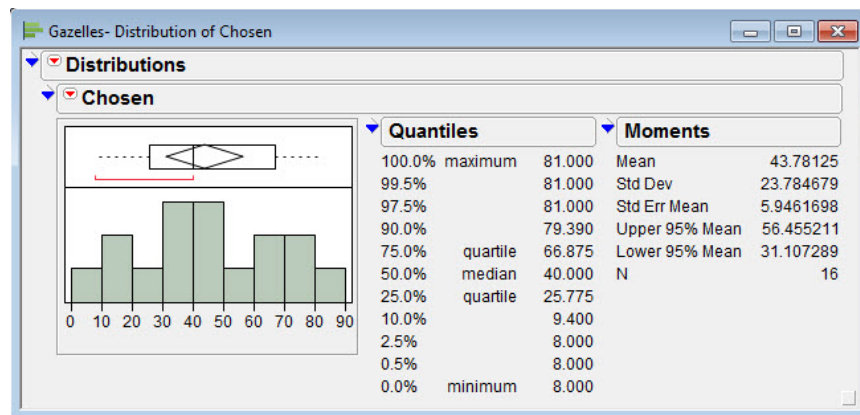
The parts of the display disappear and return as you wish—you can focus your attention on what is important to you at the time. This is also a nice feature when you want to transfer JMP output to your favorite word processor via copy and paste. JMP will transfer only the disclosed parts.

To repeat, the hot-spot icons offer different paths or a combination of paths, consistent with the context.

6. Click the **Distributions** hot-spot icon to see these further options:
 - Uniform scaling
 - Stack
 - Script

If you select the **Stack** hot-spot icon the plots will display horizontally, as shown in figure 1.14. The **Chosen** hot-spot icon provides options specific to the **Chosen** gazelles variable.

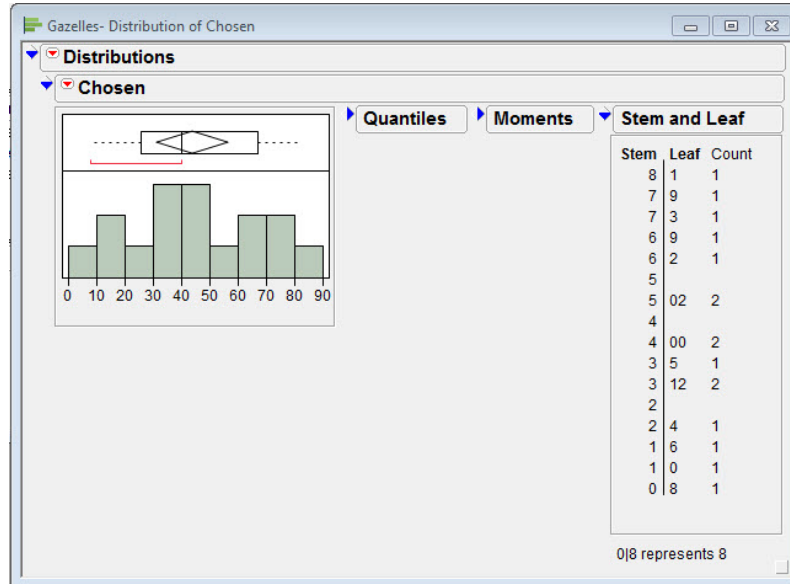
Figure 1.14 Stack chosen



7. Select **Stack**.
8. Click the **Quantiles** and **Moments** disclosure icons to close them.
9. Click the **Chosen** hot-spot icon to display the options. You will see the **Display Options**, **Histogram Options**, **Normal Quantile Plot**, and so on.
10. From the list, select **Stem and Leaf** to add the stem and leaf plot to the histogram and box plot.

You should see a display similar to that in figure 1.15.

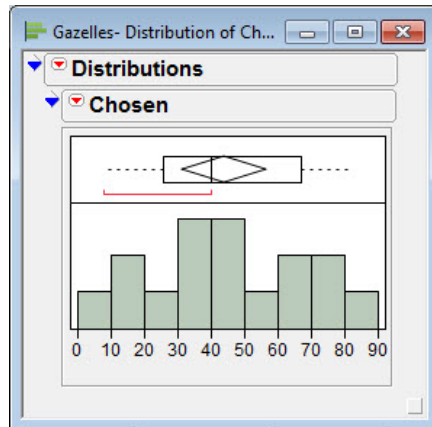
Figure 1.15 Some univariate options



Before we proceed further, we will simplify our window somewhat.

