

AP Statistics 1-Day Workshop: Experienced

July 22, 2018--Houston, TX

Dave Ferris, Noblesville High School (IN)

Introductions:

The Great American Pastime: Baseball!

2014 Data:

Team	SprWP	RegWP
Arizona	0.48	0.395
Atlanta	0.4	0.488
Baltimore	0.591	0.593
Boston	0.393	0.438
Chicago Cub	0.455	0.451
Chicago Sox	0.391	0.451
Cincinnati	0.452	0.469
Cleveland	0.69	0.525
Colorado	0.517	0.407
Detroit	0.556	0.556
Houston	0.444	0.432
Kansas City	0.429	0.549
LA Angels	0.621	0.605
LA Dodgers	0.389	0.58

<u>2018</u>	Spring Training:	1 st ½ Reg. Season:
Houston Astros:	0.700	0.646
Cincinnati Reds:	0.345	0.448

Overall Question: Is there any relationship between the winning percentages in spring training vs. the regular season for **2014**?

Question #2: What are some of the variables that contribute to each winning percentage?

Question #3: List some ways this data can be analyzed to answer the Overall Question.

Question #4: Would NFL teams' winning percentages in pre- and regular season behave similarly? Explain.

The Walls of Plataea

Measuring Activity

NAME _____



Plataea was an ancient city, located in Greece in southeastern Boeotia, south of Thebes. It was the location of the Battle of Plataea in 479 BC, in which an alliance of Greek city-states defeated the Persians and ended the Persian Wars.

During this battle/siege, a group left the city and escaped over enemy walls by building ladders equal in height to the enemy's walls. The height of the walls was determined by counting the layers of bricks on portions of the wall that they could see. Using several counts by different people, they were able to estimate the height of the wall, and thus the length of the ladders they needed to build.

1. Using the data below, estimate of the number of layers in the wall. Explain your reasoning.

Number of layers from top to bottom:

{20, 22, 23, 20, 22, 22, 21, 24}



2. What other information would be useful to improve the accuracy of the estimate?

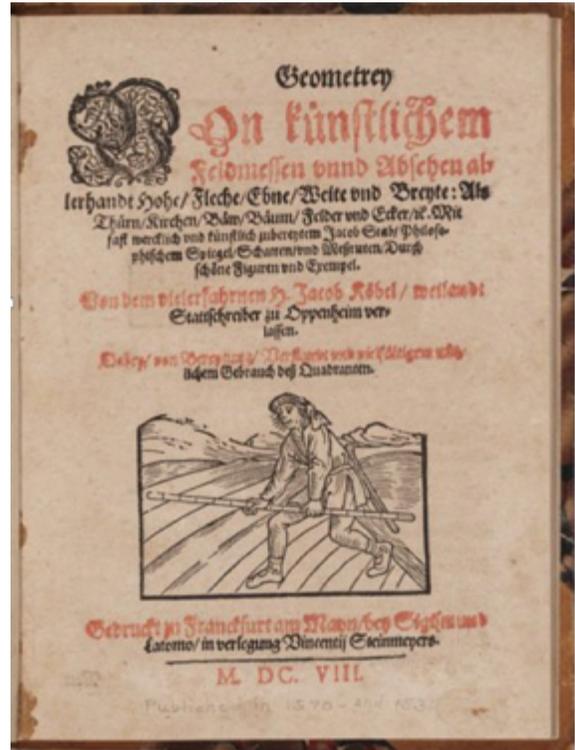
How long is a rood? (rute, rod)

Measuring Activity

NAME _____

In a [book on surveying by Jacob Köbel](#) (1460-1533), he mentions that a surveyor should request that on leaving the church service, 16 men should stop as they come out and stand in a line with their left feet touching the others, heel to toe. Then the length of the 16 feet gives the "right and lawful" rood.

Using Köbel's method and the feet of people present today, calculate an accurate estimate of the length of a rood.



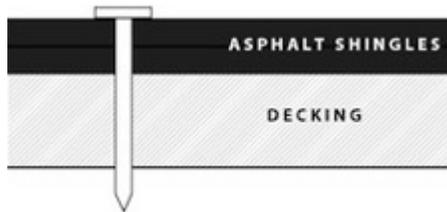
Roofing Nails Problem

True story: I was getting ready to tear off my old roof and put on a new roof, when I wondered what size nails I should use. Roofing nails come in sizes of 1", 1¼", 1½", 1¾", etc. So I decided to measure the length of the nails that I could see from our attic (see picture). The nails shown have penetrated the underlayment (plywood decking), measured at 7/16", as well as two layers of shingles estimated at 0.25". From measuring the length of the protruding nails, I planned to estimate the length of the previously used nails. Then I would know what size to buy for my new shingles.



I measured a convenience sample of 24 nails above the attic stairs under the assumption that the plywood and the shingles were a consistent thickness as well as the lengths of the nails previously used.

The mean length was 13.42mm, which is 0.53 inches. By adding the 7/16" plywood and the 0.25" shingle thickness, I calculated an average nail length of 1.22 inches. So I decided to buy 1¼" roofing nails for my new roof.



1. Comment on some of the potential problems with my methodology, assumptions, and conclusions.

2. What additional information would be needed to ensure an accurate estimate of nail length?

Cannonballs:

Soon to be taking over an enemy's fort, a commanding officer wants to have the correct cannonball size in order to defend the fort with cannons they will soon be seizing. Spy scouts return from behind enemy lines and report the size cannonballs the enemy uses for their cannons. Two sizes are reported with great confidence: 8 inches in diameter and 9 inches in diameter.

Which size should the commander order? Describe a good decision making process using the data.

Thermometers Problem:

You go to your local hardware store to purchase an outdoor thermometer. Upon inspection, you notice that not every thermometer is recording the same temperature. How will you decide which thermometer to purchase? Explain.

