Minicourse #5: Teaching Introductory Statistics
Joint Mathematics Meetings 2016 – Seattle

Session 1: Wednesday, January 6, 9:00 a.m.–11:00 a.m.
Session 2: Friday, January 8, 9:00 – 11:00 am
Metropolitan A, 3rd Floor, Sheraton Seattle Hotel

Carolyn Cuff, Westminster College ccuff@westminster.edu

This minicourse, intended for instructors new to teaching statistics, exposes participants to the big ideas of statistics and the ASA-endorsed Guidelines for Assessment and Instruction in Statistics Education (GAISE) report. It considers ways to engage students in statistical literacy and thinking, and contrast conceptual and procedural understanding in the first statistics course. Participants will engage in many of the classic activities that all statistics instructors should know. A set of approximately 6 - 8 hands-on classroom-ready activities will be given to participants. Parts of each activity will be done by the participants, other parts will be summarized by the presenter and the main statistical ideas of the activity will be explained to the participants. The activities have been chosen so that they require minimal adaptation for a wide variety of classrooms and are easy to implement. Each activity includes goals, key ideas, prerequisite skills and concepts, connection to other statistical concepts, objectives, known student difficulties and assessment questions. Internet sources of real data, activities, and best practices articles will be examined. Participants will find out how they can continue to learn about the best practices for the first course in Statistics by becoming involved in statistics education related conferences, newsletters, and groups.

SCHEDULE

Day One: Wednesday, January 6th

9:00 - 9:25	Introductions to Workshop
9:25 - 10:00	Data Production and Exploratory Data Analysis
10:00-10:40	Sampling
10:40-11:00	Experimental and Observational Studies: Cause and Effect

Day Two: Friday, January 8th

9:00 - 9:10	Questions from Session 1
9:10 – 9:45	Sampling Distributions and the Central Limit Theorem
9:45 - 10:20	Confidence Intervals
10:20 - 10:45	Hypothesis Testing
10:35 - 11:00	Summary, Discussion and Resources

Why Statistics?

1. Art Benjamin's TED talk (2009) about statistics and probability as the pinnacle of mathematics education: http://www.ted.com/talks/arthur_benjamin_s_formula_for_changing_math_education.html

2. McKinsey Report (2011) on "Big Data: The Next Frontier for Innovation, Competition, and Productivity" http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation)

3. "By 2018, the United States alone could face a shortage of 140,000 to 190,000 people with deep analytical skills as well as 1.5 million managers and analysts with the know-how to use the analysis of big data to make effective decisions"NYTimes article (2009) "For Today's Graduate, Just One Word: Statistics" (http://www.nytimes.com/2009/08/06/technology/06stats.html).

4. Hal Varian, Chief Economist at Google, says "…the sexy job in the next 10 years will be statisticians. And I'm not kidding!"

5. Common Core State Standards push statistics, probability, and data analysis earlier into the curriculum (http://www.corestandards.org/Math/Content/HSS/introduction/)

6. Data Scientist is one of the best jobs in engineering and IT, 15% job growth predicted, background in statistics needed, Forbes, Oct. 29, 2015, (http://www.forbes.com/sites/susanadams/2015/10/29/the-best-jobs-in-engineering-and-it/)

In this workshop, we cannot cover all the content in the introductory statistics class nor can we show all the activities or examples that you might use. However, we hope to introduce you to how a modern statistics course could be run.

Key concepts for a first course in statistics

- Data vary – If we correctly sample from a population, data usually vary in a predictable pattern allowing us to infer information about that population.
- Correlation does not imply causation.
- Conclusions drawn from observational studies and experimental studies are different. Observational studies have potential confounding variable; good experiments control for confounding variables. Experiments are not always possible or ethical.

GAISE: Guidelines for Assessment and Instruction in Statistics Education
http://www.amstat.org/education/gaise/

The GAISE report consists of two reports - PreK-12 and College. Endorsed by ASA (American Statistical Association) in 2005, it provides guidance for statistical education.

The focus of **PreK-12** is to develop statistical literacy. The premise is that foundational statistical concepts should be introduced and nurtured in the elementary grades and then strengthened and expanded throughout the middle, high school, and postsecondary grades. Emphasis is placed on the role of variability and the importance of context. The Pre-K-12 report was published in book form in 2007. Process components for the three levels of mastery can be summarized as follows:

Process Component	Level A	Level B	Level C
Formulate Question	Beginning awareness of the *statistics question distinction*	Increased awareness of the *statistics question distinction*	Students can make the *statistics question distinction*
Collect Data	Do not yet *design for differences*	Awareness of *design for differences*	Students make *designs for differences*
Analyze Data	*Use* particular properties of *distributions* in context of specific example	Learn to *use* particular properties of *distributions* as tools of analysis	Understand and *use distributions* in analysis as a global concept
Interpret Results	Do not *look beyond the data*	Acknowledge that *looking beyond the data* is feasible	Able to *look beyond the data* in some contexts

For those who teach K-8 or train K-8 teachers, Bridging the Gap Between Common Core State Standards and Teaching Statistics (http://www.amstat.org/education/btg/index.cfm) will be helpful. It provides 20 data analysis and probability investigations consist with the CCSS.

The **college** report includes a brief history of the introductory college course. It begins by summarizing George Cobb's 1992 report which has become a generally accepted set of recommendations for teaching these courses. Results of a survey on the teaching of introductory courses are summarized along with a description of current versions of introductory statistics courses. A list of goals for students is given, based on what it means to be statistically literate. Six recommendations for the teaching of introductory statistics that build on the previous recommendations from Cobb's report are given. These recommendations are summarized as:

1. **Emphasize statistical literacy and develop statistical thinking;**
2. **Use real data;**
3. **Stress conceptual understanding rather than mere knowledge of procedures;**
4. **Foster active learning in the classroom;**
5. **Use technology for developing conceptual understanding and analyzing data;**
6. **Use assessments to improve and evaluate student learning.**

The report concludes with suggestions for how to make these changes, and includes numerous examples in the appendix to illustrate details of the recommendations.

Data Production and Exploratory Data Analysis

Goals: Students will be able to analyze data and comment on shape, measure of central tendency, measure of dispersion, and outliers or other oddities.

Key Ideas:

1. Simple graphs are useful.
2. It may be misleading to evaluate the data in isolation.
3. Comparison in relation to another data set often offers more insight.
4. Comparison is a skill used outside of statistics.

Prerequisite skills and concepts:
Explanation of shape, measures of central tendency, measures of dispersion, definition of outliers.

Connection to other statistical concepts:

- With few data points, graphical analysis helps to determine if inferential statistics should be used.
- Sometimes statistical inference is not required to state that there is a difference between the means of two groups.
- Variation occurs within, as well as, between two groups.

Objectives:

- Choose (and construct) graphs for comparison.
- Correctly use and understand appropriate comparative words.

Known Student Difficulties:

- Correctly pair mean and standard deviation, median and IQR
- IQR is one number, 25% of data are less than Q1, 50% of data are in IQR, 25% of data are greater than Q3
- When we compare two different data set's IQR, the larger IQR does not have more observations in it. All IQRs contain 50% of data.

Overview of Activity - Paper Helicopter
The basic experiment compares the accuracy of two types of helicopters. The activity offers a good opportunity to discuss variation.

An extension of this activity can be found in Nolan, Deborah and Speed, Terry, *Stat Labs: Mathematical Statistics Through Applications*, Springer, 2000

Assessment Activities:

- Exam questions can be taken from AP Statistics Free Response questions
- Projects are often used for assessment activities.

Activity - Paper Helicopter Experiment

Materials needed: For each pair of students, template for helicopter, paper clip, four feet of string, ruler, paper to mark target. Template can be found here: https://productiondifference.files.wordpress.com/2011/07/helicopter-template1.jpg

Overview: Helicopter will be dropped from waist height at target. The linear distance of the center of the helicopter shaft to the target is measured. A paper clip is attached to the shaft and the experiment is repeated.

