

One and Two Variable Statistics – Review

Time Required
40-60 minutes

ID: 11490

Activity Overview

In this activity, students will review the concepts that they have learned thus far in statistics. The first part of the activity includes one-variable topics such as graphing quantitative variables, calculating measures of central tendency and spread, and making comparisons. The second part includes two-variable topics, including linear regression, correlation, and interpolation. The coefficient of determination will be introduced.

Topic: Review

- *One-variable statistics*
- *Two-variable statistics*

Teacher Preparation and Notes

- *This can be used as a review lesson or an assessment. Students must have a working knowledge of graphing histograms and box plots, measures of central tendency and spread, quartiles, scatter plots, lines of best fit, linear regression, and interpolation.*
- *Calculator skills – Students should be able to use the TI-Nspire to find measures of central tendency and spread, graph histograms and box plots, graph scatter plots, and find lines of best fit using the movable line and linear regression.*
- *The student TI-Nspire document has the data set and directions. It should be used in conjunction with the student worksheet.*
- *Notes for using the TI-Nspire™ Navigator™ System are included throughout the activity. The use of the Navigator System is not necessary for completion of this activity.*
- *This activity uses data from the World Health Organization. Go to <http://apps.who.int/ghodata/?vid=3200&theme=country> for more data about each country and other countries.*
- ***To download the student and solution TI-Nspire documents (.tns files) and student worksheet, go to education.ti.com/exchange and enter “11490” in the keyword search box.***

Associated Materials

- *1and2VarReview_Student.doc*
- *1and2VarReview.tns*
- *1and2VarReview_Soln.tns*

Suggested Related Activities

To download any activity listed, go to education.ti.com/exchange and enter the number in the keyword search box.

- *Analyzing the Steelers 2008 Super Bowl year (TI-Nspire technology) — 13790*
- *Yankees Vs. Mets (TI-Nspire CAS technology) — 10997*
- *A Data and Statistics Example (TI-Nspire technology) — 9674*

Data

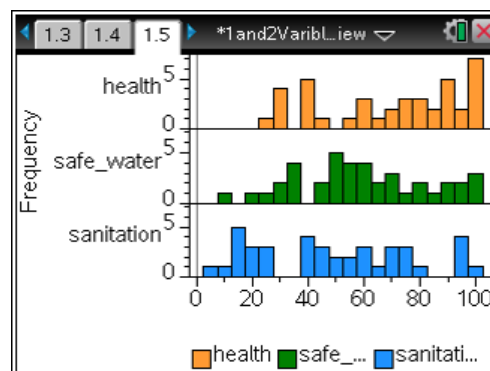
The spreadsheet on page 1.3 contains data for health indicators for different Asian and African countries. The first part of the activity will focus on the first three sets of data (columns B, C, D), which represent the percentages of people who have access to health services, safe water, and sanitation. The second part of the activity will begin by looking at the data in all of the columns, but then focus on comparing life expectancy and the infant mortality rate.

	country	health	safe_w...	sanitati...
1	Algeria	90	70	60
2	Angola	30	34	18
3	Benin	30	54	42
4	Botswan...	89	60	42
5	Cambod...	53	37	15

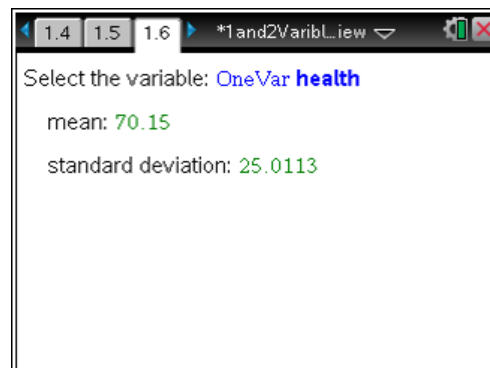
One-Variable Statistics

Students are asked to compare the three distributions that describe the percentages of people who have access to health services, safe water, and sanitation.