\$2.98 Circular



\$1.98 Rectangular



Circular thermometer:

71, 71, 72, 72, 74°

Rectangular thermometer:

73, 73, 74, 74, 77, 77, 77°

The Kristen Gilbert Case

Kristen Gilbert worked as a nurse in the intensive care unit of the VA hospital in Northampton, Massachusetts in the 1990's. During her shifts, other nurses became suspicious that she was killing patients by injecting them with epinephrine, a heart stimulant. An analysis of 1641 eight-hour shifts was presented as evidence at her trial (Cobb and Gelbach, 2005¹). Is there statistical evidence from this table that Gilbert murdered patients?

		Death on shift?		
Gilbert Present?	Yes	No	Total	
Yes	40	217	257	
No	34	1350	1384	
Total:	74	1567	1641	

¹ <http://www.stat.ucla.edu/~nchristo/statistics100B/article.pdf>

Can Facebook Identify Terrorists?²

Suppose Facebook has developed an algorithm that can identify Facebook users who are twice as likely to be involved in terrorist activities as the typical Facebook user. Your neighbor has been identified as one of the 100,000 Facebook users that fit the profile of a terrorist. Assuming there are 200 million Facebook users, 10,000 of which really are terrorists, should you be worried about your neighbor? Explain.³

² Ellenberg, Jordan (2014-05-29). *How Not to Be Wrong: The Power of Mathematical Thinking* (p. 167). Penguin Group US. Kindle Edition.

³ Two questions worth considering: 1) What is the likelihood that your neighbor would make the list if, in fact, he is not a terrorist? 2) What is the likelihood that your neighbor is not a terrorist given that he is on the list?

Selected Student Samples from 2018 Exam

1.

- (a) Identify and interpret in context the estimate of the intercept for the least-squares regression line.

~~When~~

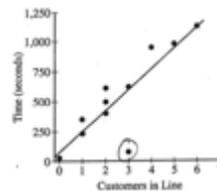
When there are 0 customers in line, the time it takes to finish checkout is 72.95 seconds.

- (b) Identify and interpret in context the coefficient of determination, r^2 .

r^2 is given by the output as 70.37%, or 0.7037. This indicates that approximately 70.37% of the variance of results from the expected times (regression line) is accounted for by the least-squares regression line.

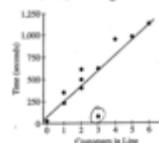
- (c) One of the data points was determined to be an outlier. Circle the point on the scatterplot and explain why the point is considered an outlier.

This point is considered to be an outlier because it has a very large, negative residual. It's also an influential point meaning it would drastically change the slope of the regression line.



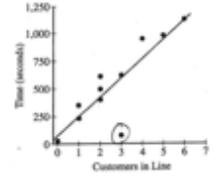
- (c) One of the data points was determined to be an outlier. Circle the point on the scatterplot and explain why the point is considered an outlier.

The point is an outlier because its y value is so far off from its \hat{y} value.



- (c) One of the data points was determined to be an outlier. Circle the point on the scatterplot and explain why the point is considered an outlier.

This is considered an outlier because it deviates from the linear pattern. The regression line doesn't pass through or run near this data point so it is an outlier.



2.

- (b) Given the method used by the environmental science teacher to collect the responses, explain how bias might have been introduced and describe how the bias might affect the point estimate of the proportion of all students at the school who would respond yes to the question.

There is bias in this sample because students do not want the teacher to know that they do not recycle. Therefore, the proportion of the sample of student who regularly recycle plastic bottles is higher than the true proportion.

- (b) Given the method used by the environmental science teacher to collect the responses, explain how bias might have been introduced and describe how the bias might affect the point estimate of the proportion of all students at the school who would respond yes to the question.

There could be bias when the environmental science teacher (who's in favor of recycling) asks the students if they recycle. The student may be embarrassed to say no or pressured to say yes in fear of the teacher judging or disapproving of them. This is an example of response bias. This could cause an over prediction of the population proportion of students at that school who recycle.

3.

- (c) A random sample of 20 children born in the region will be selected. What is the probability that the sample will have at least 3 children who are left-handed?

$$P(L \geq 3) = 1 - P(L < 3) = \text{binomcdf}(20, .067, 2) \quad \begin{matrix} \text{trials} & \text{prob} & X \\ & & \end{matrix}$$

$$= 2.26 \times 10^{-7}$$

The probability that out of 20 children in the region, at least three are left handed is 2.26×10^{-7} .

- (c) A random sample of 20 children born in the region will be selected. What is the probability that the sample will have at least 3 children who are left-handed?

$X = \text{number of left-handed children in 20 child sample}$

$N(.11385, .071)$



$$V_{\hat{p}} = p = .11385$$

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{.11385(1-.11385)}{20}} = .071$$

$$P(X \geq 3) = \text{normalcdf}(.15, \infty, .11385, .071) \approx .305$$

large enough

- (c) A random sample of 20 children born in the region will be selected. What is the probability that the sample will have at least 3 children who are left-handed?

$$p(x \geq 3) = 1 - p(x \leq 2)$$

$$p(x \leq 2) = \text{binomcdf}(20, .11385, 2)$$

$$p(x \leq 2) = .59785$$

$$p(x \geq 3) = 1 - .59785$$

$$p(x \geq 3) = .40215$$

Binary ✓
Independent ✓
Num trials ✓
Success probability ✓

4.
(a) Based on the design of the study, would a statistically significant result allow the medical center to conclude that the new procedure causes a reduction in recovery time compared to the standard procedure, for patients similar to those in the study? Explain your answer.

Yes because the study done was a good experiment that satisfied control (standard), randomization (procedures randomly assigned), and replication (210 patients to reduce variability). As a result, a good experiment means that correlation is equal to causation.

Do the data provide convincing statistical evidence that those who receive the new procedure will have less recovery time from the surgery, on average, than those who receive the standard procedure, for patients similar to those in the study?

IO: 2 sample t-test

Ho: $\mu_{\text{new}} = \mu_{\text{standard}}$

HA: $\mu_{\text{new}} < \mu_{\text{standard}}$

let μ_{new} = true mean recovery time for new procedure

let μ_{standard} = true mean recovery time for standard procedure

Conditions:

- Random sampling
- At least 1100 patients who took standard procedure; at least 1000 patients who took new procedure
- Independence between standard and new procedure

If you need more room for your work in part (b), use the space below.

