|  |
| --- |
| * http://textbooks.cpm.org/images/cc3/chap07/CC3_7.3.3title.pngIn previous lessons, you described the association between two numerical variables, such as the amount of fertilizer used and the height of the plant.  Some variables, however – such as gender, eye color, names of countries, or weather conditions – are not numerical.  Since the data are in categories, non-numerical variables are called **categorical variables**.  Another type of categorical variable occurs when numerical variables are lumped into categories, such as in age groups.  Today you will look for relationships in variables that are not numerical. * **7-107.** The experiments students did with plants and different growing conditions in Lesson 7.1.3 caught the attention of a local farmer.  He was interested in whether the type of soil made a difference in the height of his corn crop.  He planted 2500 stalks of corn and collected the following data.  This type of table is called a **frequency table** because it shows counts, or frequencies, in each of the cells of the table.   1. http://textbooks.cpm.org/images/cc3/chap07/CC3_7-107_table.pngMake a conjecture about the effect of soil type on the height of the corn.   2. The table of data on corn height used counts because it counted the number of stalks of corn.  When analyzing categorical data, percents are much easier to analyze.  But first, you need to determine the independent variable so that you can determine, “Percent of what?”  What is the independent variable in this situation?   3. Create a second table that contains the data above but as percentages instead of counts.  The percentage should be the count for each height category out of the total number for that soil type.  For example, the height of the bar for 0-3 ft will be 9.0%, because  is 0.090.  The third row of the table is not needed in this case.   **http://math-lessons-collab.wikispaces.com/file/view/1st_quadrant_graph.GIF/33331357/420x333/1st_quadrant_graph.GIF7-108.** The farmer is interested in the height of his corn for each of the two soil types.  A straightforward way to compare the effect of the independent variable is to make a different bar graph for each independent variable.   * 1. Make a bar graph for the sandy soil.  The horizontal axis should be the height categories.  The vertical axis will represent the percent of the corn in sandy soil that grew to that height.   2. Make a similar bar graph for the other independent variable.   3. Is the height of corn associated with the soil type?  That is, does the soil type have an impact on height?  Report your conclusions to the farmer.   4. Why did you make a bar graph instead of a histogram for the height of the corn?   **http://math-lessons-collab.wikispaces.com/file/view/1st_quadrant_graph.GIF/33331357/420x333/1st_quadrant_graph.GIF7-109.** The nutrition staff at CPM Diabetic Institute is interested in the impact of three new energy bars on blood sugar levels.  The staff conducted a study on 1000 volunteers and collected the following data.   * 1. Complete the table by computing row and column totals.  What is the independent variable?   2. http://math-lessons-collab.wikispaces.com/file/view/1st_quadrant_graph.GIF/33331357/420x333/1st_quadrant_graph.GIFhttp://math-lessons-collab.wikispaces.com/file/view/1st_quadrant_graph.GIF/33331357/420x333/1st_quadrant_graph.GIFA **relative frequency table** displays the percents in a table instead of a bar graph.  Change the frequency table at the beginning of this problem to a relative frequency table by changing the counts to percents.  For example, the 30 Mighty Bars that lowered blood sugar will be displayed as 20%.   3. http://textbooks.cpm.org/images/cc3/chap07/CC3_7-109_table.pngUse your data from the relative frequency table to make a bar graph for each of the independent variables.  The horizontal axis should be the dependent variable.   4. Is there an association between blood sugar level and the choice of energy bar? |