Using the GAISE report to improve student projects

Jared Derksen

November 6, 2015

CMC

Palm Springs, California

Technology Resources

GAISE report from ASA

http://www.amstat.org/education/gaise/

StatKey

http://lock5stat.com/statkey/index.html

Regression resource

http://seeit.ucdavis.edu

More regression at

http://www.desmos.com

Plenty of stats tools at

http://geogebra.org

Free instructor's software, $5/student/year

http://www.statcrunch.com

Free graphs

http://www.plot.ly

DC

448.1	116.6
534.9	66.5
134.2	107.5
251.2	70.5
291.0	60.0
184.0	40.2
200.1	19.8
206.9	14.3
162.8	23.6
107.3	3.0
108.2	10.5
76.0	1.7
	0.2

Marvel

623.4	222.3	18.3
409.0	157.3	24.4
403.7	154.7	33.8
336.5	176.7	28.8
373.6	146.4	8.1
262.0	131.9	
232.5	132.2	
259.7	134.8	
202.8	115.8	
312.4	102.5	
206.4	82.3	
318.4	70.1	
234.4	52.4	
214.9	51.8	
181.0	48.1	

Domestic Box Office (millions)

GAISE

The PreK-12 report is a very helpful document to assist teachers in the teaching of Statistics education. The document can be found at http://www.amstat.org/education/gaise/. AP Statistics teachers will recognize that good friends of our AP program helped write this document!

Here are some helpful excerpts.

The Difference between Statistics and Mathematics

"Statistics is a methodological discipline. It exists not for itself, but rather to offer to other fields of study a coherent set of ideas and tools for dealing with data. The need for such a discipline arises from the *omnipresence of variability*." (Moore and Cobb, 1997)

A major objective of statistics education is to help students develop statistical thinking. Statistical thinking, in large part, must deal with this omnipresence of variability; statistical problem solving and decision making depend on understanding, explaining, and quantifying the variability in the data.

It is this focus on *variability in data* that sets apart statistics from mathematics.

The Nature of Variability

There are many sources of variability in data. Some of the important sources are described below.

Measurement Variability—Repeated measurements on the same individual vary. Sometimes two measurements vary because the measuring device produces unreliable results, such as when we try to measure a large distance with a small ruler. At other times, variability results from changes in the system being measured. For example, even with a precise measuring device, your recorded blood pressure could differ from one moment to the next.

Natural Variability—Variability is inherent in nature. Individuals are different. When we measure the same quantity across several individuals, we are bound to get differences in the measurements. Although some of this may be due to our measuring instrument, most of it is simply due to the fact that individuals differ. People naturally have different heights, different aptitudes and abilities, and different opinions and emotional responses. When we measure any one of these traits, we are bound to get variability in the measurements. Different seeds for the same variety of bean will grow to different sizes when subjected to the same environment because no two seeds are exactly alike; there is bound to be variability from seed to seed in the measurements of growth.

Induced Variability—If we plant one pack of bean seeds in one field, and another pack of seeds in another location with a different climate, then an observed difference in growth among the seeds in one location with those in the other might be due to inherent differences in the seeds (natural variability), or the observed difference might be due to the fact that the locations are not the same. If one type of fertilizer is used on one field and another type on the other, then observed differences might be due to the difference in fertilizers. For that matter, the observed difference might be due to a factor we haven't even thought about. A more carefully designed experiment can help us determine the effects of different factors. [GAISE, pages 6 & 7]

Statistical problem solving is an investigative process that involves four components:

I. Formulate Questions

- → clarify the problem at hand
- → formulate one (or more) questions that can be answered with data

II. Collect Data

- → design a plan to collect appropriate data
- → employ the plan to collect the data

III. Analyze Data

- → select appropriate graphical and numerical methods
- → use these methods to analyze the data

IV. Interpret Results

- → interpret the analysis
- → relate the interpretation to the original question

[GAISE, page 11]

The *Framework* uses three developmental Levels: A, B, and C. Although these three levels may parallel grade levels, they are based on development in statistical literacy, not age. Thus, a middle-school student who has had no prior experience with statistics will need to begin with Level A concepts and activities before moving to Level B. This holds true for a secondary student as well. If a student hasn't had Level A and B experiences prior to high school, then it is not appropriate for that student to jump into Level C expectations. The learning is more teacher-driven at Level A, but becomes student-driven at Levels B and C. [GAISE, page 13]