11. Click the **Chosen** hot-spot icon.
12. Repeatedly select **Display Options** and deselect these options: **Quantiles**, **Moments**, **Stem and Leaf**.
13. Change the histogram and box plot to a comfortable size by clicking and dragging on the lower right corner of the box plot or the histogram.

Figure 1.16 displays my current window.

Figure 1.16 Histogram and box plot

At this point we will experiment a bit and modify the graphs. You may have noticed that the box plot contains a diamond and a horizontal red bracket. Although these are informative, they are not box plot features we initially teach our students. The diamond presents the sample mean and a 95% confidence interval for the population mean. The horizontal red bracket indicates where the most concentrated 50% of the data are.

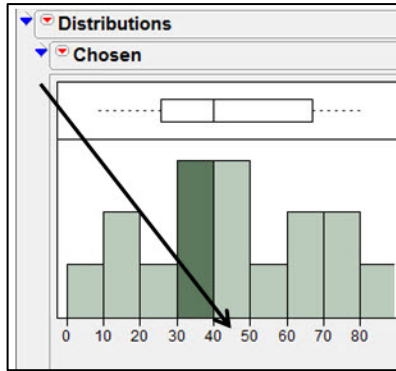
14. Right-click in the box plot area and deselect these options: **Mean Confid Diamond; Shortest Half Bracket.**

We will also change the appearance of the histogram. Move the mouse over the scale area of the histogram (as shown by the arrow in figure 1.17). Depending on where the mouse is, you will see a differently pointing right hand.

15. Left-click and drag with the mouse positioned at different positions in the scale.

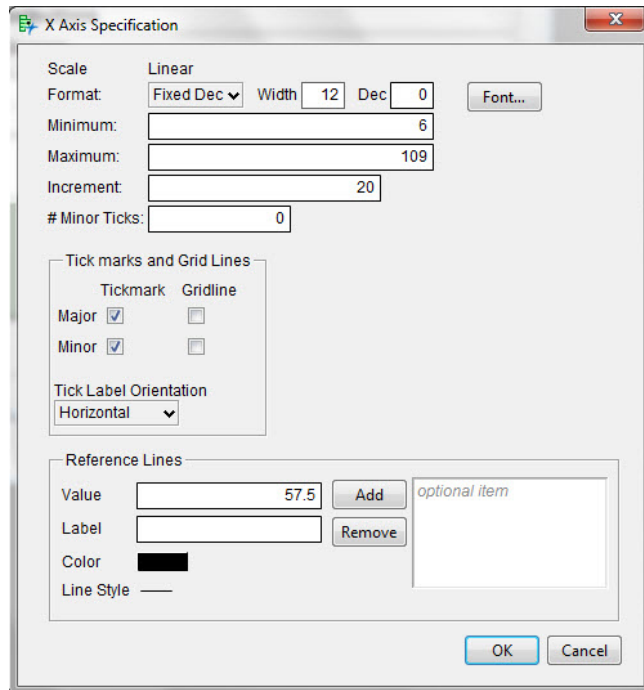
You will quickly see the effects on the scale. You may have changed the bin size while experimenting with the scale, and some of the histogram bars may have a darker color, something like that shown in figure 1.17. Not to worry! We can undo that change.

Figure 1.17 After scale manipulation



16. Double-click on the scale area (arrow in figure 1.17), and the **X Axis Specification** panel will appear as shown in figure 1.18.

Figure 1.18 X axis specification



You can adjust various features of the scale in this window. (With bivariate data, there is also a **Y Axis Specification**.)

