Exploring Statistics Using Fathom
Correlation Coefficient and Least-Squares Homework

Homework:

1. **In Fathom:**

    (a) Calculate the Pearson correlation coefficient, r, for this data.

      i. Download the Fathom file `GunDeaths.ftm` from

       `http://paws.wcu.edu/emcnelis/StatsExamples.html`.

      ii. Generate new attributes for the z-scores of the Year and Rate attributes, and the product of those z-scores.

      iii. Make a scatterplot of the z-scores for the death rate versus the z-scores for the year.

      iv. Use these new attributes to calculate the Pearson correlation coefficient, r, for this data using Fathom.

    (b) Moveable lines, sum of squares, and least-squares lines

      i. Create a scatterplot of the death rate versus the year, and add a moveable line and show the squares. Adjust the line so that you have achieved the smallest sum of squares that you can manage. Add the least-squares line and compare your answer to this.

      ii. Repeat this process on your scatterplot of the z-scores.

2. **The Report:**
   You are to address the following issues in a formal write-up, written in Word or another text editor. Make sure that you answer the questions completely and use images from your Fathom worksheet to support your argument. You will be graded on your accuracy, the thoroughness of your argument, and your presentation (including grammar).

    (a) Explain why the correlation coefficient, r, value is not surprising given the scatterplot of your original data or the scatterplot of your z-scores. Interpret the meaning of your correlation coefficient in terms of the relationship between the year and the death rate due to guns (be thorough here). Use illustrations in your explanations and arguments.

    (b) Explain the connection between the magnitude of the correlation coefficient and the ability to "fit" a straight line to the data.

    (c) Although the scatterplots of our original data and of our associated z-scores are identical, the equations of their best fit lines and the associated sum of squares are not identical. Give your best explanation for why this might be the case.