

Graphs from Questions, Questions from Graphs: Looking at Backpack Data

The Background: Some researchers were interested in learning about the weight of the backpack or book bag carried by math teachers. They were also interested in how well math teachers guessed the weight of their backpack or book bag. In 2006, the researchers conducted a study at a mathematics conference that attracted math teachers from across the US.

The Study: The researchers found a convenience sample of 52 teachers to participate in the study. Each math teacher was asked to estimate the weight of their book bag or backpack in pounds. The item was then weighed with a bathroom scale.

The Data:

Five pieces of information were recorded on each individual in the study:

Guess: The guessed weight of an individual's backpack (in lbs.)

Actual: The actual weight of an individual's backpack (lbs.)

Difference: Guess – Actual (lbs.)

Gender: Gender of each individual (m or f)

Under_3_lbs. Whether the individual's guess was less than 3 lbs. away or at least 3 lbs. away from the actual weight.

This Data can be found in the Fathom file “backpack_data_good” in the resources.

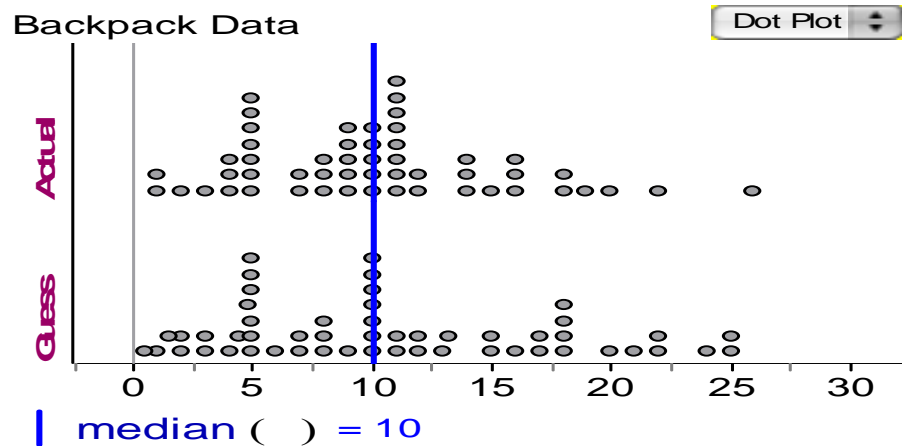
The questions:

The researchers want to use visual displays of the data to answer five questions about math teachers:

- 1) Overall, how heavy are math teachers' backpacks?
- 2) Which group, males or females, do a better job at guessing the weight of their backpack?
- 3) What fraction of people missed the correct weight of their backpack weight by less than 3 pounds?
- 4) Are math teachers good at guessing their backpack weights?
- 5) When math teachers guess the weight of their backpack incorrectly, do they tend to overestimate or underestimate?

Part I: Introductory Activity

Using a computer program, the researchers were able to make many different visual displays of the data. Here's one of the visual displays they made:



The blue bar on each dot plot indicates the median value of each distribution. In this case, the median for each sample was 10.

- a) In the first question, the researchers asked about the overall heaviness of math teachers' backpacks. If we assume that the subjects are a good representation of all math teachers, can we use this display to answer this question?
 - i) If so, make a conclusion in context, and cite *specific features of the display* that support your conclusion.
 - ii) If not, then explain why. In addition, *what question could this visual display answer?*

- b) For the other four research questions, answer parts **i)** and **ii)** again. Here are the remaining questions:
 - *Which group, males or females, do a better job at guessing the weight of their backpack?*
 - *What fraction of people missed the correct weight of their backpack weight by less than 3 pounds?*
 - *Are math teachers good at guessing their backpack weights?*
 - *When math teachers guess the weight of their backpack incorrectly, do they tend to overestimate or underestimate?*

Part II:

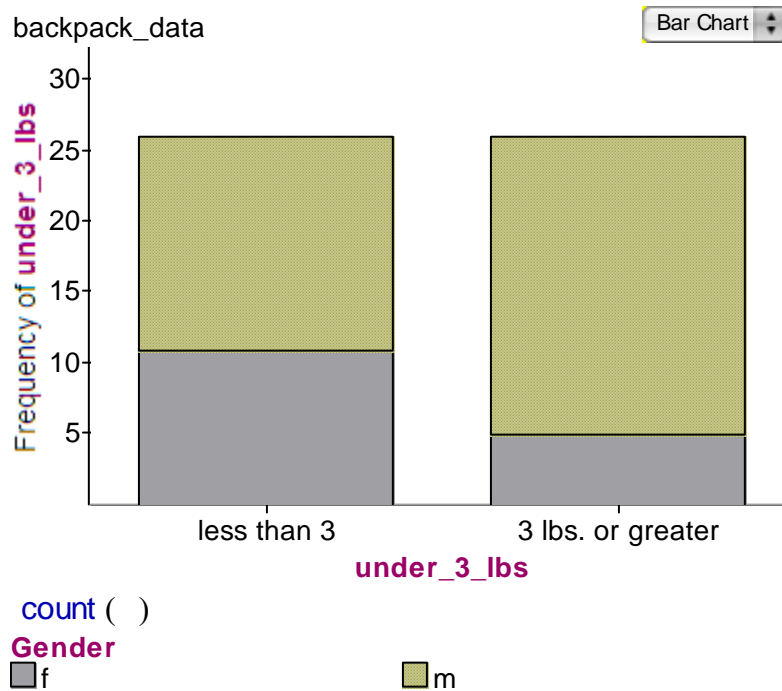
The researchers of the study created eight visual displays, in hopes of answering their research questions. These displays are found on the pages that follow.

3) Each display may help the researchers provide an answer to some, one, or none of the questions. For each display (labeled A-H), determine whether it can be used to help the researchers answer each of the five research questions (labeled 1-5) below. To keep track of your decisions, you can use the chart below.

Question	Display A	Display B	Display C	Display D	Display E	Display F	Display G	Display H
1. Overall, how heavy are math teachers' backpacks?								
2. M or F: who do a better job at guessing the weight of their backpack?								
3. What fraction of people missed the correct weight of their backpack weight by less than 3 pounds?								
4. Are math teachers "good" at guessing their backpack weight?								
5. Do math teachers tend to overestimate, or underestimate their backpack weight?								

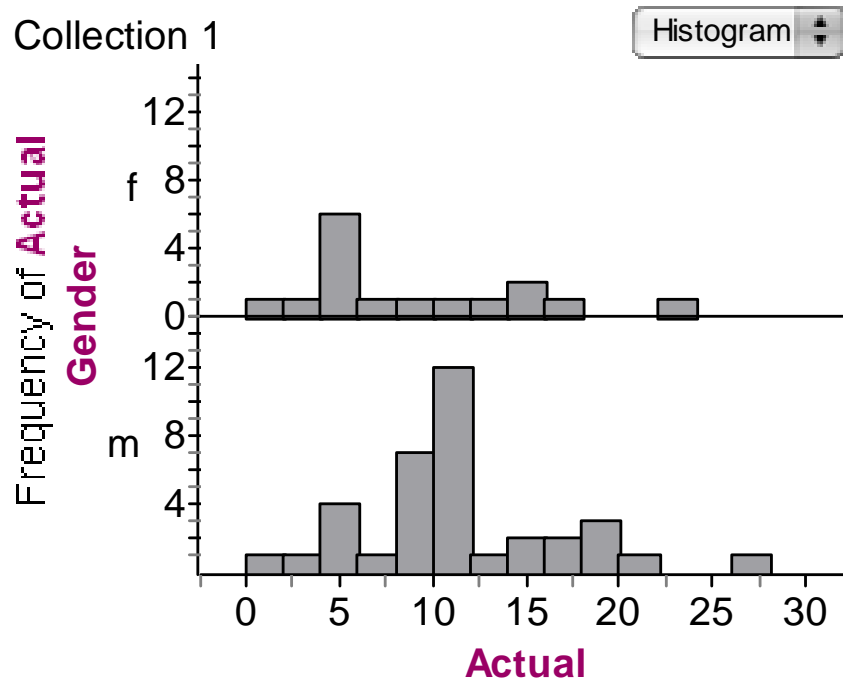
4) Look at the pages that follow. Under each visual display, each research question has been included. If appropriate, make a *specific conclusion about mathematics teachers* that could be included in a newspaper or research article. Justify your conclusion with specific evidence from the visual display.