Specific instructions: Determine who will be the person (launcher) dropping the helicopter. The launcher creates a plumb line by attaching the paper clip to the string. The launcher holds the plumb line at waist height from where the helicopter will be dropped while the other partner (measurer) marks the target on the paper as the point vertically underneath the launch point. Launcher drops the helicopter and remains in place while the measurer records the distance the center of the shaft is from the target. For the second launch the paper clip is attached to the shaft. The launcher drops the helicopter; the measurer records the distance.

1. Drop each of your helicopters 10 times. Record the distance to the closest half inch from the center of the helicopter's shaft to the target in the table below.

Helicopter type	1	2	3	4	5	6	7	8	9	10
With paper clip	18	2	4.5	7.5	3	16.5	2	17.5	19.5	2
Without paper clip	23.5	11.5	17.5	33	14.5	23	18	20	17.5	27

Calculate each measure of central tendency and dispersion using StatCrunch.

Helicopter type	Mean	Standard deviation	Q1	Median	Q3
With paper clip	9.25	7.63	2.00	6.00	17.5
Without paper clip	20.55	6.28	17.50	19.00	23.5
Overall for the 20 drops	14.90	8.94	6.00	17.50	19.75

2. Create a dotplot, by hand, for all 20 paper clips drops. On your dotplot mark each measure of central tendency.

3. Above the dotplot you've create, create a boxplot, by hand, for the 20 paper clip drops.

4. What does the boxplot show that the dotplot does not show as well?

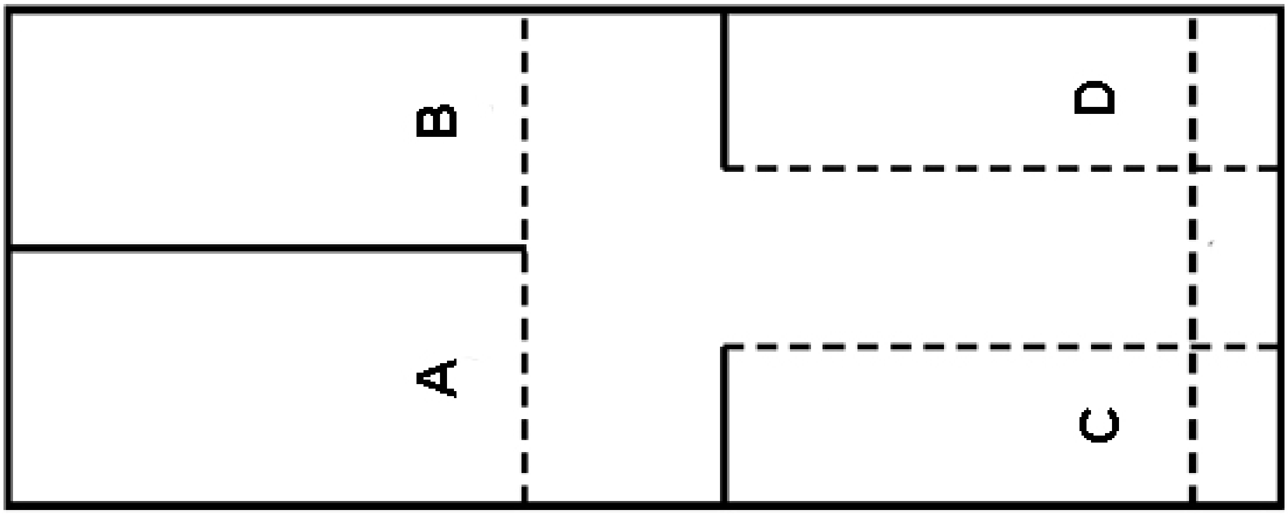
5. What does the dotplot show that the boxplot does not show as well?

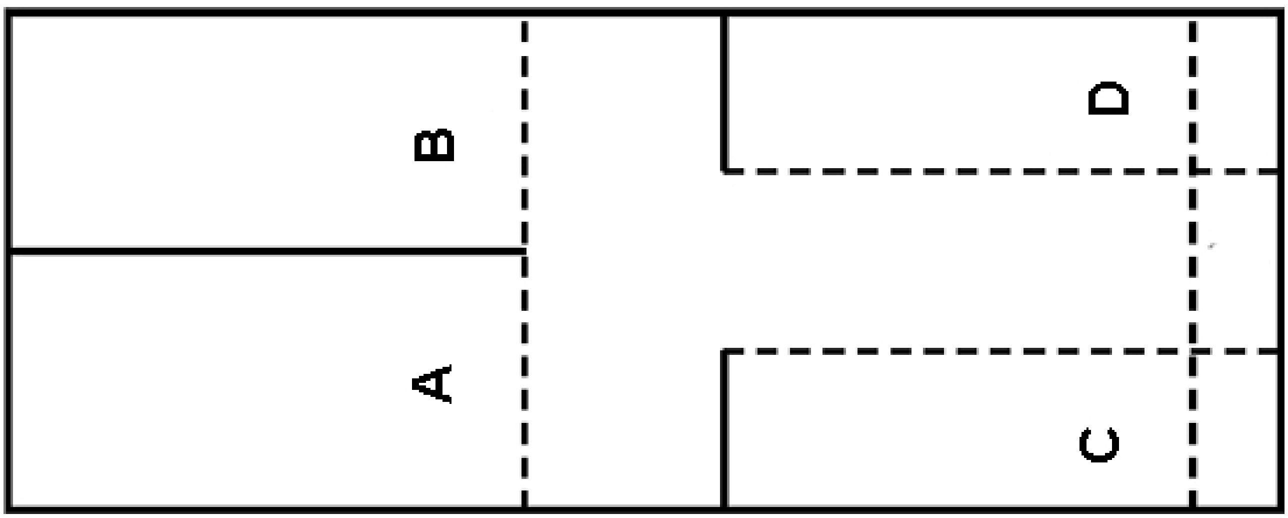
6. Create a boxplot for the helicopter drops with the paper clip.

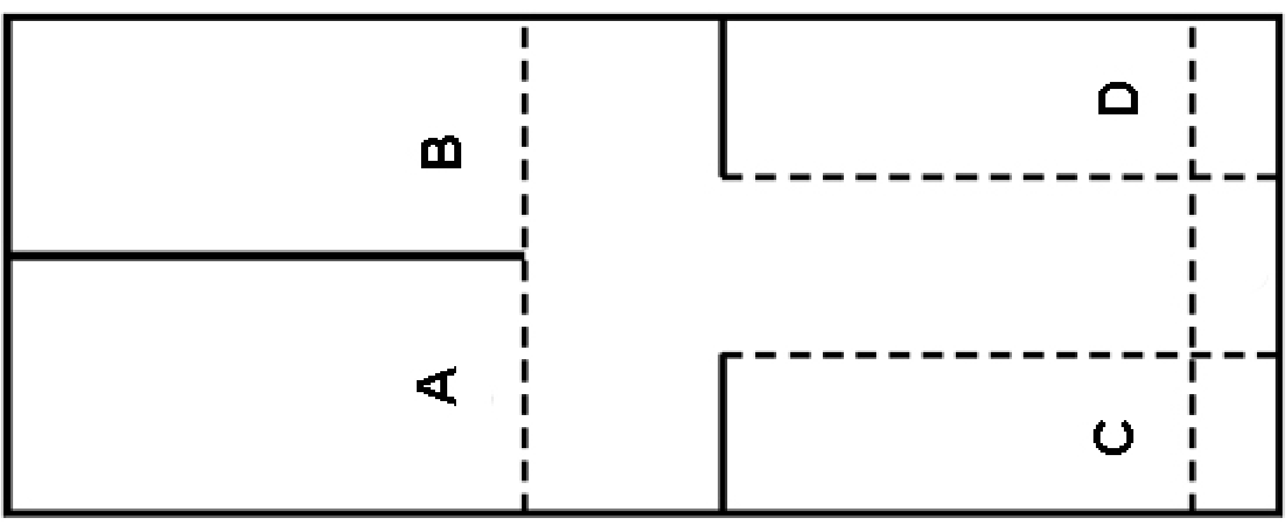
7. Using the same scale and underneath the boxplot created above, create a boxplot for the helicopter drops without the paper clip. Mark the mean on each graph.