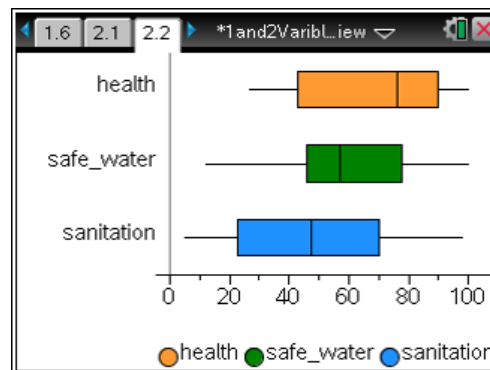
They will need to first create histograms, on page 1.5, and box plots, on page 2.2. To do this, they need to press **MENU > Plot Properties > Add X Variable** and select **health** from the list. Repeat the command, selecting **safe_water** and then selecting **sanitation**. Then change the plot type to Histogram. (**MENU > Plot Type > Histogram**) If students adjust the x-axis or the y-axis, it will change for all of the graphs, not just one.



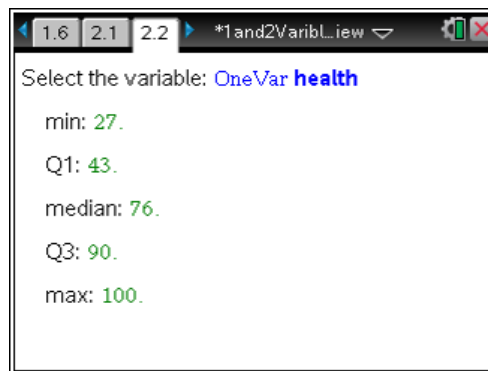
Students should then determine the measures of central tendency and spread. For the histograms, students can use page 1.6 by changing the bold word in the first line to that of the desired list. (Delete the current list name, press **[var]**, and select the list.) The mean and standard deviation values will automatically update.



Explain to students when comparing the box plots that the scale is the same for each. Students can roll over the parts of the box with the cursor and the numbers of the five-number summary will appear.



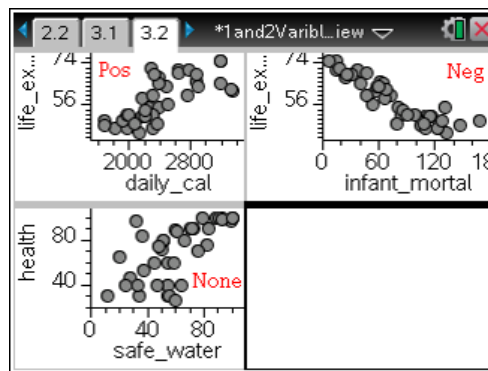
Or students can use page 2.2 to calculate the five-number summary, in the same manner they used page 1.6. Note: Students should only change the name of the list in the first line. The other numbers will automatically update.



Two-Variable Statistics

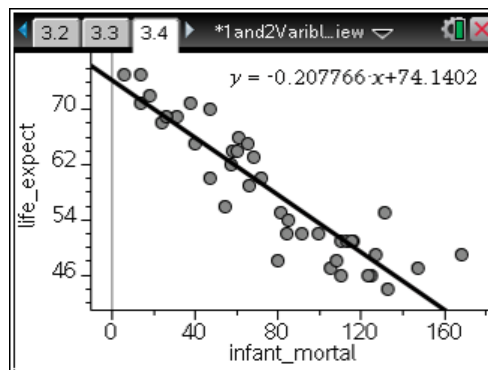
Students are asked to investigate relationships between two variables, finding a positive correlation, negative correlation, and no correlation.

They will probably find the two variables that show these types of correlation through trial and error, but encourage them to think about how one variable might affect another. For example, one may think that if you drink better water, then you will have a healthier life and live longer, so life expectancy and safe water might have a correlation.