Graph A



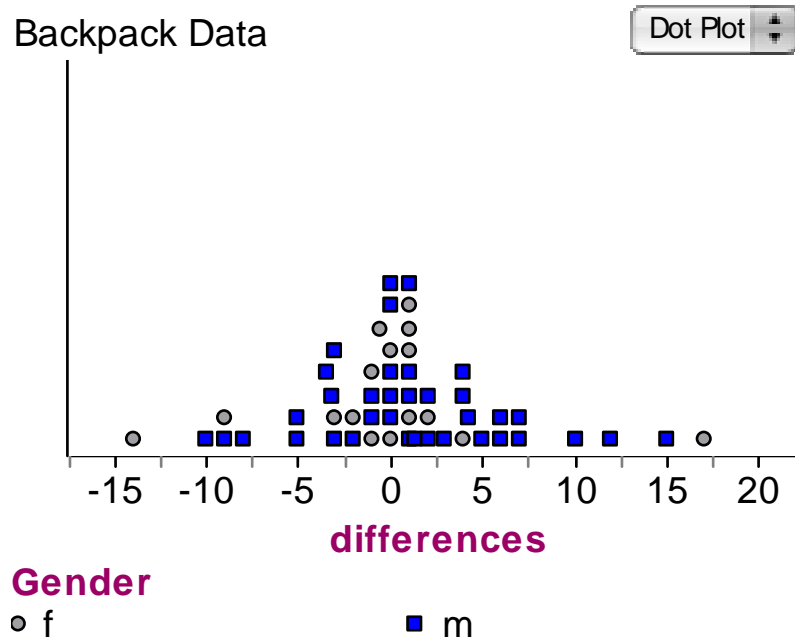
- 1) Overall, how heavy are math teachers' backpacks?
- 2) Which group, males or females, do a better job at guessing the weight of their backpack?
- 3) What fraction of people missed the correct weight of their backpack weight by less than 3 pounds?
- 4) Are math teachers good at guessing their backpack weights?
- 5) When math teachers guess the weight of their backpack incorrectly, do they tend to overestimate or underestimate?

Graph B



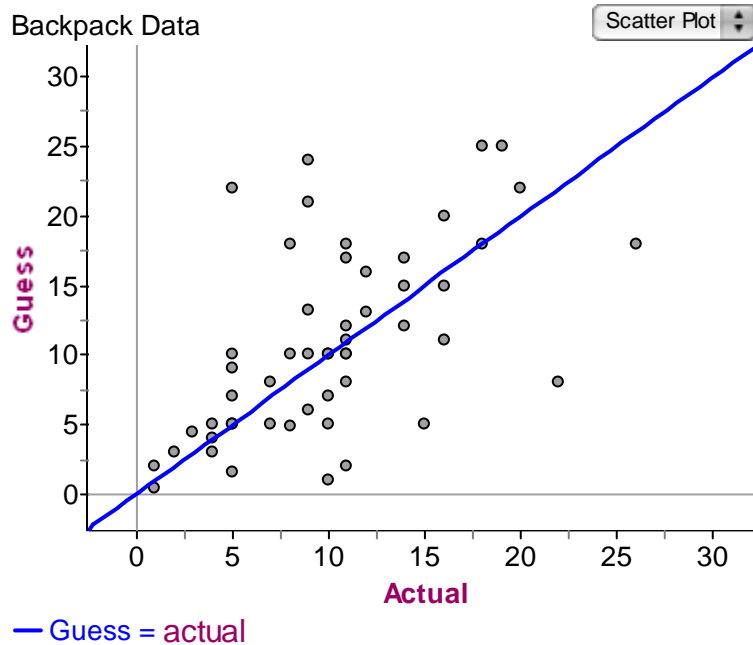
- 1) Overall, how heavy are math teachers' backpacks?
- 2) Which group, males or females, do a better job at guessing the weight of their backpack?
- 3) What fraction of people missed the correct weight of their backpack weight by less than 3 pounds?
- 4) Are math teachers good at guessing their backpack weights?
- 5) When math teachers guess the weight of their backpack incorrectly, do they tend to overestimate or underestimate?