8. Write a few sentences comparing the center and dispersion for each set of data.

9. Compare your data with another pair of students' observations. What are you willing to conclude about helicopters design based on this experiment?


B
A
D
C


B
A
D
C


B
A
D
C

Overview of Related Activity

Question of Interest: Do Double Stuf® Oreo cookies have twice the amount of filling as original Oreos®?

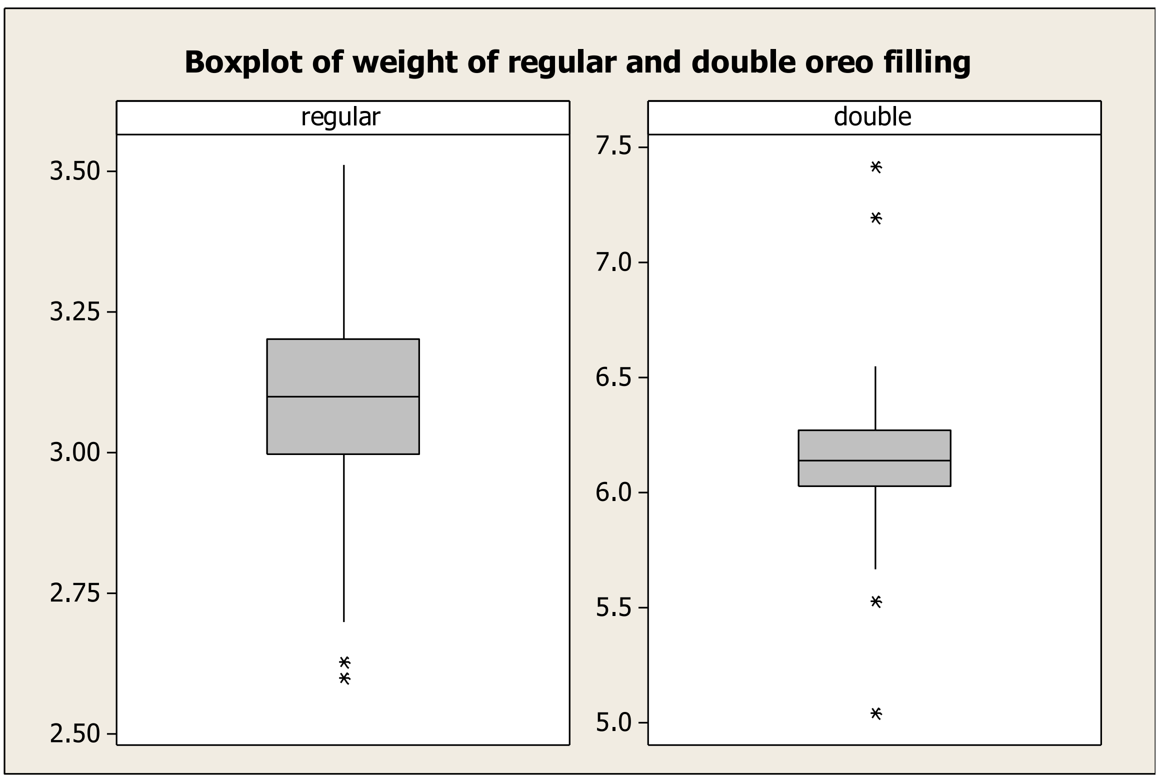
Procedure: Scrape the filling off and weigh filling.

My class results: We found no overlap between boxplots of grams of filling per Double Stuf® and grams of filling per Original. 95% confidence interval original (3.06371, 3.13416), 95% confidence interval Double Stuf(6.09123, 6.22371) Data available as a public file on StatCrunch entitled "weight of oreo filling."

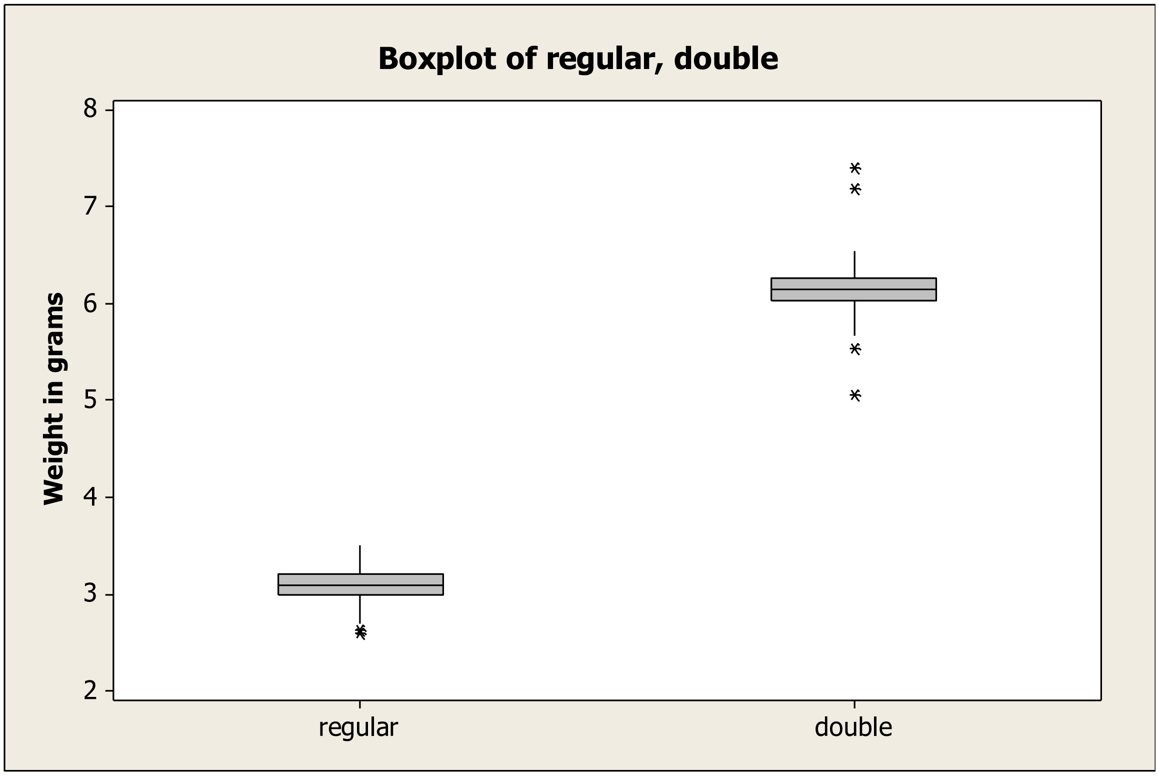
Adapted from "Double Stuf® Dilemma," Mathematics Teacher, 1999

Descriptive Statistics: regular, double

Variable	N	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
regular	94	3.0989	0.1720	2.6000	2.9975	3.1000	3.2025	3.5100
double	79	6.1575	0.2957	5.0500	6.0300	6.1400	6.2700	7.4100



GO ON TO THE NEXT PAGE.



Assessment Activities

Exam Assessment

AP Exam questions are carefully developed, rich problems with realistic data. The problem contexts do not need substantial explanations so student can focus on statistical concepts.

The 2015 AP Statistics Exam required a comparison of starting salaries for accountants at two large corporations. Graphs were given and students were expected to compare boxplots. (The question and grading rubric and explanations follow.)

Exam questions can be easy modifications of AP questions.

