Exploring Statistics Using Fathom Graphical Displays of Data

Consider the data given in the Fathom file, SATData.ftm¹, available at

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

- 1. Create a scatterplot of the mean math scores in 2004 versus the mean math scores in 2003 (MeanM03).
- 2. Approximate the value of r, the Pearson correlation coefficient, for this data. Give justification for your answer (i.e. explain why you believe such a value is accurate).

3. Does there appear to be any pattern or relationship between the math scores in 2003 and 2004 given this scatterplot? If so, explain the relationship you've noticed; if not explain why there appears to be no relationship and what this means in terms of yearly math scores.

4. What assumptions do you feel safe in making about states with low math scores on the SAT's?

¹The data, and the experiment explanation given below, were taken from the College Board Website: http://www.collegeboard.com/about/news_info/cbsenior/yr2004/reports.html

5. Create a new attribute, ParticipLevel04, and define it as follows:

$$if(ParticipRate04 < 25) \begin{cases} "low" \\ if(ParticipRate04 < 75) \end{cases} \begin{cases} "med" \\ "high" \end{cases}$$

6. Drop the new ParticipLevel04 attribute name directly on top of your scatterplot. How does this change the graph? Does this change your previous assumptions about states with low math scores? What does this additional information tell us? Explain.

- 7. Respond to the following arguments provided by the College Board web site along with this data:
 - (a) "The SAT is a strong indicator of trends in the college-bound population, but it should never be used alone for such comparisons because demographics and other non-school factors can have a strong effect on scores."

(b) "If ranked, schools and states that encourage students to apply to college may be penalized because scores tend to decline with a rise in percentage of test takers."