Graph C



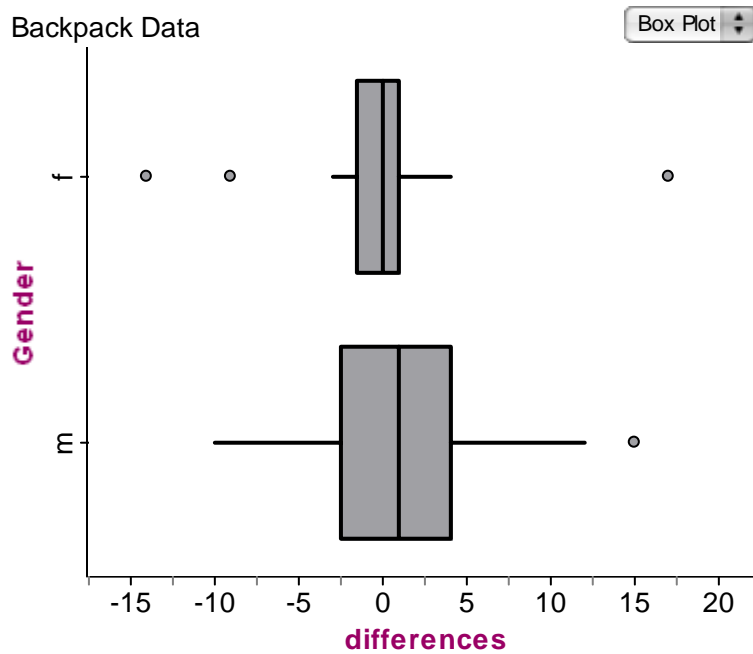
- 1) Overall, how heavy are math teachers' backpacks?
- 2) Which group, males or females, do a better job at guessing the weight of their backpack?
- 3) What fraction of people missed the correct weight of their backpack weight by less than 3 pounds?
- 4) Are math teachers good at guessing their backpack weights?
- 5) When math teachers guess the weight of their backpack incorrectly, do they tend to overestimate or underestimate?

Graph D



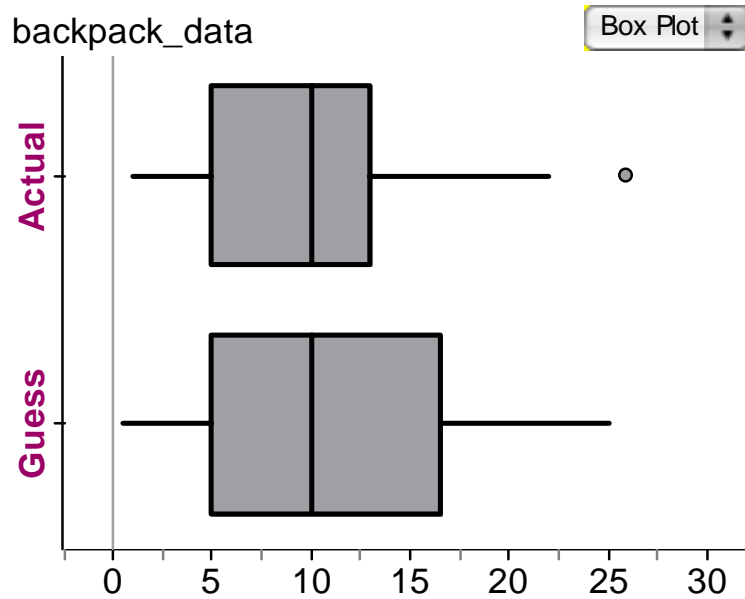
- 1) Overall, how heavy are math teachers' backpacks?
- 2) Which group, males or females, do a better job at guessing the weight of their backpack?
- 3) What fraction of people missed the correct weight of their backpack weight by less than 3 pounds?
- 4) Are math teachers good at guessing their backpack weights?
- 5) When math teachers guess the weight of their backpack incorrectly, do they tend to overestimate or underestimate?