17. Change the **Minimum** to 0 and the **Maximum** to 109; change the **Increment** to 20, and click **OK**.

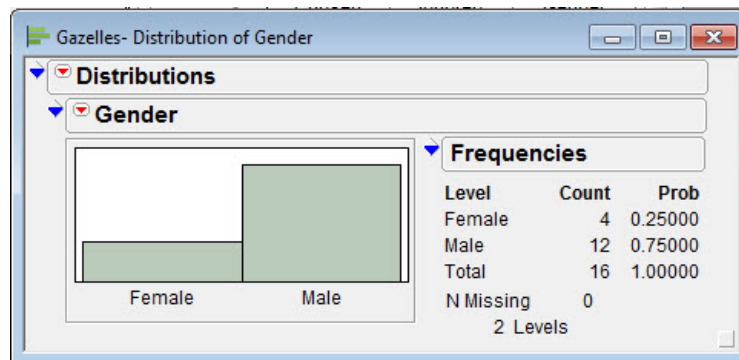
Figure 1.16 will reappear.

Thus far I have presented a short analysis of a distribution of a continuous variable. What about categorical variables? The procedures for a categorical variable unfold in a manner very similar to the quantitative variable. (This is another nice design feature of JMP: Learning is transferable.)

18. Select **Window** → **Close All Reports**.
19. Select **Analyze** → **Distribution** → **Gender** → **Y, Columns** → **OK**.

JMP presents a bar graph and a table in figure 1.19 with counts and proportions. I have used the **Distributions** hot-spot icon and selected **Stack** to change the orientation of the graph from vertical to horizontal.

Figure 1.19 Distribution of gender



How did JMP know to generate a bar chart rather than a histogram? Recall that JMP recognized the **Modeling Type** of the variable as “nominal” (categorical) when you entered the data; JMP still remembers and once again presents the correct graph for the categorical data.

There is also an alternate, and probably better, procedure to display a bar chart.

1. Select **Window** → **Close All Reports**.
2. Select **Graph** → **Chart**.
3. Select **Gender** → **Categories, X, Levels** → **OK**.

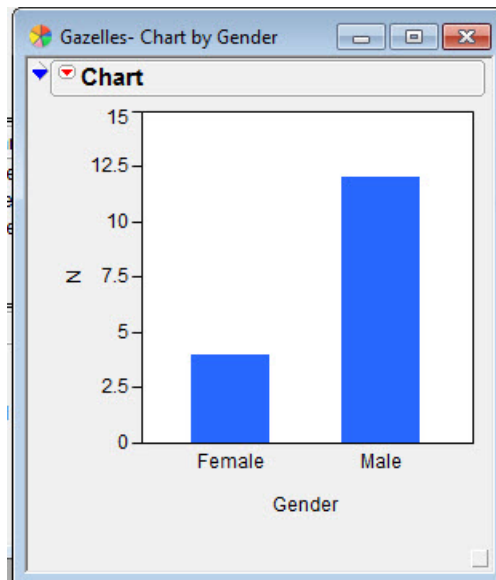
Notice the **Chart** hot-spot icon. Let us see what sort of options we have there. We can get a **Horizontal [bar] Chart** and a **Pie Chart**, among others.

4. On the **Chart** hot-spot icon, select **Horizontal Chart**.
5. Then click again, this time on **Pie Chart**.

Before moving on to bivariate data, you may wish to experiment with options presented in **Y Options**, **Level Options**, and **Label Options**.

Here I make a brief but important teaching point about this presentation of the distribution of gender. I prefer what is shown in figure 1.20 for a bar chart. Experienced statisticians may be comfortable with the plot in figure 1.19, but it is something of a problem for those who are still learning statistics. The confusion between bar charts and histograms is mentioned often in research articles about students' interpretation of graphs. Frequently students will attempt to infer something about shape from a bar graph and even try to calculate means and standard deviations if the data are coded as integers. (See, for example, Bright and Friel [1998].)

Figure 1.20 A better bar chart



delMas, Garfield, and Ooms (2005) reported that even college students confuse histograms of numeric data with bar charts of categorical data. Students in this study felt they could use bar graphs of categorical data to estimate center, shape, and spread of a distribution. The graph in figure 1.19 is sort of a hybrid; it is a graph of categorical data but has the appearance of a histogram. Good information is presented nicely in this

plot—but it can be something of a problem for students. A clear implication for teaching exists here. It is apparent that distinguishing categorical and numerical data is a topic that needs some attention. The explicit presentation of the data types in the data table in JMP supports the importance of making this distinction.