- Both standard procedure distribution and new procedure distribution are approx normal because sample sizes are greater than 40 for each

$$(110 > 40) \text{ and } (100 > 40)$$

$$t = \frac{(186 - 217) - 0}{\sqrt{\frac{34^2}{110} + \frac{29^2}{100}}} \rightarrow t = \frac{-31}{4.35} \rightarrow t = -7.127$$

$$df = 207.129$$

$$t_{cdf}(-10, -7.127, 207.129) = 8.358 \times 10^{-12}$$

Because p-value of 8.358×10^{-12} is less than .01, we can reject the null hypothesis at the 1% level. There is strong evidence that the new procedure has a shorter mean recovery time than the standard procedure.

4. The anterior cruciate ligament (ACL) is one of the ligaments that help stabilize the knee. Surgery is often recommended if the ACL is completely torn, and recovery time from the surgery can be lengthy. A medical center developed a new surgical procedure designed to reduce the average recovery time from the surgery. To test the effectiveness of the new procedure, a study was conducted in which 210 patients needing surgery to repair a torn ACL were randomly assigned to receive either the standard procedure or the new procedure.
- (a) Based on the design of the study, would a statistically significant result allow the medical center to conclude that the new procedure causes a reduction in recovery time compared to the standard procedure, for patients similar to those in the study? Explain your answer.

Yes because with a statistically significant result the medical center will be able to see which procedure works better.

Do the data provide convincing statistical evidence that those who receive the new procedure will have less recovery time from the surgery, on average, than those who receive the standard procedure, for patients similar to those in the study?

H_0 : The data does not provide statistical evidence that the new procedure will have less recovery time.

H_a : The data does provide statistical evidence that the new procedure will have less recovery

2-sample t-test

① We have 210 randomly assigned patients.

②

$$t = 7.12707882$$

$$p\text{-value} = 8.358 \times 10^{-12}$$

$$df = 207.174044$$

If you need more room for your work in part (b), use the space below.

Since the p-value is less than .05 or α we reject H_0 ; there is enough evidence to prove that the new procedure will reduce recovery time after the surgery

5.

- (a) The median teaching year for one high school is 6, and the median teaching year for the other high school is 7. Identify which high school has each median and justify your answer.

High School A = median of 7

High School B = median of 6

High School B is skewed right and has a much higher frequency of teachers who have a smaller teaching year.

This would make the median of high school B be less than that of high school A. High School B's shape is unimodal and skewed right while A is not.

- (a) The median teaching year for one high school is 6, and the median teaching year for the other high school is 7. Identify which high school has each median and justify your answer.

The median teaching year for High School A is 7 years and the median teaching year for high school B is 6 years. When calculated, the median year for A falls between 7 and 10 and the median year for B falls between 4 and 7.

- (c) The standard deviation of the teaching year for the 221 teachers at High School B is 7.2. If one teacher is selected at random from High School B, what is the probability that the teaching year for the selected teacher will be within 1 standard deviation of the mean of 8.2? Justify your answer.

This means the teaching years will be 8.2 ± 7.2 , (1, 15.4)

$$\text{Normalcdf}(1, 15.4, 8.2, 7.2) = .6827$$

↓ ↓ ↓ ↓
LB UB μ σ

$P(\text{within 1 standard deviation}) = .6827$

6.

The actual mean systolic blood pressure of all employees at the corporation is 125 mmHg, not the hypothesized value of 122 mmHg, and the standard deviation is 15 mmHg.

(c) Using the actual mean of 125 mmHg and the results from part (b), determine the probability that the null hypothesis will be rejected.

$$p = \text{Ncdf}(124.5, -\infty, 125, 1.5) = \boxed{.631}$$

(e) Suppose the size of the sample of employees to be selected is greater than 100. Would the probability of rejecting the null hypothesis be greater than, less than, or equal to the probability calculated in part (c)? Explain your reasoning.

It would be greater because increasing sample size decreases α and β , thereby increasing power, which is equal to $1 - \beta$, so the power increases. Alternatively, increasing sample size would decrease the standard error, meaning the actual mean has an even greater z-score and the likelihood of rejection of the null hypothesis increases because the p-value is even smaller.

Desmos demo of Errors and Power using THIS problem!

desmos.com/calculator/onat3s3fyd

Research Summary:

“...researchers often find that few students properly interpret their results or draw valid conclusions based on the statistical results they obtain. They seem to complete procedures mechanically, without developing statistical thinking.”

(Ben-Zvi et al)

What is Statistical Thinking and How is It Developed?
by Sharon J. Lane-Getaz
©2006, NCTM's Sixty-Eighth Yearbook:
Thinking and Reasoning with Data and Chance

Specific Suggestions:

- Transition students from “local view” to “global view” of data.
- Develop the idea of a sample.
- Encourage the need for a statistical term before formal introduction
- Progress from small to large data; then continuous distributions.

General Implications for Teachers:

- Sequence activities carefully
- Focus on the “big ideas”
- Use good, real data
- Encourage speculation and exploration
- Encourage making and testing conjectures
- Use appropriate technology
- Discuss and reflect in class

Research in the Statistics Classroom: Learning from Teaching Experiments
by Dani Ben-Zvi, Joan B. Garfield and Andrew Zieffler
©2006, NCTM's Sixty-Eighth Yearbook:
Thinking and Reasoning with Data and Chance

Garfield and Ben-Zvi's Eight Learning Principles:⁴

1. Students learn by constructing knowledge.
2. Students learn by active involvement in learning activities.
3. Students learn to do well only what they practice doing.
4. It is easy to underestimate the difficulty students have in understanding basic concepts of probability and statistics.
5. It is easy to overestimate how well students understand basic concepts.
6. Learning is enhanced by having students become aware of and confront their errors in reasoning.
7. Technological tools should be used to help students visualize and explore data, not just to follow algorithms to pre-determined ends.
8. Students learn better if they receive consistent and helpful feedback on their performance.