Graph E



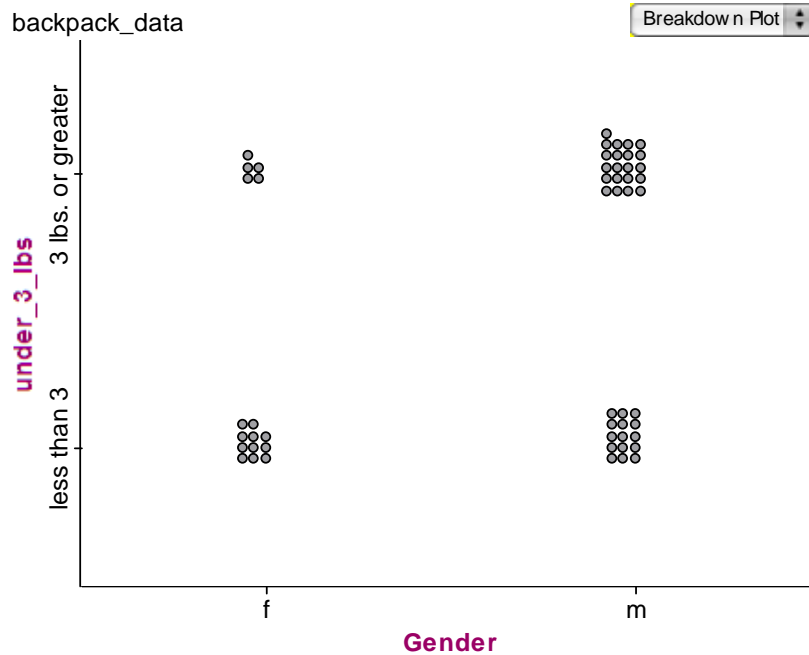
- 1) Overall, how heavy are math teachers' backpacks?
- 2) Which group, males or females, do a better job at guessing the weight of their backpack?
- 3) What fraction of people missed the correct weight of their backpack weight by less than 3 pounds?
- 4) Are math teachers good at guessing their backpack weights?
- 5) When math teachers guess the weight of their backpack incorrectly, do they tend to overestimate or underestimate?

Graph F



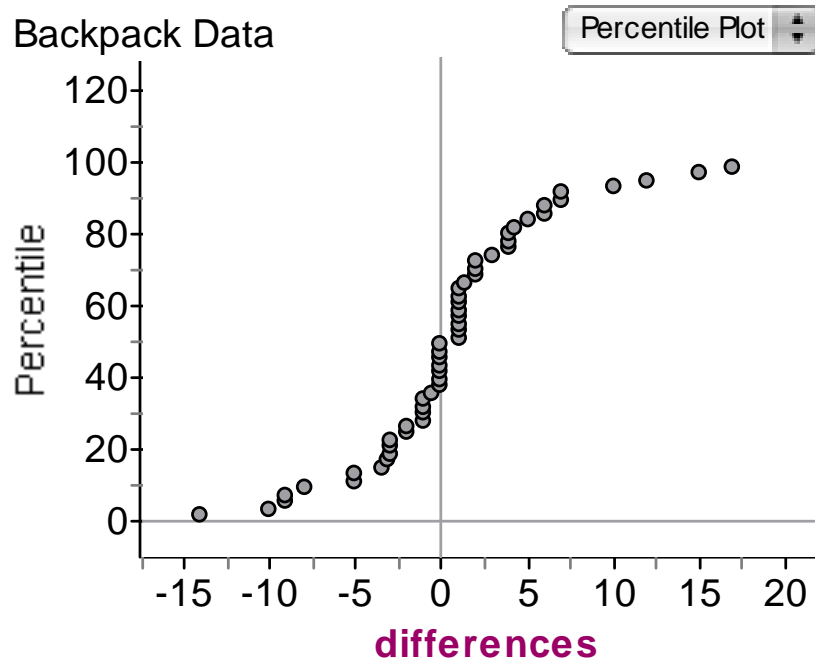
- 1) Overall, how heavy are math teachers' backpacks?
- 2) Which group, males or females, do a better job at guessing the weight of their backpack?
- 3) What fraction of people missed the correct weight of their backpack weight by less than 3 pounds?
- 4) Are math teachers good at guessing their backpack weights?
- 5) When math teachers guess the weight of their backpack incorrectly, do they tend to overestimate or underestimate?

Graph G



- 1) Overall, how heavy are math teachers' backpacks?
- 2) Which group, males or females, do a better job at guessing the weight of their backpack?
- 3) What fraction of people missed the correct weight of their backpack weight by less than 3 pounds?
- 4) Are math teachers good at guessing their backpack weights?
- 5) When math teachers guess the weight of their backpack incorrectly, do they tend to overestimate or underestimate?

Graph H



- 1) Overall, how heavy are math teachers' backpacks?
- 2) Which group, males or females, do a better job at guessing the weight of their backpack?
- 3) What fraction of people missed the correct weight of their backpack weight by less than 3 pounds?
- 4) Are math teachers good at guessing their backpack weights?
- 5) When math teachers guess the weight of their backpack incorrectly, do they tend to overestimate or underestimate?