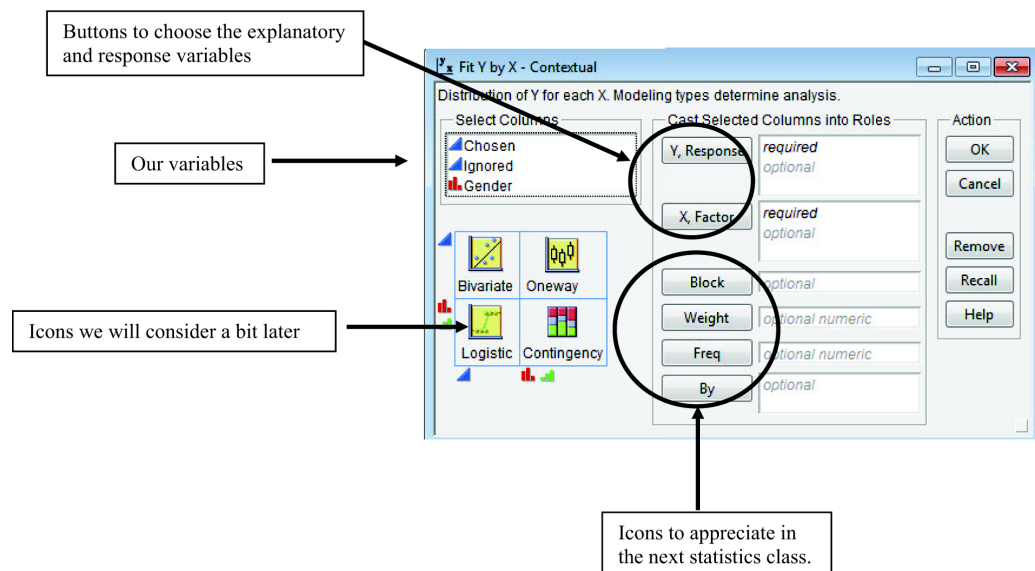
Numbers and Pictures—Bivariate Data

The premier bivariate method in elementary statistics is simple regression. Our two quantitative gazelle variables provide an opportunity to demonstrate a simple regression analysis in JMP. If you have not already done so:

1. Select **Window** → **Close All Reports**.
2. Select **Analyze** → **Fit Y by X**.

The **Fit Y by X** window will appear (figure 1.21). You will notice features of windows we have already seen as well as some new features.

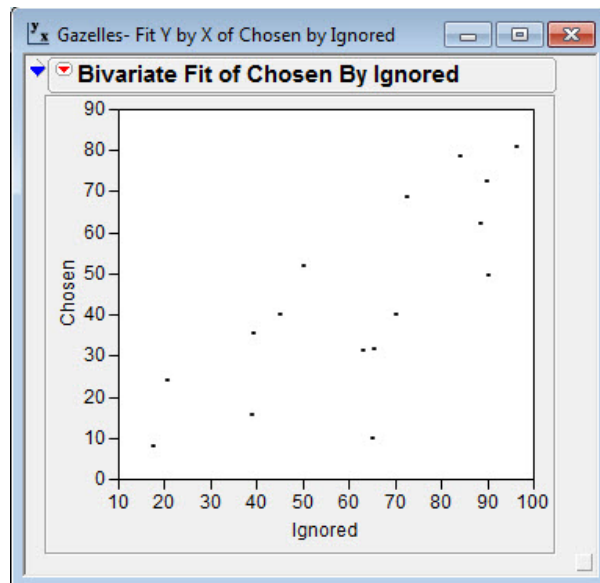
Figure 1.21 Fit Y by X



Now we must select our explanatory and response variables. We will choose the **Ignored** gazelle vigilance level as the explanatory variable, and the **Chosen** gazelle vigilance level as the response variable (see figure 1.22):

1. Select **Chosen** → **Y, Response**.
2. Select **Ignored** → **X, Factor**.
3. Select **OK**.

Figure 1.22 Chosen versus Ignored



You may wonder why JMP uses the term *Factor* instead of *explanatory* or *independent* variable. This is because *Factor* is the term equivalent to “explanatory” that is usually used in the context of analyzing experiments; JMP is written to facilitate analyzing experiments as well as the analysis of data.