Other useful parts of the AP web site include:

- AP Statistics Home page
- An analysis of the questions explaining what concepts were examined
- Sample questions and scoring guidelines – access to free response questions

STATISTICS
SECTION II
Part A
Questions 1-5
Spend about 65 minutes on this part of the exam.
Percent of Section II score — 75

Directions: Show all your work. Indicate clearly the methods you use, because you will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

1. Two large corporations, A and B, hire many new college graduates as accountants at entry-level positions. In 2009 the starting salary for an entry-level accountant position was $36,000 a year at both corporations. At each corporation, data were collected from 30 employees who were hired in 2009 as entry-level accountants and were still employed at the corporation five years later. The yearly salaries of the 60 employees in 2014 are summarized in the boxplots below.

(a) Write a few sentences comparing the distributions of the yearly salaries at the two corporations.

(b) Suppose both corporations offered you a job for $36,000 a year as an entry-level accountant.

(i) Based on the boxplots, give one reason why you might choose to accept the job at corporation A.

(ii) Based on the boxplots, give one reason why you might choose to accept the job at corporation B.

Question 1

Intent of Question

The primary goals of this question were to assess a student's ability to (1) compare features of two distributions of data displayed in boxplots and (2) identify statistical measures that are important in making decisions based on data sets.

Solution

Part (a):

The median salary is approximately the same for both corporations. The range and interquartile range of the salaries are greater for Corporation A than for Corporation B. The two highest salaries at Corporation A are outliers while Corporation B has no outliers.

Part (b):

(i) Five years after starting, at least 3 out of 30 (10%) of the salaries at Corporation A are greater than the maximum salary at Corporation B. If I accept the offer from Corporation A, I might be able to make a higher salary at Corporation A than at Corporation B.

(ii) Five years after starting, the minimum salary at Corporation B is greater than at Corporation A In fact, at Corporation A it looks like some people are still making the starting salary of $36,000 and never received a raise in the five years since they were hired. So if I work at Corporation A, I might never receive a raise in salary.

Scoring

Parts (a) and (b) are scored as essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is scored as follows:

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

1. A correct comparison of center.
2. A correct comparison of spread.
3. A discussion of the outliers for Corporation A
4. The response is in context.

Partially correct (P) if the response includes only three of the four components.

Incorrect (I) if the response includes at most two of the four components.

Note: Any mention of shape should be ignored because complete shape information cannot be determined from a boxplot.

Question 1 {continued)

Part (b) is scored as follows:

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

1. In part (b-i) a relevant statistical measure is identified (or described) or a relevant statistical comparison is provided that supports the choice of Corporation A
2. In part (b-i) an explanation is provided for why the measure or comparison is relevant.
3. In part (b-ii) a relevant statistical measure is identified (or described) or a relevant statistical comparison is provided that supports the choice of Corporation B.
4. In part (b-ii) an explanation is provided for why the measure or comparison is relevant.

Partially correct (P) if the response includes only two or three of the four components.

Incorrect (I) if the response includes none or one of the four components.

Note: If a response does not provide a statistical measure or comparison in part (b-i) or (b-ii), the second and fourth components can still be satisfied if an acceptable explanation is provided that would follow from a relevant statistical measure or comparison. For example, if the response in part (b-i) only states "At Corporation A, I have the potential to earn a higher salary," the second component is satisfied.

4 Complete Response

Both parts essentially correct

3 Substantial Response

One part essentially correct and one part partially correct

2 Developing Response

One part essentially correct and one part incorrect

OR

Both parts partially correct

1 Minimal Response

One part partially correct and one part incorrect

Project Type Assessment

Overview - Pose a question that can be answered with a statistical analysis.

Your first project will involve collecting and analyzing data on a random variable from a population of your choice. The parameters of interest will be a MEAN and a DIFFERENCE IN MEANS for some quantitative variable. Additional statistical analysis will be done for the second project. More information on exactly what is required for the additional analysis will be given later. So that you can do that project using this data, will want to be able to get a sample of at least 40 values and you should be able to divide your sample into two groups (by some categorical factor) to compare their means.

After your project idea has been approved ... Collect your data to estimate a population mean and the difference in means between two subgroups.

Your final write-up for this project should include:

1. An introduction. Describe what variable you are measuring, the population involved, and any relevant background information. Describe in detail your method of sampling.
2. Data and graphical display. Include the raw data (in an electronic file via email) and a graph for the distribution as a whole. Also give a graphical display comparing your two subgroups. If you got the data from some source, that source should be cited in a bibliography.
3. Numerical Analysis. Give the appropriate statistics for your entire sample, as well as each group individually.

Due Dates and Breakdown

Due Monday November 7th - via the form below

1. Pose a question that can be answered with statistical analysis. State that question.

2. State the variable of interest and describe how you plan to sample. Be sure that you are selecting a valid numeric variable to measure (not a yes/no type proportion) - you should get a number for every sample point.

Due Friday November 18th.

3. A well written analysis with appropriate graphical and numerical analysis.
****
Note: You MUST have an approved project idea (including a guess for the population mean) BEFORE actually getting your sample..
****

Grading criteria:

Part 1

10 Appropriate question is posed
(Deduct 3 points for every attempt. Deduct all points if I need to give you a question.)

Part 2

10 Variable and sampling method are correct and explained
10 Graphical Analysis is appropriate and correct
10 Numerical Analysis is appropriate and correct
10 Writing – Spelling, Grammar, Organization, concise but sufficient detail is given

****

Some sources of data for the project.

1. Published data (newspapers, magazines, journals, catalogs, almanacs, World Wide Web,...)
2. Student/Faculty research projects
3. Local businesses
4. Athletic teams
5. Personal sources

QUESTION: Does the mean number of passes thrown in a football game differ between Division I college and professionals?
QUANTITAIVE VARIABLE TO STUDY: Number of passes thrown in football games
POPULATION: College and professional football games
SUBGROUPS: NCAA vs. NFL
METHOD OF SAMPLING: Data will be found from boxscores available at ESPN's website. Every third game from the several weeks in the season will be selected until we have a sample of 25 games from each group. The total number of passes thrown (by both teams) in each game will be determined from the boxscore.

--QUESTION: Is the average length of time between clothes washing different for male and female WC students?
QUANTITATIVE VARIABLE TO STUDY: # of days since last washed clothes
POPULATION: WC Student
SUBGROUPS: Male, Female
METHOD OF SAMPLING: Survey of students in our class. Ask “How many days ago did you last wash clothes”? (Also ask for gender.)

Sampling

Goal: Students will understand what is meant by random selection (sampling), take a simple random sample, identify poor sampling and suggest confounding variables.

Key ideas:

1. Sampling is an efficient way of inferring something about a population parameter. Random selection guarantees that inferences are generalizable to the population of interest. Sampling bias can lead to inaccurate or invalid estimates and is eliminated by random selection.
2. Random assignment to treatment is an important component of experimentation that eliminates the impact of confounders (will not be covered in the minicourse).
3. A confounding variable changes the apparent relationship between the explanatory and response variable.

Prerequisite skills and concepts:
Calculating mean and variance (or alternatively, none.)

Connection to other statistical concepts

1. Sampling Distributions
2. Central Limit Theorem
3. Interpretations of Confidence Intervals
4. Conclusions of Hypothesis Tests

Objectives

1. Learn how to take a random sample of data.
2. Examine bias from non-random sampling.

Known Student Difficulties:

1. Convince the need for simple random sampling

Overview of Activity – Random Rectangles

A population of N=100 rectangles of various areas are provided. Students will estimate the average area by guessing, choosing a representative sample (n=5), and random sampling (n=5). They will learn how to select a random sample and show that only the latter method produces unbiased results.

Assessment

- Students should be able to identify good and poor sampling techniques and discuss the bias found in the poor sampling techniques
- Students should be able to explain the difference between random sampling and random assignment.