Teacher Notes, *Graphs from Questions, Questions from Graphs***Objectives:**

This activity was designed to help students engage in the kinds of statistical thinking activities typically missing from text book lessons on visual displays of data. The activity is also designed to help students develop their statistical reasoning and literacy skills by creating well-justified explanations to their conclusions.

Background:

In 2000, Robert Del Mas articulated three different instructional domains in teaching statistics: statistical *literacy*, statistical *reasoning*, and statistical *thinking*. A description of these different domains can be summarized with the following table:

Verbs that may distinguish the three instructional domains (DelMas, 2002)

BASIC LITERACY	REASONING	THINKING
IDENTIFY DESCRIBE REPHRASE TRANSLATE INTERPRET READ	WHY? HOW? EXPLAIN (THE PROCESS)	APPLY CRITIQUE EVALUATE GENERALIZE

In short, statistical literacy involves an understanding of the definitions, procedures, and key concepts of statistical inquiry. Statistical *reasoning* requires the use of logic to make conclusions about the context the statistical tools describe. In statistical *thinking*, the student incorporates the overall investigative process: making/asking a question, creating a data collection plan, choosing how to analyze the data, and deciding what/how to conclude. For a more complete understanding of these domains, refer to DelMas (2000) or Thill(2007).

Although there is clear overlap among these three domains, this particular lesson emphasizes statistical *thinking*, because students select appropriate tools of analysis (visual displays) in order to appropriately address specific questions. Students develop their statistical reasoning and literacy skill in the process of justifying their conclusions throughout the activity.

Activity Time: One 90-minute block session, or two 45-minute sessions.

Setting: Part I will work as a whole-class discussion. Part II may best be implemented in a group setting, as different students will need to express their own conclusions and evaluate the conclusions and explanations made by their group mates.

Statistics Prerequisites:

- Students should be literate with a variety of visual displays: dot plot, histogram, box plot, scatter plot, two-way tables and plots, percentile plot.
- Students understand how to describe a distribution by its shape, center, and spread.
- Familiarity with comparing groups, making conclusions about groups with evidence.

Fathom Prerequisites:

Students do not need access to Fathom to complete Parts I and II of this activity. Optionally, students can further explore the data collected in this study with the accompanying Fathom data file: backpack_data_good.ftm

Procedure:

About the data: In this activity, students are asked to match up different statistical questions from the PCMI backpack data to an appropriate visual display. Students should be familiarized with the data collection process for the backpack data: 52 adults participating in the 2006 Park City Mathematics Institute were asked to estimate the weight of their book bags in pounds. Each subject's book bag was then weighed with a typical bathroom scale. The guessed and actual weights were recorded in pounds.

Part I: The first activity can be done in a whole-class setting so that students understand what will be expected in their conclusions and justifications for part II. However, some questions can be answered by more than one graph. Students should be prepared not only to choose graphs, but also to prepare an answer for each question and justify their conclusion with specific evidence from the graph(s) they chose. Some small groups may elect to answer a question in more than one way, but should be encouraged to provide at least one complete response for each question.

Part II: In groups of 4, invite students to determine which questions can be answered by each graph. In monitoring groups, press participants with prompts that invite them to more deeply justify their reasoning:

- So tell me, which specific feature of this display leads to your conclusion?
- Are there any other visual displays that also answer this question?
- What does your group mean by a “good guess?” Why is this a good criterion for “good guessing?”

The whole class can then compare the different conclusions found from each group, and discuss their explanations. Because some of the questions are broad enough to invite interpretation, the instructor should look for opportunities to highlight different interpretations of the same question. Students should be encouraged to ask questions that require their classmates to communicate their reasoning with detail and clarity.

Extending the lesson: If students have access to Fathom or another data analysis program, you can require them to create a new statistical question, choose (or create) two graphs to answer their question, and make conclusions based on those graphs.

Possible Answers, Part II: The table below can be used to determine how to use each graph to answer the question. If the graph is not usable, the box is blank.