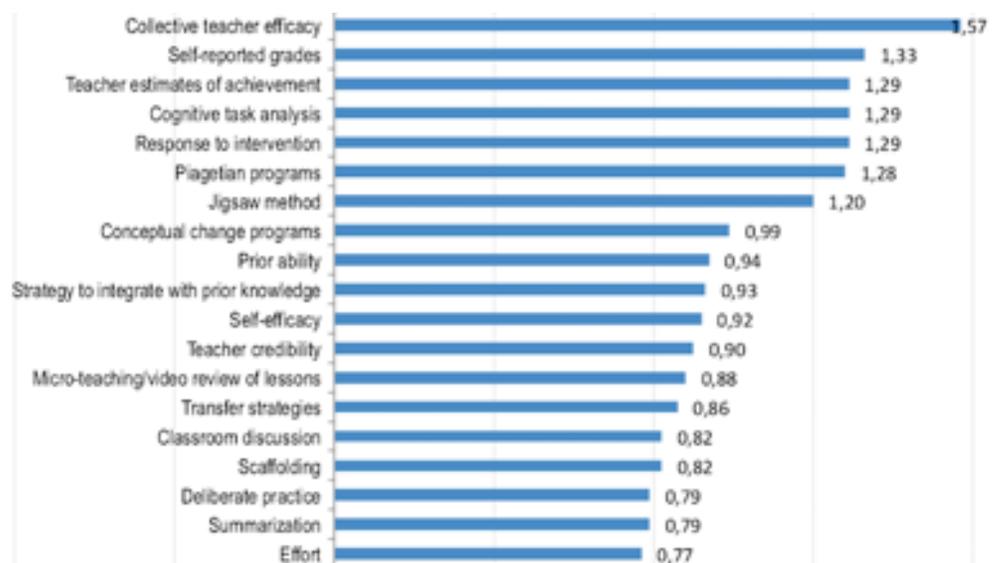
Marzano's Keys to Student Engagement--Students must answer in the affirmative:

How do I feel? Am I interested?
[Getting their attention.]



Is it important? Can I do it?
[Getting them engaged.]

John Hattie's Visible Learning Research:



⁴ Garfield, Joan, and Dani Ben-Zvi. "How Students Learn Statistics Revisited: A Current Review of Research on Teaching and Learning Statistics." *International Statistical Review* 75.3 (2007): 372-96

Mathematics Teaching Practices

Establish mathematics goals to focus learning. Effective teaching of mathematics establishes clear goals for the mathematics that students are learning, situates goals within learning progressions, and uses the goals to guide instructional decisions.

Implement tasks that promote reasoning and problem solving. Effective teaching of mathematics engages students in solving and discussing tasks that promote mathematical reasoning and problem solving and allow multiple entry points and varied solution strategies.

Use and connect mathematical representations. Effective teaching of mathematics engages students in making connections among mathematical representations to deepen understanding of mathematics concepts and procedures and as tools for problem solving.

Facilitate meaningful mathematical discourse. Effective teaching of mathematics facilitates discourse among students to build shared understanding of mathematical ideas by analyzing and comparing student approaches and arguments.

Pose purposeful questions. Effective teaching of mathematics uses purposeful questions to assess and advance students' reasoning and sense making about important mathematical ideas and relationships.

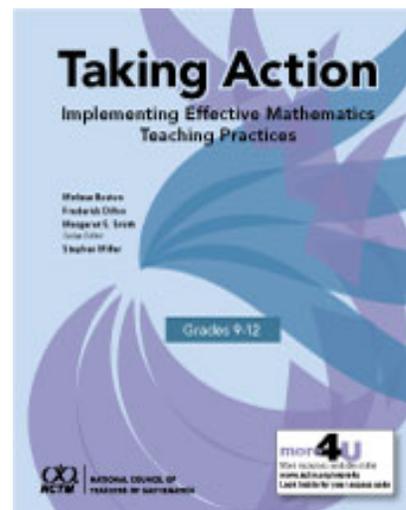
Build procedural fluency from conceptual understanding. Effective teaching of mathematics builds fluency with procedures on a foundation of conceptual understanding so that students, over time, become skillful in using procedures flexibly as they solve contextual and mathematical problems.

Support productive struggle in learning mathematics. Effective teaching of mathematics consistently provides students, individually and collectively, with opportunities and supports to engage in productive struggle as they grapple with mathematical ideas and relationships.

Elicit and use evidence of student thinking. Effective teaching of mathematics uses evidence of student thinking to assess progress toward mathematical understanding and to adjust instruction continually in ways that support and extend learning.

From NCTM's *Taking Action* series...

The eight "effective mathematics teaching practices"...are intended to guide and focus the teaching of mathematics across grade levels and content areas. ***Decades of empirical research in mathematics classrooms support these teaching practices.***



Walking the Hall Activity

Thinking about and collecting data

NAME _____

1. If we all walked the same short distance in the hall (less than 30 seconds) and recorded the number of steps we took, predict what the graph (dotplot) will look like. Justify your answer.

2. If we all walked the same LONG distance in the hall (more than one minute) and recorded the number of steps we took, predict what the graph (dotplot) will look like. Justify your answer.

3. What if we created a scatterplot of everyone's short distance step count vs. their long distance step count. What would this look like? Justify your answer.

4. Let's go walk the halls! Two distances are already marked out, so use your normal walking pace and count the number of steps for both distances.

Short: _____ steps

Long: _____ steps

Record these data pairs on the board or poster paper. Also include a column for each person's initials.

Create graphs to see if your predictions were correct!

5. What are some of the features of the dotplots of steps? Was there anything unexpected?

What would account for the variation in the number of steps?

6. Describe the association between number of steps in the short walk vs. the number of steps in the long walk (look at a scatterplot). Discuss anything unusual that appears.

7. Now we will measure each person's approximate leg length and compare to the step counts of each walk. What should we expect to see? Which one will have the strongest relationship? Explain.

Collect this data and add it to the previously collected step data.

Create two scatterplots, using leg length as the explanatory variable for each.

Compare what the scatterplots reveal compared to the predictions. Discuss any anomalies or unexpected results.

