Working Hard or Hardly Working?

About the Lesson

Students will analyze univariate and bivariate data. Questions are posed for discussion, further research, and algebraic problem solving. As a result, students will:

- Analyze a box plot.

- Analyze a histogram.

- Use a scatter plot to generate a line of best fit.

- Use a linear regression equation to answer questions.

Vocabulary
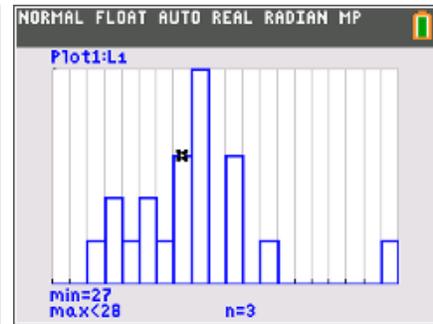
- line of best fit
- linear regression

Teacher Preparation and Notes

- It is expected that students have already encountered box plots, histograms, and scatter plots. The concepts of median, frequency, and mean are called upon in this activity.

- Students will manipulate a manual-fit line on the scatter plot to obtain their best visual fit for the data and use it to answer questions.

- Students will also create the linear regression and use it to evaluate and to solve problems.

Activity Materials

- Compatible TI Technologies:

    TI-84 Plus*
    TI-84 Plus Silver Edition*
    🌐TI-84 Plus C Silver Edition
    🌐TI-84 Plus CE

* with the latest operating system (2.55MP) featuring MathPrint™ functionality.

Tech Tips:

- This activity includes screen captures taken from the TI-84 Plus CE. It is also appropriate for use with the rest of the TI-84 Plus family. Slight variations to these directions may be required if using other calculator models.
- Watch for additional Tech Tips throughout the activity for the specific technology you are using.
- Access free tutorials at http://education.ti.com/calculators/pd/US/Online-Learning/Tutorials
- Any required calculator files can be distributed to students via handheld-to-handheld transfer.

Lesson Files:

- Working_Hard_Student.pdf
- Working_Hard_Student.doc
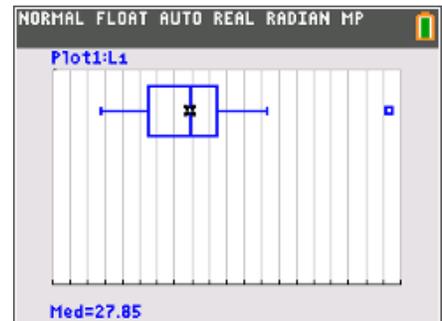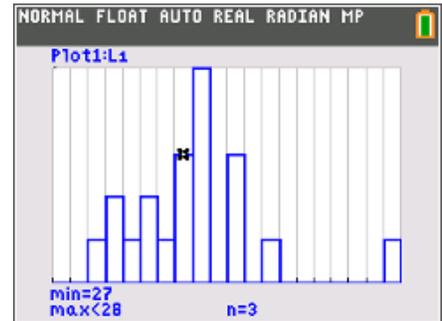- L1.8xl
- L2.8xl
- L3.8xl

Tech Tip: Before beginning the activity, the files L1.8xl, L2.8xl, and L3.8xl need to be transferred to the students' calculators via handheld-to-handheld transfer or transferred from the computer to the calculator via TI-Connect™ CE Software.

Problem 1 – Histograms and Box Plots

Students are given the average commute time for workers in 20 US cities (**L1**). Students will examine this data as a histogram and a box plot.

For more data on commute times and other facts, go to http://factfinder.census.gov.

To take the data further, have students do a survey of employees in the building of their commute times. Have students compare this information to the data given in the list.

1. Examine the histogram. How is the data distributed?

    Sample Answer: The data is distributed towards the center.

2. According to the histogram, how many cities have an average commute time between 25 and 29 minutes?

    Answer: 9

3. What is the median commute time? Press ⟨trace⟩ and use the arrow keys.

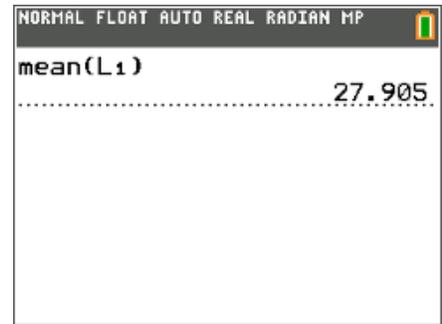   **Answer:** 27.85

To find the mean commute time, students can use the **mean(** command (2nd [list] and arrow to the MATH menu).

Ask:

- *What factors may affect the commute times? [population density, public transportation, carpooling, weather, etc.]*

```
NORMAL FLOAT AUTO REAL RADIAN MP
mean(L₁)
                              27.905
```

4. What is the mean commute time?

Answer: 27.905

5. Are the median and mean good representations of the data? Why or why not?

Sample Answer: They are both good representations because they are very close in value.

6. What is the longest commute time? How do you know?

Answer: The longest commute time is 39.3 because it is the data point furthest to the right in the box plot.

7. Which graph should be used if we wanted to know how often a certain time occurs in the data, the box plot or the histogram? Explain your reasoning.

