Illustrating the Statistical Process with Regression

Josh Tabor
Canyon del Oro High School
joshtabor@hotmail.com

Daren Starnes
The Lawrenceville School
dstarnes@lawrenceville.org

At the end of the session, participants will

1. Become familiar with the Common Core State Standards for analyzing two quantitative variables.
2. Understand that statistics is a problem-solving process, not a set of isolated skills.
3. Obtain engaging, classroom-tested examples that illustrate the statistical process using regression contexts.
4. Learn how to use graphing calculators and online applets to analyze two quantitative variables.

The Statistical Problem-Solving Process

1. Ask Questions: Clarify the research problem and ask one or more valid statistics questions. *A valid statistical question is based on data that vary.*

2. Collect Data: Design and carry out an appropriate plan to collect the data. *Choices include surveys, observational studies, and experiments.*

3. Analyze Data: Use appropriate graphical and numerical methods to analyze the data. *Bar charts, dotplots, scatterplots, means, standard deviations, and so on.*

4. Interpret Results: Draw conclusions based on the data analysis. *Answer the question using inferential methods, including significance tests and confidence intervals.*

See the *Guidelines for Assessment and Instruction in Statistics Education* (GAISE) for much more information: http://www.amstat.org/education/gaise/.

Grade 8 » Statistics & Probability

Investigate patterns of association in bivariate data.

8.SP.A.1
Construct and interpret scatter plots for bivariate measurement data to investigate patterns of association between two quantities. Describe patterns such as clustering, outliers, positive or negative association, linear association, and nonlinear association.

8.SP.A.2
Know that straight lines are widely used to model relationships between two quantitative variables. For scatter plots that suggest a linear association, informally fit a straight line, and informally assess the model fit by judging the closeness of the data points to the line.

8.SP.A.3
Use the equation of a linear model to solve problems in the context of bivariate measurement data, interpreting the slope and intercept. *For example, in a linear model for a biology experiment, interpret a slope of 1.5 cm/hr as meaning that an additional hour of sunlight each day is associated with an additional 1.5 cm in mature plant height.*

High School » Interpreting Categorical & Quantitative Data (S-ID)

Summarize, represent, and interpret data on two categorical and quantitative variables

HSS.ID.B.6
Represent data on two quantitative variables on a scatter plot, and describe how the variables are related.

HSS.ID.B.6.A
Fit a function to the data; use functions fitted to data to solve problems in the context of the data. Use given functions or choose a function suggested by the context. Emphasize linear, quadratic, and exponential models.

HSS.ID.B.6.B
Informally assess the fit of a function by plotting and analyzing residuals.

HSS.ID.B.6.C
Fit a linear function for a scatter plot that suggests a linear association.

Interpret linear models

HSS.ID.C.7
Interpret the slope (rate of change) and the intercept (constant term) of a linear model in the context of the data.

HSS.ID.C.8
Compute (using technology) and interpret the correlation coefficient of a linear fit.

HSS.ID.C.9
Distinguish between correlation and causation.

Example 1: Seating Charts

Ask Questions

Collect Data

Analyze Data

	1	1	1	1	2	2	2	2	3	3	3	3	4	4	4	4	4
	76	77	94	99	83	85	74	79	90	88	68	78	94	72	101	70	79
	5	5	5	5	5	6	6	6	6	7	7	7	7				
	76	65	90	67	96	88	79	90	83	79	76	77	63				

Applet available at: www.tinyurl.com/SPAapplets

Interpret Results

Example 2: Ford F-150s

Ask Questions

Collect Data

Analyze Data

	70,583	129,484	29,932	29,953	24,495	75,678	8,359	4,447
	21,994	9,500	29,875	41,995	41,995	28,986	31,891	37,991
	34,077	58,023	44,447	68,474	144,162	140,776	29,397	131,385
	34,995	29,988	22,896	33,961	16,883	20,897	27,495	13,997

Interpret Results