Probability Activities:

Coinless Cokes Activity:

A vending machine has been programmed to only take \$1 bills. Sometimes a customer will get a free Coke, and sometimes they pay \$1. Over time, Cokes purchased will have an average cost of less than \$1.

Use the TI calculator to determine "FREE" or "Pay Dollar," then enter a "1" or a "0" into Fathom. **We want to try to determine the average cost of a Coke in the long run.** This will demonstrate an important principle of probability.

Coin Flip Walking Activity

1. Flip a fair coin 30 times and record the results.
2. Stand side-by-side in a straight line.
3. On teacher's command, take a step:
Heads = forward
Tails = backward
4. What will the line look like after 20 steps? 100? 1000? 1,000,000?



Casino War Activity



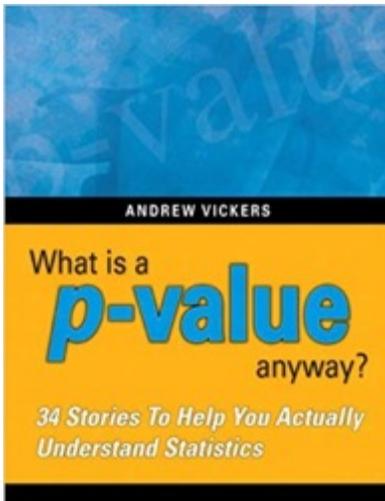
1. Shuffle two decks.
2. Compare cards one at a time.
3. What is the probability of an EXACT match sometime before both decks are exhausted?
4. Try the theoretical approach first.
5. Then get two decks and simulate it!
6. Can a computer simulate this?
7. Back to theoretical...a surprise?

Tommy John and Errors

Famous pitcher Tommy John once made three errors on a single play: he bobbled a grounder, threw wildly past first base, then cut off the relay throw from right field and threw past the catcher.

In a scientific paper describing a clinical trial comparing a new pain drug with a placebo, the authors wrote something like this: ***“Although there was no difference in baseline age between the groups ($p = 0.458$), controls were significantly more likely to be male ($p = 0.000$).”***

This statement is worse than Tommy John’s worst day because there are actually four errors in this sentence (or maybe even 4.5 errors). See if you can find them.



From *What is a P-value, Anyway?*
By Andrew Vickers

Quizlet Live Game (good for practicing terms):

https://quizlet.com/_2no4wj

SOLUTIONS to Selected Student Samples from 2018 Exam:

1a) Student E: Partially correct. (No mention of “predicted.”)

Part (a) is scored as follows:

Essentially correct (E) if the response satisfies the following three components:

1. Correctly identifies 72.95 as the intercept.
2. Communicates the concept of a y intercept AND in a context that includes both time and zero customers.
3. Includes wording to indicate that the value of the intercept is a prediction, such as “predicted”, “estimated”, or “average” value of y .

Partially correct (P) if the response satisfies only two of the three components required for E.

Incorrect (I) if the response satisfies at most one of the components required for E.

1b) Student F: Incorrect. (Wrong value and wrong variable referenced. Should have said “observed times” instead of “expected times.”)

Part (b) is scored as follows:

Essentially correct (E) if the response satisfies the following three components:

1. Correctly identifies 73.33% as the coefficient of determination.
2. Provides a correct (possibly generic) interpretation of r^2 .
3. Interpretation includes context.

**Partially correct (P) if the response satisfies only two of the three components required for E;
OR**

if the response satisfies all three components, but reverses the roles of number of customers in line and time to finish checkout in the definition.

Incorrect (I) if the response satisfies at most one of the components required for E.

1c #1) Student D: Partially correct (explanation has no comparison to other points)

Essentially correct (E) if the response satisfies the following two components:

1. Correctly identifies the outlier.
2. Describes an unusual feature of the identified scatter plot point, relative to the remaining data points, that is sufficient to identify it as the outlier. Examples include:
 - The combination of x and y values is unusual compared to the other points.
 - The value of y is much lower than would be expected (or predicted), given the remaining data.
 - The residual for the point is unusually large relative to the other residuals.

Partially correct (P) if the response satisfies component 1 but does not satisfy component 2.

Incorrect (I) if the response does not meet the criteria for E or P.

1c #2) Student C: Partially correct (explanation has no comparison to other points)

1c #3) Student G: Essentially correct (compares outlier to “the linear pattern,” which suggests a comparison to remaining points. Communication could have been more clear, but the essential comparison to the rest of the data was deemed present.)

2b #1) Student C: Partially correct (no mention of HOW responses might differ from the truth)

Essentially correct (E) if the response includes the following three components:

1. Explains why the responses to the survey might differ from the truth about student recycling in this context (e.g., the survey was not anonymous, the question was asked by an authority figure)
2. Explains how the responses to the survey might differ from the truth about student recycling (e.g., “students might say yes when they actually don’t recycle”, “students lie and say yes”, “students don’t recycle but lie to the teacher”)
3. Describes the effect of the bias on the point estimate (or the proportion, percentage, number of “yes” responses in the sample) and doesn’t contradict the bias described.

Partially correct (P) if the response includes only two of the components required for E.

2b #2) Student F: Partially correct (Again...no mention of HOW responses might differ from what students really believe.)

3c #1) Student O: Partially correct (wrong probability; wrong answer based on their work... they forgot to subtract from 1)

Part (c)

Essentially correct (E) if the response includes the following components:

1. Uses a calculation based on the binomial distribution to find the probability of the number of children in the sample who are left-handed.
2. Specifies appropriate values for n and p .
3. Uses correct endpoint value for the probability.
4. Uses correct direction to calculate the probability of at least 3 left-handed children.
5. Correctly calculates a binomial probability consistent with the previous work.

Partially correct (P) if the response includes component 1 and two or three of the other components required for E;

OR

if components 2, 3, 4, and 5 are met and the response does not explicitly indicate the binomial distribution is used by name or formula.