Poor sampling techniques are readily available. Here is a classic example of poor sampling to get you started:

- http://www.learner.org/resources/series65.html (Video on Demand) Program 17 time 5:00 shows the Literary Digest's poll predicted Alf Landon to beat Roosevelt in 1936

Students need to understand why we take a sample.

- Taking a census doesn't necessarily solve the problem either. http://serc.carleton.edu/sp/cause/interactive/examples/18173.html

Count the F's in the following paragraph.

THE NECESSITY OF TRAINING HANDS FOR FIRST-CLASS
FARMS IN THE FATHERLY HANDLING OF FARM LIVESTOCK
IS FOREMOST IN THE MINDS OF FARM OWNERS. BECAUSE
THE FOREFATHERS OF THE FARM OWNERS TRAINED
THE FARM HANDS FOR THE FIRST-CLASS FARMS IN THE
FATHERLY HANDLING OF FARM LIVESTOCK, THE OWNERS OF
THE FARMS FEEL THEY SHOULD CARRY ON WITH THE FAMILY
TRADITION OF TRAINING FARM HANDS IN THE FATHERLY
HANDLING OF FARM LIVESTOCK BECAUSE OF THEIR BELIEF
THAT IT IS THE BASIS FOR GOOD FUNDAMENTAL FARMING.

Activity – Random Rectangles

This example is from Activity-Based Statistics. Second edition. By Schaeffer, Watkins, Witmer, and Gnanadesikan.

Random Rectangles Worksheet

Look at the following page for a few seconds. What is your best guess for the average area of the rectangles?

____

Look at the following page again. Choose five rectangles that you think are representative. Write down their ID#. Calculate the area for each one. Then calculate the average area.

	Rectangle 1	Rectangle 2	Rectangle 3	Rectangle 4	Rectangle 5
ID#					
Area					

_______ = Average Area of the five rectangles

Look at the following page again. Select five random rectangles using the following table of random digits [or using a computer program]. Write down their ID#. Calculate the area for each one. Then calculate the average area.

	Rectangle 1	Rectangle 2	Rectangle 3	Rectangle 4	Rectangle 5
ID#					
Area					

_______ = Average Area of the five rectangles

Join together with three other students to form groups of four. Write down the average of the five random rectangles each group member has calculated. Take the average of these averages to get the average area for 20 rectangles.

	Member 1	Member 2	Member 3	Member 4
Average				

_______ = Average Area of the four group members

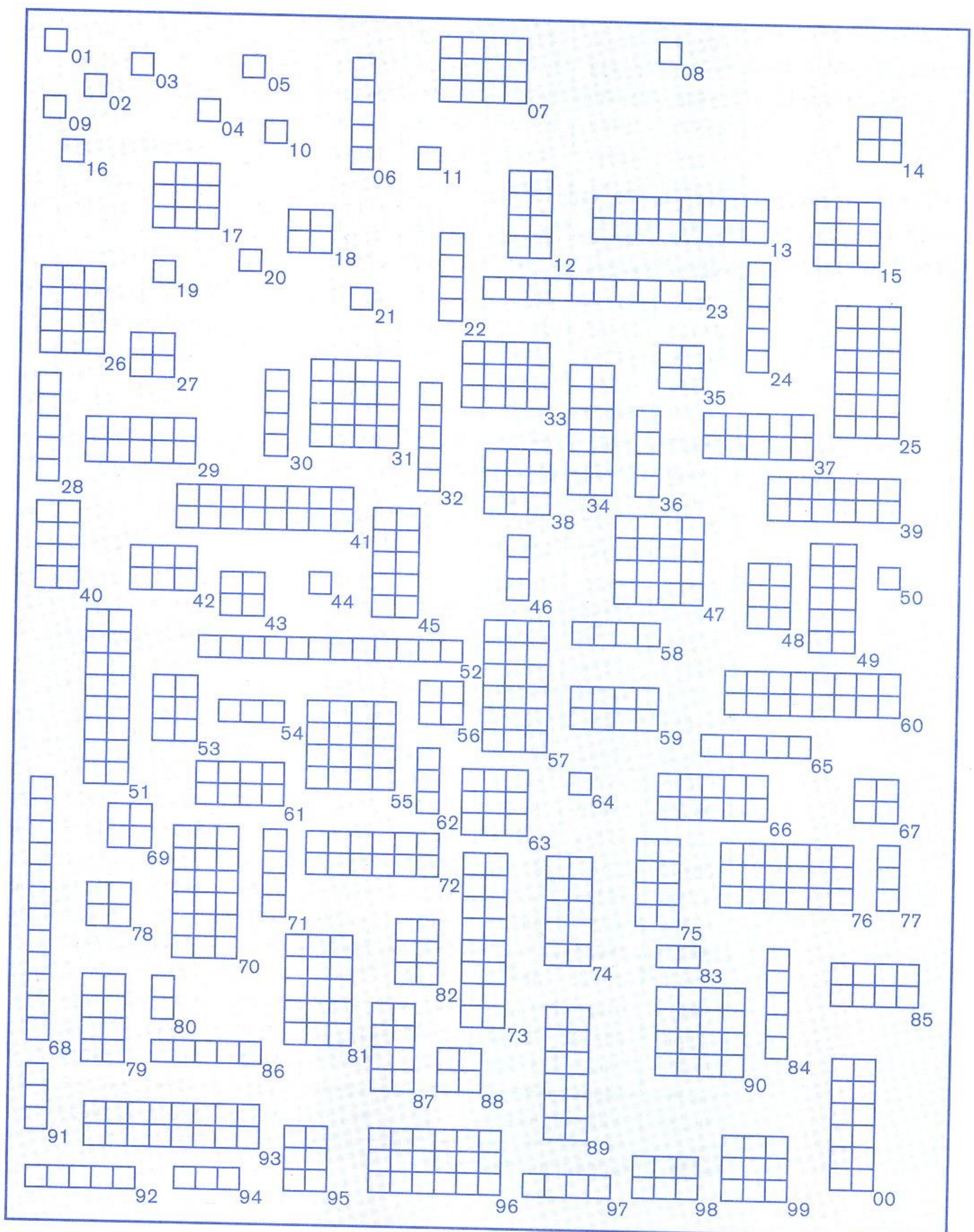
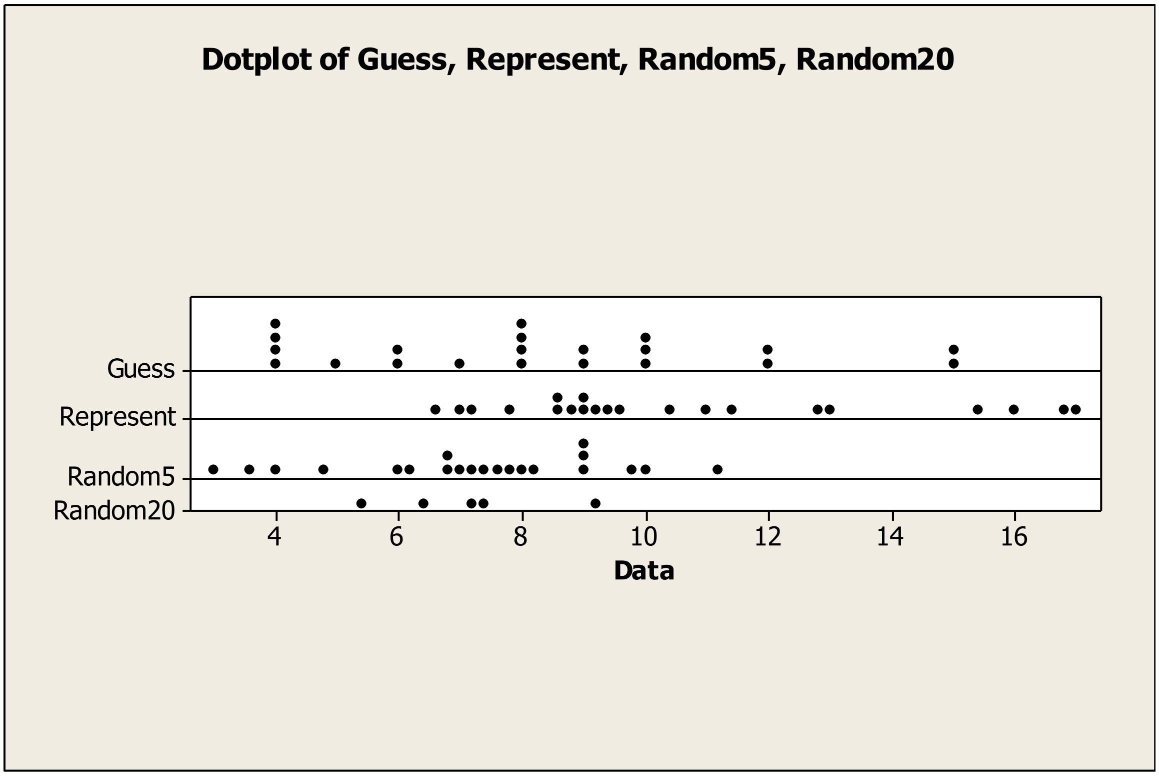