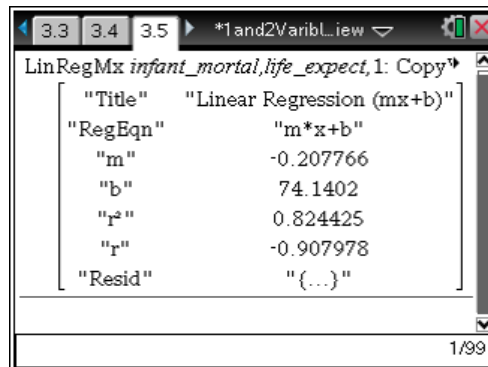


Students will answer Questions 7 through 14 using the scatter plot of infant mortality vs. life expectancy. They should have **infant_mortal** on the bottom as the independent variable and **life_expect** on the left side as the dependent variable.

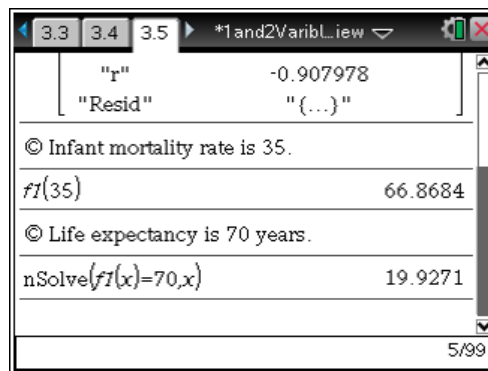
After finding their own line of best fit, students will find the linear regression equation. They can show the line on the scatter plot by selecting **MENU > Analyze > Regression > Show Linear(mx+b)**.



However, students will also need to calculate the regression on a *Calculator* page so that they can see the values of r and r^2 . (**MENU > Statistics > Stat Calculations > Linear Regression (mx+b)**)



The *Calculator* page can also be used to answer Questions 13 and 14 to predict the life expectancy and mortality rates for the numbers given.



Solutions to Student Worksheet

One-Variable Statistics

	Health Services	Safe Water	Sanitation
Best measures of central tendency and spread	Min = 27 Q1 = 43 Median = 76 Q3 = 90 Max = 100	Mean = 59.55 sx = 23.33	Min = 5 Q1 = 22.5 Median = 47.5 Q3 = 70 Max = 98
Justification	Skewed to the left, so the median and 5-number summary better describe the data.	Symmetric, so mean and standard deviation best describe the data.	Skewed to the right, so the median and 5-number summary better describe the data

1. Sample response: Sanitation needs the greatest improvement since it has the lowest measure of central tendency. It also has the lowest minimum value. Over 50% of the countries have less than 50% of the people with access to sanitation. 75% of the countries have less than 70% of the people with access to sanitation.
2. Angola (46); Central Africa Republic (47); Zaire (52)
3. Students should justify any statement using a graph and/or a numerical statistic.
Sample response: More money should be given to sanitation in these countries. This is the greatest need. This can be seen, as it has the lowest median value of the three variables measured.

Two-Variable Statistics

4. Answers will vary. Life expectancy vs. safe water, sanitation, daily calorie intake and health all work. These variables give a positive correlation because an improvement in any of these variables should lead to a longer life. Daily calorie intake has the strongest relationship.
5. Answers will vary. Life expectancy vs. Infant mortality. Practically, as a country can reduce the death of infants, the life expectancy should increase.
6. Answers will vary. Safe water and health have very little relationship. Practically, countries with minimal finances will have to choose between health care and safe water for their people.
7. It appears linear with a negative correlation. It is a relatively strong relationship.
8. Answers will vary. Sample equation: $y = -0.22x + 75$
Their justification for fit should be specific and should reference either the sum of the residual squares or the properties of a line of best fit.
9. $f_1(x) = -0.207766x + 74.1402$
The slope means that for each increase of .2 deaths per 1000 (or 2 per 10000) the life expectancy will decrease by 1 year. The units for slope are years per deaths per 1000.
If there is an infant mortality rate of 0, then the life expectancy will be 74.14 years.
10. $r = -0.90797$
11. $r^2 = 0.824425$
12. It tells us that 82.44% of the variation in life expectancy is due to the infant mortality rate.
13. Life expectancy is 66.87 years.
14. The infant mortality rate is about 20.