Answer: The histogram because it shows the frequency of small ranges when the box plot does not.
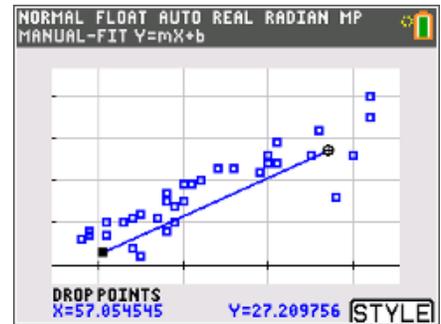
Problem 2 – Bivariate Data

Students will examine data from a survey about teachers' age and number of years' experience teaching (**L2** and **L3**). Students should be able to mention the ideas of linear behavior, an increasing or positive slope, and a strong or a weak correlation. Some students may mention that there are some dots that appear to be farther away from the general shape of the distribution. These special cases can be discussed. For example, a woman may have stayed at home while her children were young and then returned to teaching. Or a person who was unhappy in a corporate job or a research position came back to education after that work experience.

8. Set up a scatter plot as shown. What trends do you see in the scatter plot of the data?
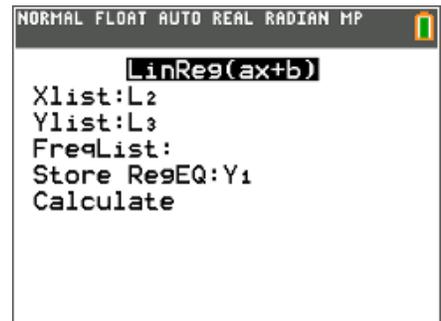
Answer: The data has a positive correlation.

Students will use a manual-fit line to find the equation of a line that appears to 'fit' the data best. Once the scatter plot is set up, press zoom and select **ZoomStat**. Then press stat, move to the CALC menu, and select **Manual-Fit Y=mX+b**. Enter **Y2** (vars **> Y-VARS > Function…**) for Store EQ and press enter twice. The cursor will appear on the screen and students should move the cursor around to place one point of the manual-fit line where desired. Then, move the other end to the second location to create a line. The equation will then be shown at the top of the screen or in **Y2**. Once the line is placed, students can modify **a** and **b** to adjust their line.
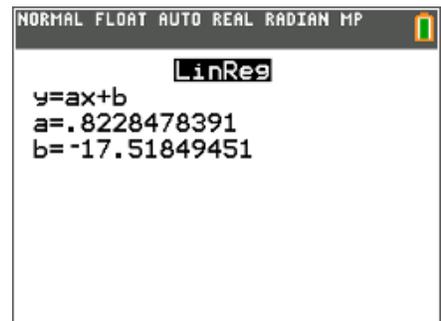

9. What is the equation of your line?

Sample Answer: $y = 1.01x - 24.57$

Next, students will display the linear regression of the data (stat, arrow to the CALC menu and select **LinReg(ax+b)**). It is assumed that the students have performed a regression at this point. Rich discussion can take place now that students understand the slope (change in years of experience for each year older) of linear functions.

Correlation is not shown as part of this task. If you do not want students to see the values of r^2 and r, have them press 2nd [catalog], select **DiagnosticOff**, and then press enter on the Home screen.

10. What is the equation of the regression line?

Answer: $y = 0.8228x - 17.5185$

11. How does the regression line compare to the line you 'fit' above?

Sample Answer: My line has a slightly higher slope than the regression line, but they are close.

12. In this situation, what is the significance of the value for *a* in the regression equation?

Sample Answer: The value for *a* is the ratio of the teacher's age and the number of years experience they have.

Be sure to show students how to perform the associated algebra for solving the next few problems. To find the mean age, students can type **mean(L2)** on the Home screen.

Students should remember, in the form $y = mx + b$, *x* represents age. So, students will substitute the mean age for the variable *x* in the regression equation and simply evaluate the right side. Their answer will be the corresponding number of years of teaching experience.
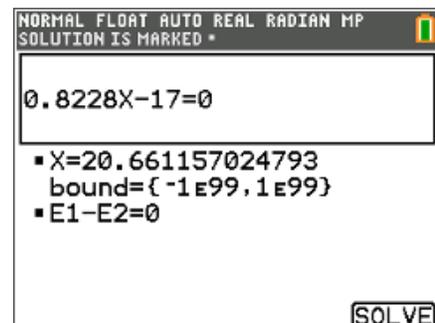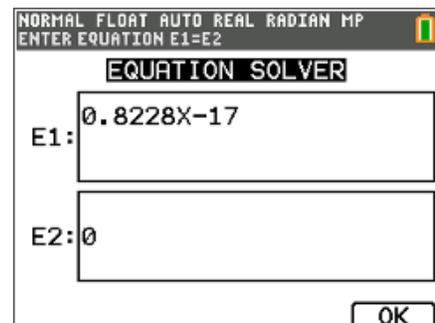
13. What is the mean age? Use this age and the regression line to determine the number of years' experience the "average"-aged teacher has.

Answer: The mean age is 42.4 years. The number of years' experience the average-aged teacher has is 17.36 years.

Then, students will solve to find at what age the typical teacher begins his or her career and at what age a teacher is about to retire. This may be surprising to some students!

You choose how you wish for students to solve these problems. One option is the **Numeric Solver** command under the math menu. Enter **0.8228X–17** into **E1** and **0** for **E2** and press enter.

On the next screen, move the cursor to the **X** variable and press graph (for solve).

14. According to the regression equation, what is the age that a teacher begins his/her teaching career?

Answer: 17

15. Teachers are considered "retirement age" after about 30 years' teaching experience. At what age is a teacher who is about to retire?

Answer: 57