Figure 1

Line	**TABLE OF RANDOM DIGITS**								
1	79690	86718	33541	82828	12426	51929	14259	60592	55815
2	60945	95508	52743	53401	51154	20627	15127	11701	79304
3	16159	34564	59373	90106	41635	51518	59939	29712	98439

(note: if you get to the end of a line and are still involved in a simulation, continue down on the next line)

Sample MINITAB data of results from one class (for the population of 100):



One-Sample T: Guess, Represent, Random5, Random20

```
Test of mu = 7.42 vs not = 7.42


Variable    N    Mean  StDev  SE Mean        95% CI           T      P
Guess      21   8.286  3.349    0.731  (6.761,  9.810)   1.18  0.250
Represent  21  10.690  3.246    0.708  (9.213, 12.168)   4.62  0.000
Random5    21   7.252  2.137    0.466  (6.280,  8.225)  -0.36  0.723
Random20    5   7.124  1.362    0.609  (5.433,  8.815)  -0.49  0.652
```

Summary of Random Rectangles activity:

- Typically, guesses have the greatest variance
- Typically, representative groups are biased, but more precise
- Typically, random samples are unbiased
- Random20 is unbiased but <u>always</u> more precise than random5
- You have to be able to deal with random variation from “typical” occurrences “on the fly” with real data
- The random5 and random20 are not independent tests of each other, since we didn’t choose a new set of 20 observations.

Assessment Ideas

The best assessment of student's understanding of random sampling is often through asking them to collect data on a question of their choice that can be answered via a statistical study.

Many good assessment items can be found on ARTIST – Assessment Resource Tools for Improving Statistical Thinking (https://apps3.cehd.umn.edu/artist/index.html) and on the AP Statistics website.

There are three different levels of assessment on the ARTIST site; literacy, reasoning and thinking. Below are one example of each for the category Sampling.

Literacy - Question 2443
If the characteristics of a sample approximate the characteristics of its population in every respect, the sample is:

a. random
b. accidental
c. stratified
d. systematic
e. representative

Reasoning - Question 1436
Shelly is going to flip a coin 50 times and record the percentage of heads she gets. Her friend Diane is going to flip a coin 10 times and record the percentage of heads she gets. Which person is more likely to get 20% or fewer heads?

a. Diane because the more you flip the closer you get to 50%.
b. Shelly because the greater the sample size, the greater the variability in results.
c. Neither because each coin flip is a separate event and the probability of heads is not affected by the number of times flipped.

Thinking - Question 0114

Four students at a local high school conducted surveys. Shannon got the names of all 800 children in the high school and put them in a hat, and then pulled out 60 of them. Jake asked 10 students at an after-school meeting of the computer games club. Adam asked all of the 200 children in Grade 10. Claire set up a booth outside of the school. Anyone who wanted to stop and fill out a survey could. She stopped collecting surveys when she got 60 students to complete them. Who do you think has the best sampling method? Why?

Shannon – the sample size is reasonable. More importantly, she's chosen a method (simple random sampling) that allows every student to have the same chance of being chosen. Her method allows for equal representation of all students.

Jake will only have input from his computer games group. They may not be representative of all the students. Adam will not have representation from any except 10th graders. Claire is taking a convenience sample which is likely to include only those willing to stop by her booth.

Experimental and Observational Studies:

Cause and Effect

Goals:

1. Students will understand the difference between an experimental and an observational study.
2. Students will understand confounding variables and be able to identify potential confounding variables in a study.
3. Students will understand that association (correlation) does not necessarily imply causation.
4. Students will understand the components of experimentation in order to eliminate confounding variables and establish causation.

Key Ideas:

1. The use of random assignment to treatment as an important component of experimentation that eliminates the impact of confounders.
2. Confounding variables can explain an observed relationship between the explanatory and response variable.

Prerequisite skills and concepts: Obtaining a simple random sample or specifying a random assignment.

Connection to other statistical concepts: Sampling from a population or specifying random assignments, interpretations of confidence intervals, conclusions of hypothesis tests.

Objectives:

1. Identify explanatory and response variables.
2. Identify the type of study performed.
3. Show examples of confounding.

Known Student Difficulties (major and typical)

- Identifying variables.
- Identifying study design.
- Explaining confounding.
- Incorrectly inferring causation.

Overview of Activity 1: Identifying explanatory, response, and potential confounding variables. Several relationships are given and students are asked to identify confounding variable(s).

Related Activities and Resources:

- “Fighting cancer with raspberries: demonstrating the value of random assignment” webinar with applet: https://www.causeweb.org/webinar/activity/2009-05/
- A fun video to watch: Bonjour Paris L’ecole (https://www.youtube.com/watch?v=tS55WeYwPfA)
- Students read an actual published research paper and answer questions about study design..

Activity – Identifying Confounding Variables

- A **confounding variable** is a variable that is related to the explanatory variable and affects the response variable. Non-experiments are very susceptible to confounders.
- **Lurking variables** are unmeasured. Lurking variables can also be confounding variables.

People who attend church have lower blood pressure than those who stayed home and watched religious services on television. Maybe those who attended church had lower blood pressure as a result of a strong social support network rather than from attending church regularly. If this is the case, “amount of social support” is the confounding variable. Two conditions must be met for “amount of social support” to qualify as a confounding variable.

- Amount of social support (confounding variable) affects blood pressure (response variable)
- Amount of social support (confounding variable) is related to attending church regularly (the explanatory variable) because the same people who attend church regularly are likely to have a strong social support network.

For each of the following, identify the confounder:

- Ice cream sales go up when more baseball games are played
- The average number of speeding tickets that a student has is less than that of his/her parents.
- Typically, students do worse on exam problems that they spend a lot of time on. Clearly spending more time on a problem doesn’t make you do worse! What’s going on here?
- Mortality rate and country – specifically the mortality rate in the US is higher than most other North and South American countries
- Studies have shown that the more children a woman has, the lower her risk of breast cancer.

Answers: Season, Age, Difficulty of Problem, Age Distribution, Breast Feeding

Related Activity Resource: Students read an actual published research paper and answer questions about study design

JAMA. 2007 Jul 4;298(1):49-60. http://jama.jamanetwork.com/article.aspx?articleid=207783

Effects of low habitual cocoa intake on blood pressure and bioactive nitric oxide: a randomized controlled trial. Taubert D, Roesen R, Lehmann C, Jung N, Schömig E.

Department of Pharmacology, University Hospital of Cologne, Cologne, Germany.
dirk.taubert@medizin.uni-koeln.de

CONTEXT:
Regular intake of cocoa-containing foods is linked to lower cardiovascular mortality in observational studies. Short-term interventions of at most 2 weeks indicate that high doses of cocoa can improve endothelial function and reduce blood pressure (BP) due to the action of the cocoa polyphenols, but the clinical effect of low habitual cocoa intake on BP and the underlying BP-lowering mechanisms are unclear.