	<i>How Heavy</i>	<i>Better Guessers, M or F?</i>	<i>What fraction missed by less than 3 lbs?</i>	<i>Are people good at guessing?</i> Note: "Good" will need to be defined by students. Answers will vary.	<i>Overestimate/Underestimate?</i>
Display A		Women. A majority of women and under 3 lbs, but a majority of men are 3 lbs or more.	$26/52 = 50\%$ missed by three lbs. or under.	If we accept "under 3 lbs" as "good," then about half are "good."	
Display B	Combine frequencies for M and F, to notice a center weight near 11 lbs. Most individuals within 4-5 pounds. Some right skewness. Student may also notice that men carry heavier backpacks.				
Display C		Women. 13/16 of women within 5 lbs, but a much smaller proportion of male dots are within this range.	By making off -3 and +3 on the "differences" axes, about half of the points inside these bounds.	This graph is most useful. Students can define good in many ways and use this graph	The number of individuals with difference > 0 is slightly higher than the number < 0 . Perhaps a slight tendency to overestimate.
Display D	The x-coordinates of most of the points on the scatter plot show "actual" near 10, most within 4-5 pounds of 10 lbs. This is hard to see,.		Draw the lines $guess = actual + 3$, $guess = actual - 3$. Observe about half the points inside these two boundaries.	This graph is also useful. Students can define good in many ways and use this graph to determine % points sufficiently close to "actual = guessed."	The number of points above "guessed = actual" is slightly higher than the number below. Perhaps a slight tendency to overestimate, but unclear if true for population.
Display E		The IQR for women is much smaller than for men. So women's guesses are more consistent.			Medians for each box plot are both near zero, with the men slightly skewed right. Perhaps a slight tendency to overestimate for men, not for women.
Display F	"actual" boxplot shows median near 10 lbs. slightly skewed right in the top quartile, but 1/4 of all between 10 and 13 lbs.				
Display G		For women 11/16 under 3 lbs, a much higher proportion than 15/36 for men in the study.	26/50 dots are in the under 3 lbs category.	If we accept "under 3 lbs" as "good," then about half are "good."	
Display H				If "Good" means within 2 lbs. Then find percentile for difference = 2, and difference = -2.	50 th %ile at 0, but 20 th at -2, and 80 th at +4. Suggest when incorrect, overestimates are larger than underestimates.

More Issues for discussion:

- a) **Sureness of Conclusions:** When students make conclusions, how confident are they about their answers? What facts would allow them to be more confident about their conclusions? More data? More information about the data collection process? More precise definitions?

- b) **Percentile Plot:** Although students may be familiar with percentiles, they may have never seen a percentile plot. These are also called *cumulative frequency plots*, or *ogives*. It will be valuable to monitor group discussions so that all students understand how to read these graphs. It's also useful to have students construct a box plot from the percentile plot as a way to assess their understanding.

- c) **Data Collection** The data from the math teachers were collected between 7:30 am and 8:00 am before morning sessions at a summer mathematics conference. Do these facts influence the conclusions?

- d) **Measuring percentage errors:** In defining a "good guess," Some students may want to define "error" in terms of a percentage of the actual weights. This idea is indeed an effective way of measuring errors. Only the scatter plot can allow us to measure individual's percentage error. For example, if a student wants to know which points fall 20% above or 20% below the actual weight, they should graph the lines "guess" = 1.2 "actual" and "guess" = 0.8 actual," respectively.

- e) **Creating quantitative definitions.** The question about "good guessing" is deliberately vague. As a result, students need to create a consistent definition of "good." Perhaps this means guessing your backpack within 2 pounds, or within 10% of the actual weight. Students should be expected to clearly describe their rule, justify why it means the guess is "good," and make correct conclusions based of their definition of "good."

References:

- DelMas, R. (2002). Statistical Literacy, Reasoning, and Learning: A Commentary. *Journal of Statistics Education*, 10(2). [Electronic version] website: http://www.amstat.org/publications/jse/v10n2/delMas_introduction.html
- Franklin, C., Garfield, J. (2006). The GAISE Report: Developing Statistics Education Guidelines for Pre-K-12 and College Courses. In P. Elliott (Series Ed.) & G. Burrill (Vol. Ed.), (Vol. 68, pp. 345-376) Reston, VA: National Council of Teachers of Mathematics.
- Key Curriculum Press (2007) . Fathom Dynamic Data Software, Version 2.1. Key Curriculum Press Technologies.
- Thill, W. Implementing Statistical Inquiry in the Classroom (2007). Unpublished paper, Completed in partial fulfillment of Master of Arts in Secondary Education, California State University, Northridge, August 2007.