3c #2) Student L: Partially correct (used Normal model instead of Binomial, but parts 2-5 are all correct based on their work)

- A normal approximation to the binomial is not appropriate since $np = 20 \times 0.11385 = 2.277 < 5$. A response using the normal approximation can score at most P. To score P, the response must include all of the following:
 - a correct mean and standard deviation based on the binomial parameters
 - clear indication of boundary and direction with a z-score or diagram
 - the probability computed correctly.

3c #3) Student L: Partially correct (n and p are not labeled)

4. #1) Student D: Earned a holistic score of 3 (EEE, but no conditions)

4a #2 b) Student I: IPP = 1

Section 1 is Incorrect (see below)

- Section 1 is scored in correct (I). Component 1 is satisfied because the response states that a causal conclusion can be made. Component 2 is not satisfied because the response does not provide justification for the causal relationship. Component 3 is satisfied because the response is in context. In order for section 1 to be scored partially correct (P), component 1 must be satisfied and weak justification must be provided according to the examples in the rubric. The response has not provided weak justification; therefore, section 1 is scored incorrect (I).

Section 2 is Partially correct (parameters are not clearly defined)

Section 3 is Partially correct (student said “prove” in conclusion)

5a #1) Student G: Partially correct (no justification of why skewness identifies medians)

Essentially correct (E) if the response includes the following 3 components:

1. States that the median is 6 for High School B and the median is 7 for High School A.
2. Provides a reasonable explanation of how the decision was made.
3. Provides the definition of the median or explicitly applies the definition of a median as a criterion in reaching their decision.

OR

Essentially correct (E) if the response includes the following 3 components:

1. States that the median is 6 for High School B and the median is 7 for High School A.
2. States that High School B shows a skewed distribution (or High School A shows a less skewed distribution).
3. Provides a reasonable explanation of how the more skewed distribution (High School B) would be the one with a larger separation between the mean and median.

Partially correct (P) if the response includes the first component and one of the other two components required for E.

5a #2) Student H: Incorrect (no explanation of decision; no definition of median)

5c) Student Y: Incorrect (used a Normal distribution)

6c) Student E: Partially correct (no identification of parameters in “calculator speak”)

Essentially correct (E) if the response includes the following three components:

1. Recognizes that the null hypothesis will be rejected when $\bar{x} \geq 124.4675$, as found in part (b).
2. Provides the correct sampling distribution for the sample mean when the true mean is 125, including correct values for the mean and standard deviation, either explicitly or by plugging them into the test statistic formula.
3. Provides evidence of using the normal curve and finds the correct probability value.

Note: Components 1 and 3 can still be satisfied if errors made in finding the rejection region in part (b) are carried into part (c).

Note: Calculator speak without labeling input values does not satisfy component 2 but may still satisfy components 1 and 3.

6e) Student E: Partially correct (Component 4: no mention of minimum \bar{x} -bar decreasing)

Essentially correct (E) if the response includes the following 4 components:

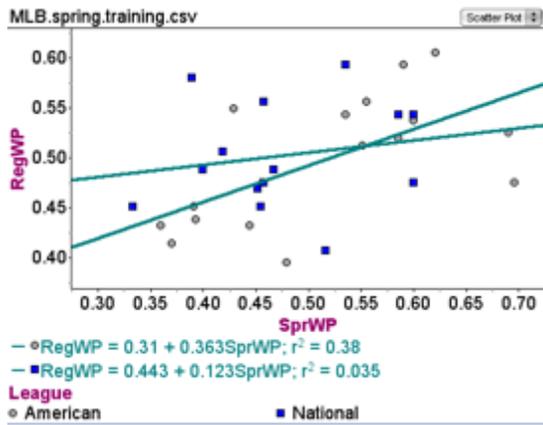
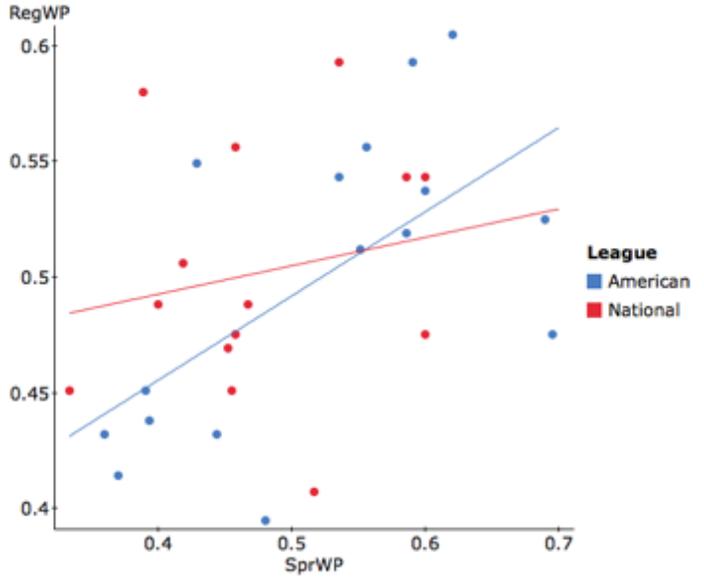
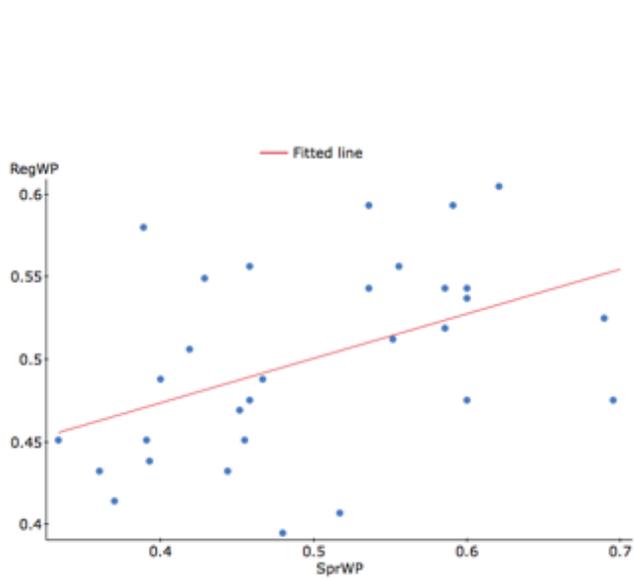
1. Part (d) specifies “power” as the name of the probability.
2. Part (e) correctly states that the probability would be greater.
3. Part (e) correctly implies that the standard deviation of the sampling distribution decreases, either explicitly or by plugging values into a formula.
4. Part (e) indicates the minimum value of \bar{x} for which the null hypothesis is rejected decreases, either explicitly or by plugging values into a formula.