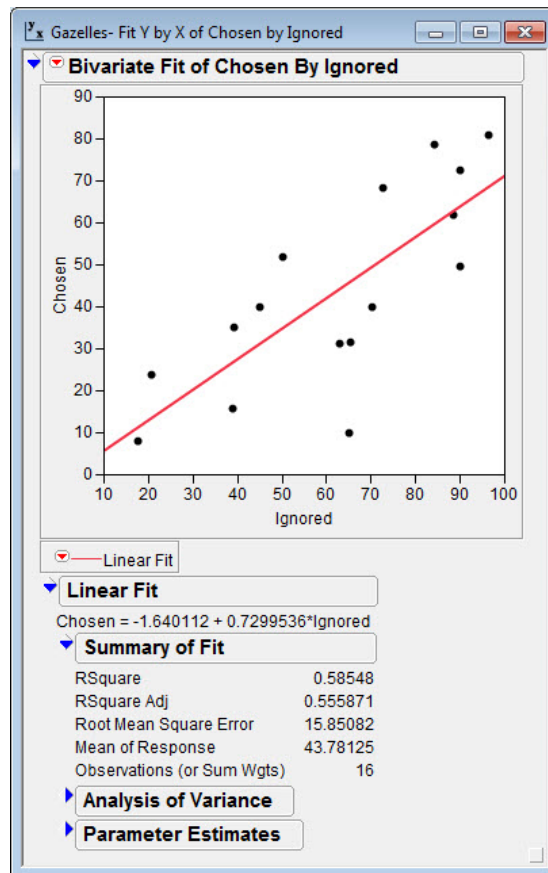
Notice the consistent presentation of graphs, especially the hot-spot icon. This hot-spot icon is where our bivariate analysis options will be presented—the options we need, and right when we need them.

Before proceeding, you might wish to revisit a JMP capability we have seen before. Check that you can modify each scale by using the right-hand icon, double-clicking on the axis areas. The points might be a bit small for some students, so we will enlarge them before adding the best-fit line to the scatterplot.

1. Right-click on the **Bivariate Fit** hot spot and select **Fit Line**.
2. Right-click on the graph area and select **Marker Size** → **4, XL**.
3. Click on the graph area and select **Line Width Scale** → **2.0**.
4. Click the **Analysis of Variance** and **Parameter Estimates** disclosure icons (the blue ones) to “undisclose” this information and thereby compact the **Fit Y by X** window.

The standard regression output is presented in figure 1.23. (You may need to enlarge your window to see it all.)

Figure 1.23 Best fit line



What Have We Learned?

In chapter 1 we learned the basic techniques of navigation with JMP and introduced the important contextual hot-spot and disclosure icons. We learned how to enter data and create files, as well as how to open and use existing files. Working through the gazelle data introduced the basic techniques of producing statistics and graphs in a univariate and bivariate analysis. It would not be too far afield to suggest that you already know 90% of how to generate univariate and regression reports commonly used in elementary statistics. In the next chapters, we will refine that knowledge, but at this point you are able to click your way anywhere and back with JMP.

References

- Bright, G. W., & S. N. Friel. (1998). Graphical representations: Helping students interpret data. In S. P. Lajoie (Eds.), *Reflections on statistics: Agendas for learning, teaching, and assessment in K–12* (pp. 63–88). Mahwah, NJ: Lawrence Erlbaum Associates.
- delMas, R. C., J. Garfield, & A. Ooms. (2005). *Using assessment items to study students' difficulty with reading and interpreting graphical representations of distributions*. Presented at the Fourth International Research Forum on Statistical Reasoning, Thinking and Literacy (SRTL-4), July 6, 2005, Auckland, New Zealand.
- Fitzgibbon, C. D. (1989). A cost to individuals with reduced vigilance in groups of Thomson's gazelles hunted by cheetahs. *Animal Behavior* 37(3): 508–510.