OBJECTIVE:
To determine effects of low doses of polyphenol-rich dark chocolate on BP.

DESIGN, SETTING, AND PARTICIPANTS:
Randomized, controlled, investigator-blinded, parallel-group trial involving 44 adults aged 56 through 73 years (24 women, 20 men) with untreated upper-range prehypertension or stage 1 hypertension without concomitant risk factors. The trial was conducted at a primary care clinic in Germany between January 2005 and December 2006.

INTERVENTION:
Participants were randomly assigned to receive for 18 weeks either 6.3 g (30 kcal) per day of dark chocolate containing 30 mg of polyphenols or matching polyphenol-free white chocolate.

MAIN OUTCOME MEASURES:
Primary outcome measure was the change in BP after 18 weeks. Secondary outcome measures were changes in plasma markers of vasodilative nitric oxide (S-nitrosoglutathione) and oxidative stress (8-isoprostane), and bioavailability of cocoa polyphenols.

RESULTS:
From baseline to 18 weeks, dark chocolate intake reduced mean (SD) systolic BP by -2.9 (1.6) mm Hg ($P < .001$) and diastolic BP by -1.9 (1.0) mm Hg ($P < .001$) without changes in body weight, plasma levels of lipids, glucose, and 8-isoprostane. Hypertension prevalence declined from 86% to 68%. The BP decrease was accompanied by a sustained increase of S-nitrosoglutathione by 0.23 (0.12) nmol/L ($P < .001$), and a dark chocolate dose resulted in the appearance of cocoa phenols in plasma. White chocolate intake caused no changes in BP or plasma biomarkers.

CONCLUSIONS:
Data in this relatively small sample of otherwise healthy individuals with above-optimal BP indicate that inclusion of small amounts of polyphenol-rich dark chocolate as part of a usual diet efficiently reduced BP and improved formation of vasodilative nitric oxide.

TRIAL REGISTRATION:
clinicaltrials.gov Identifier: NCT00421499.

Questions for Students:

1. Using appropriate statistics terms, what type of study was this study? Do not use the word "trial" in your answer!

2. From the objective listed in the abstract, state the question of interest using appropriate statistical terms.

3. Write the null and alternate hypothesis for your question of interest listed above.

4. This study was investigator blind as noted in the abstract. Could the study have been double-blind? If so, explain how. If not, explain why not.

5. Based on your null and alternate hypotheses above (and using the conclusions of the study) write a statistical conclusion which appropriately includes the correct p-value.

6. In the context of this study, what is a Type I error?

7. In the context of this study, what is a Type II error?

8. List two potential lurking variables in the cocoa and blood pressure study.

Sampling Distributions and the Central Limit Theorem

Goal: Students will understand sampling distributions and the central limit theorem of statistics.

Key ideas:

1. The sampling distribution of the sample mean (or sample proportion) represents all the possible means of random samples of size n from the population. It has a predictable shape, center and dispersion.
2. As the sample size (n) increases, the shape becomes more normally distributed and the standard error decreases by a factor of $\frac{1}{\sqrt{n}}$.

Prerequisite skills and concepts:

1. Calculating the mean and variance.
2. Random sampling.
3. Basic understanding of normal and binomial probability distributions.

Connection to other statistical concepts:
Statistical inference (hypothesis testing and confidence intervals)

Objective:
See how sampling distribution is created and how increasing ***n*** impacts its shape and precision.

Known Student Difficulties:

1. Students confuse the distribution of the sample and the distribution of the sample statistic. Students want to take many samples to create a sampling distribution.

Overview of Activity Random Pennies - Understanding sampling distributions
Examine the impact of sample size and underlying distribution on shape and precision.

Assessment

1. Consider revisiting Random Rectangles

2. Each time a problem is discussed have the student draw a graph of what they think the population, sample and sampling distribution of the sample statistics looks like.

3. Happyville, developed by Kevin Robinson, Millersville University, is an activity similar to random rectangles. A webinar can be found here:
   http://www.causeweb.org/webinar/activity/2011-07/
   the web site which includes a quiz on sampling distributions is here:
   https://www.pdffiller.com/en/project/48094691.htm?form_id=80145539

Activity Random Pennies adapted from *Activity Based Statistics: Student Guide* see references; also uses applet from the Rossman/Chance collection.

Consider the age of pennies in circulation. We expect more new pennies and fewer older pennies.

Someone has collected a large number of pennies from circulation, recorded the mint date and from that and found the age of those pennies. Sketch the graph of the distribution of this population estimating and marking the mean and standard deviation.

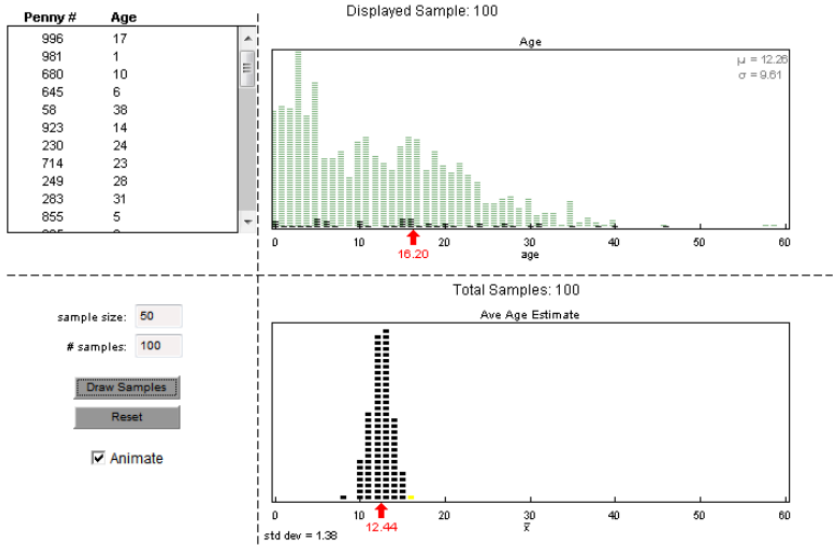
Now you take a sample of 50 from this population. Sketch the graph of your sample estimating and marking the mean and standard deviation.

We will use a computer to find the sampling distribution of sample mean age of pennies in circulation. Sketch the graph of this sampling distribution estimating and marking the mean and standard deviation**.**

Use this applet for the demonstration of the sampling distribution
http://www.rossmanchance.com/applets/SampleData/SampleData.html