Note: Component 4 can still be satisfied if a response indicates that a maximum value of \bar{x} for which the null hypothesis is rejected increases if this direction is consistent with answers in parts (b) and (c).

Note: Student E earned EPP = 2 for the two mistakes above.

Other Solutions/Notes:

MLB Spring Training vs. Regular Season:

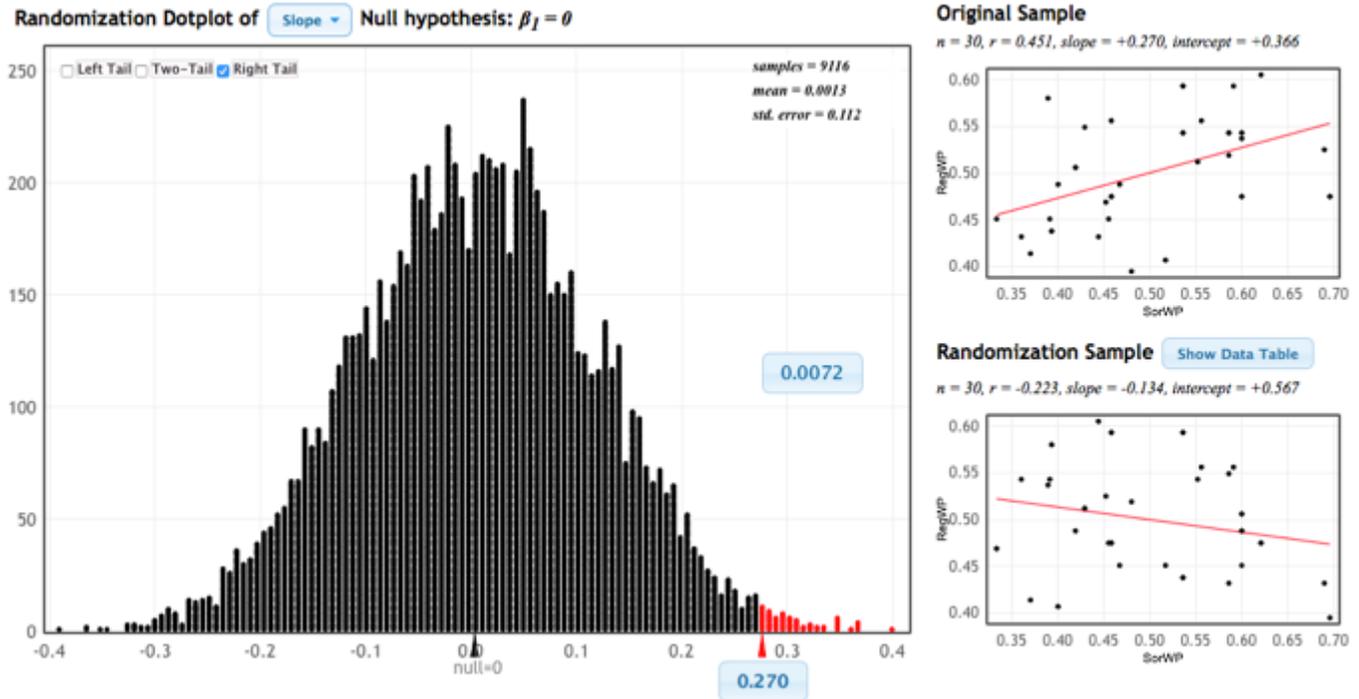


American League:	National League:
Correlation = 0.617	Correlation = 0.186
R-sq = 0.381	R-sq = 0.035

Using StatKey (free online tool) to conduct a randomization test:

Choose a Randomization Hypothesis Tests for slope. Select "Edit" and enter MLB data. Change the option of Correlation to Slope.

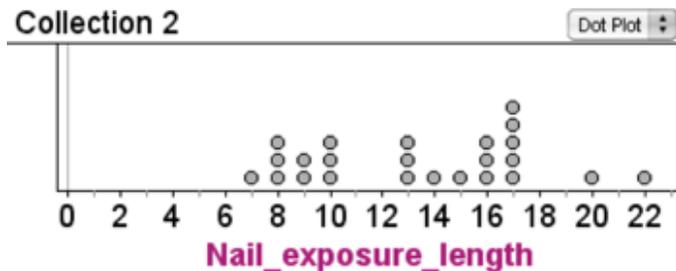
Simulate 1000's of re-randomizations. Click on "Right Tail." Change the boundary to the sample slope (0.270).



Note that after 1000's of re-randomizations, if there is really no association between these two seasons, the probability of getting a slope of 0.27 or higher due simply to random chance, is less than 1%. Therefore, we reject the null hypothesis that there is no association and conclude that there IS a positive association, although it is rather weak with $r = 0.451$.

Roofing Nails Problem:

Dotplot of Data: Notice the shape and the location of the centers of the two modes.
 (Hint: convert to inches and add $7/16''$ and $0.25''$)
 The previous roofer must have used **TWO** sizes of nails!



Facebook Terrorist Problem:

Draw a 2x2 table...

	On "the list"	Not on list	
Terrorists	10	9,990	10,000
Non terrorists	99,990	199,890,010	199,990,000
	100,000	199,900,000	200,000,000 Facebook users

You find out that your neighbor is on that list.
 Should you be worried? Explain.

ANSWER: No. Given he is on "the list," there is a 99.99% chance your neighbor is not a terrorist.

Solutions/Hints to Walking the Hall Activity:

Let students THINK by themselves first...this is good for the introverts...

Then have them DISCUSS with a partner/group...

Have a whole-class discourse following these initial student discussions...

Use the principles of good discourse!