Table 1: The Framework

Process Component	Level A	Level B	Level C
I. Formulate Question	**Beginning awareness of the *statistics question distinction*** Teachers pose questions of interest Questions restricted to the classroom	**Increased awareness of the *statistics question distinction*** Students begin to pose their own questions of interest Questions not restricted to the classroom	**Students can make the *statistics question distinction*** Students pose their own questions of interest Questions seek generalization
II. Collect Data	Do not yet *design for differences* Census of classroom Simple experiment	Beginning awareness of *design for differences* Sample surveys; begin to use random selection Comparative experiment; begin to use random allocation	Students make *design for differences* Sampling designs with random selection Experimental designs with randomization
III. Analyze Data	*Use* particular properties of *distributions* in the context of a specific example Display variability within a group Compare individual to individual Compare individual to group Beginning awareness of group to group Observe association between two variables	Learn to *use* particular properties of *distributions* as tools of analysis Quantify variability within a group Compare group to group in displays Acknowledge sampling error Some quantification of association; simple models for association	Understand and *use distributions* in analysis as a global concept Measure variability within a group; measure variability between groups Compare group to group using displays and measures of variability Describe and quantify sampling error Quantification of association; fitting of models for association

[GAISE, page 14]

Process Component	Level A	Level B	Level C
IV. Interpret Results	Students do not look *beyond the data* No generalization beyond the classroom Note difference between two individuals with different conditions Observe association in displays	Students acknowledge that *looking beyond the data* is feasible Acknowledge that a sample may or may not be representative of the larger population Note the difference between two groups with different conditions Aware of distinction between observational study and experiment Note differences in strength of association Basic interepretation of models for association Aware of the distinction between association and cause and effect	Students are able to *look beyond the data* in some contexts Generalize from sample to population Aware of the effect of randomization on the results of experiments Understand the difference between observational studies and experiments Interpret measures of strength of association Interpret models of association Distinguish between conclusions from association studies and experiments
Nature of Variability	Measurement variability Natural variability Induced variability	Sampling variability	Chance variability
Focus on Variability	Variability within a group	Variability within a group and variability between groups Covariability	Variability in model fitting

[GAISE, page 15]

Rubric Considerations

- Interesting question of inquiry
- Appropriate data collection methods, biased recognized/discussed
- Appropriate graphs, neat and labeled
- Appropriate summary statistics used in context
- Statistical ideas clearly communicated in context
- Conclusion clearly communicated in context

Project Ideas

- Compare multiple quantitative groups (boxplots, histograms, etc…)
- Compare two categorical groups (bar charts and pie charts)
- Use one quantitative variable to predict another (regression, scatterplots)
- Conduct a survey, then create a bias
- Compare two treatments and use simulation to test for a statistically significant difference (Algebra 2/Integrated III)

Exploratory Data Analysis project

Step 1

- Create a statistical question of inquiry.
- Your question must involve comparing two or more quantitative variables.
- Do some preliminary research to ensure you can gather the appropriate data.
- Write up your idea and have it approved by me.

Step 2

- Collect your data.
- Required: a minimum of 50 observations per each group being compared.
- Feel free to check with me once your data is collected.

Step 3

- Graph your data—twice. Choices: histograms, boxplots, stem and leaf plots, dotplots, choose two of these.
- Pro-tip: two boxplots count as **one** graph. Make sense?
- Your graphs should be made with the same scale so visual comparisons can be easily made.
- Graphs should be made on statcrunch.com and include labels and scales.
- No posters. Just copy and paste graphs into a word processor and print.

Step 4

- Calculate summary statistics for each group.
- Use your calculator or statcrunch and add to word processing document.

Step 5

- Compare the groups using the summary statistics and the graphs.
- Typed. Along with steps 3 & 4 in one neat document. No cover needed.
- No need to include your raw data with your report.
- Draw conclusions about the differences/similarities between the groups.
- Think carefully about the complete process and make final statement(s) about what you learned about your initial question of inquiry.