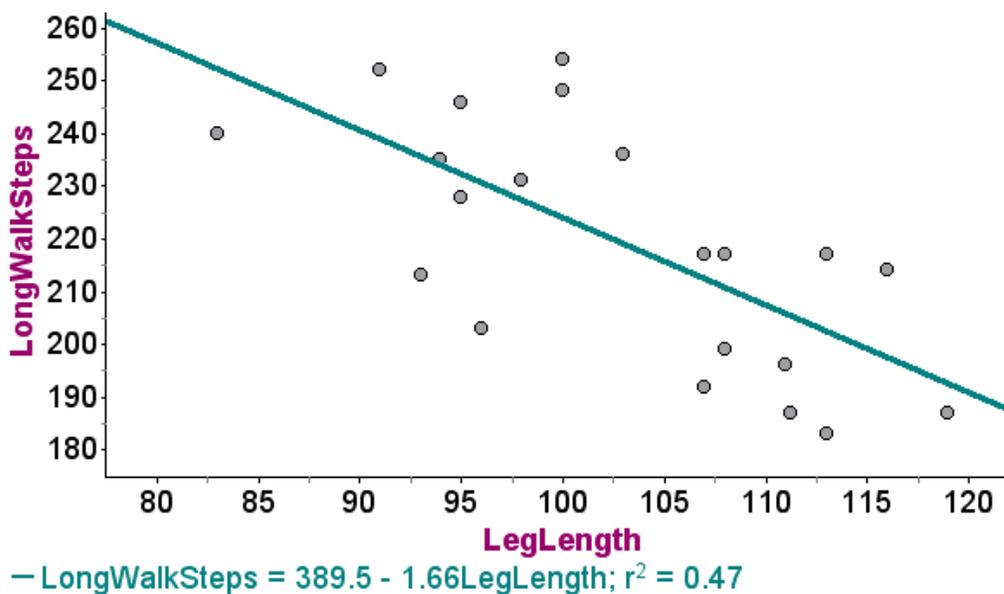
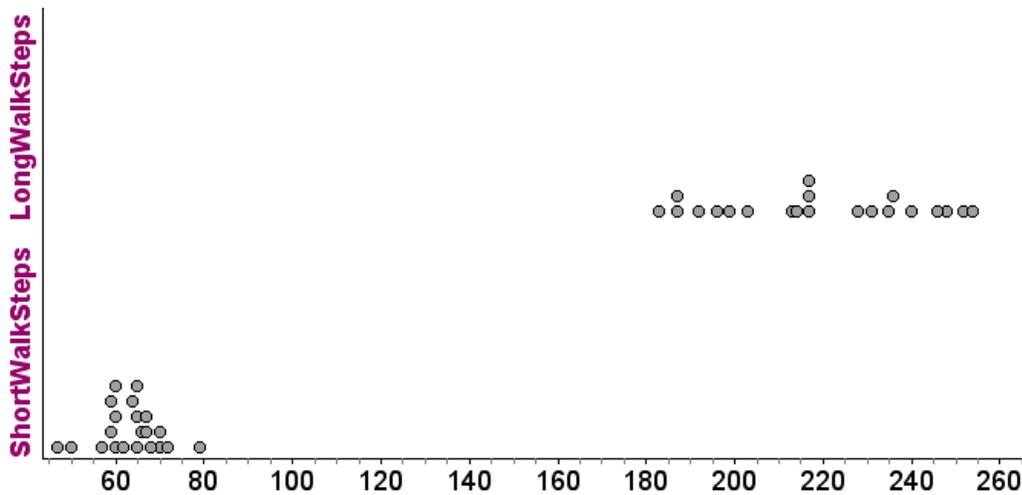
NOTES: Some of the features of the dotplots will be less predictable than others...

How do you best measure leg length? Is there another body measurement that might have a greater association with number of steps?

The weak relationship between leg length and step count lends itself to a discussion of what other variables might impact step count. Students usually identify that different people have different gaits (though they may not use the term gait, which means "stride length").

Gait/stride analysis can be used to assess deviations from normal, especially if a person's baseline gait has changed due to an injury, students got distracted during the walk, they wore new shoes, etc.

Sample graphs:



Answers to Tommy John problem:

1. Accepting the null hypothesis
 2. Giving a p-value for baseline differences between random groups (p-values test hypotheses).
 3. Inappropriate levels of precision (what do the 5 and 8 tell us?)
 4. Reporting a p-value = 0.
- 4½: Why were they measuring baseline ages anyway? All patients will be in the trial over the same time period.

Web Site List

RESOURCES, HANDOUTS, ACTIVITIES:

noblestatman.com (“**ehandouts**” and other resources)

AP Central (official **documents** and AP **Exam problems**)

apstatsmonkey.com (**clearinghouse** for many useful resources, including Best Practices Night at the AP Reading, FRAPPY’s, etc.)

amstat.org (American Statistical Association)

(**STEW lesson plans** for activities and mini-projects; GAISE report: guidelines for statistical instruction)

TOOLS:

stapplet.com (online “calculator” for all computations and inference procedures)

Against All Odds statistics **videos** (can stream for free--learner.org)

amstat.org (Census At School: **survey** and **student data**)

Rossman Chance **applets** (many good simulation applets)

onlinestatbook.com/stat_sim/sampling_dist/index.html (great sampling distribution demonstration **applet**)

getkahoot.com, quizlet.com, quizizz.com (engaging online review **games**)

StatCrunch (Stats **Software**: teacher account is free, students pay small fee)

StatKey (**simulation** website app)

tuvalabs.com (online **tool** for analyzing distributions and scatterplots)

gapminder.org (amazing online analysis **tool** of United Nations data)

Classifying Statistics Problems (Itconline—**practice** at choosing the correct inference procedure)

ARTICLES and NEWS:

fivethirtyeight.com (great current **articles** with a statistical slant; engaging graphs)

www.causeweb.org/sbi (Simulation Based Inference discussions/**blog**. This is a “trending” topic among high school and college statistics teachers.)

thisisstatistics.org (engaging information on statistics as a **career**)

tylervigen.com/spurious-correlations (funny, non-causation relationships)