Rossman/Chance Applet Collection

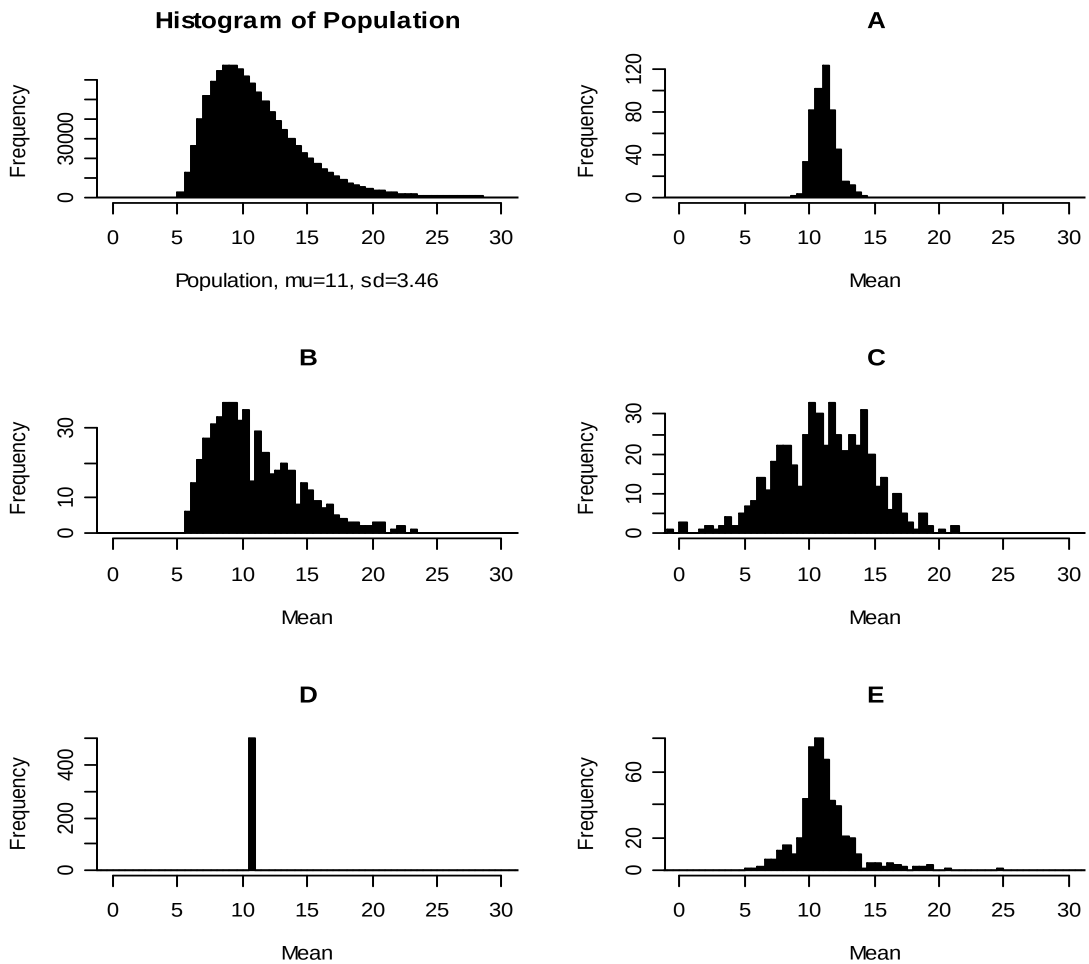
Sampling Pennies



Assessment of Sampling Distribution

From pilot study “Towards an Understanding of Prerequisite Knowledge of Sampling Distributions” by Tisha Hooks, Michael Posner, Michelle Sisto, and Dale Berger (USCOTS, 2009)

The first graph below is a distribution for a population of test scores. Each of the other five graphs, labeled A to E represent possible sampling distributions of sample means for 500 random samples drawn from the population.



1. Which graph represents a sampling distribution of sample means for 500 samples of size 1?

   A B C D E

   Explain your answer:

2. Which graph represents a sampling distribution of sample means for 500 samples of size 16?

   A B C D E

   Explain your answer:

Statistical Inference – Confidence Intervals

Goals:
1. Students will understand that confidence intervals are used to estimate possible values of a population parameter.
2. Students will be able to explain why a confidence interval is better than a point estimate.

Key Ideas:
1. A confidence interval consists of a point estimate and the margin of error.
2. The confidence level explains the proportion of times in repeated sampling we expect the population parameter to be captured by the interval created.
3. How the margin of error is affected by the confidence level and the sample size.
4. Confidence intervals can be used to determine the results of hypothesis tests.

Prerequisite skills and concepts:
1. Obtaining a simple random sample or specifying a random assignment
2. Identifying sampling errors and non-sampling errors.

Connection to other statistical concepts:
1. **S**ampling from a population
2. Normal distributions
3. Sampling distributions.

Objectives:
1. Compute a a 95% confidence interval based on sample data or a simulation.
2. Interpret a 95% confidence interval based on sample data or a simulation.

Known Student Difficulties:
- Confidence interval is interpreted to be about the sample statistic and not the population parameter.
- Confidence level and confidence interval are two different aspects of the inference about a population parameter using an interval.
- Confusing "confidence" with probability.

Overview of Activity: Reese's Pieces (Confidence Interval for a Population Proportion)
Confidence intervals are generated for the proportion of orange candies in Reese's Pieces candy. This activity reviews the sampling distribution for a population proportion. Confidence intervals are then computed from sample data. The relationship between confidence level, sample size, and margin of error are examined. The meaning of "confidence" is explored.

Activity: Reese's Pieces. Modified from *Workshop Statistics, Discovery with Data*, 2nd Edition, by Allan Rossman and Beth Chance, Key College Publishing, 2001.

Topics:

Review of Sampling Distributions for Proportions
Confidence Intervals for Proportions

Review of Sampling Distributions for Proportions:

We are going to obtain a random sample of $n = 25$ Reese's Pieces**.** Before we do so, answer questions (a) through (e).

(a) We are interested in the color of the Reese's Pieces. In particular we want to know if each piece of candy is orange or not. What type of variable is this (categorical or quantitative)?

(b) Let X = "number of orange candies in your sample." What are the possible values for X? Will everyone in class get the same value for X?

(c) Consider the total size of the population of all Reese's Pieces candies. Do you think it would be more than 20 times the size of your sample?

(d) Let p be the proportion of all Reese's Pieces candies that are orange. Is p a parameter or statistic?

(e) Do you know the exact value for p?

Now let's obtain our random sample. Be careful how you sample!

(f) Let $\hat{p} = \frac{X}{n}$ (i.e. the proportion of Reese's Pieces in your sample that are orange). Is $\hat{p}$ a parameter or statistic?

(g) Do you know the exact value for your $\hat{p}$?

(h) Did everyone in the class obtain the same value for $\hat{p}$? What do we know about the sampling distribution of $\hat{p}$ (i.e. what is its mean and what is its standard deviation)? How could we examine this theoretical distribution? Does the shape of this distribution remind you of any other distribution?

(i) What conditions should be satisfied to use a normal distribution to approximate the sampling distribution of $\hat{p}$? Do we meet those conditions?

(j) According to the empirical rule, what should be true for about 95% of the $\hat{p}$'s that we randomly sample from the population of Reese's candies?

Confidence Intervals for Proportions:

We say the statistic $\hat{p}$ is a "point estimate" for the parameter p.

When we know the sampling distribution of an estimator, such as $\hat{p}$, we can construct a confidence interval, or "interval estimate" for the population parameter.

In this case, an approximate 95% confidence interval for p is: $\hat{p} \pm 2\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$.

More generally, we use:

Point Estimate $\underline{+}$ Margin of Error

where

Margin of Error = critical value * standard error of estimator

The critical value depends on the distribution family of the estimator and the level of confidence. With proportions, the 2 came from the empirical rule of the normal distribution.

(a) Use your value of $\hat{p}$ computed above and the formula $\hat{p} \pm z^*\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$, where z^*=1.960 to find the endpoints of a 95% confidence interval for p, the proportion of all Reese's Pieces that are orange.
(NOTE: The above computation can also be done using the STAT|TESTS|1-PropZInt key on the TI-83 and TI-84 series calculators)

(b) Now we will plot our intervals on the same axes and discuss our results. In particular, about what percent of these intervals do we expect actually contain the (unknown) value of p?

(c) We will now use simulation to examine the behavior of confidence intervals. Use an Internet browser to open the "confsim" applet found at http://www.rossmanchance.com/applets/ConfSim.html

In this applet, Π will play the part of the unknown parameter p.

- Set the Method to Proportion|Binomial|Wald
- Enter 0.45 for the value of Π. (i.e. 45%)
- Enter 25 for the sample size.
- Enter 1 in the "intervals" box.
- Click the recalculate button.

If you click on the resulting interval, you should see the value of $\hat{p}$ that generated the interval, along with the endpoint values.

(d) Did all of us obtain the same interval? Why not?

(e) Click the Recalculate button again. Did you get the same interval? Did the interval capture the "true value of Π" (in this case, 0.45)?

(f) Click the Recalculate button a few more times. The interval will be red if 0.45 is not within the endpoint values. How often do you expect this to happen?

(g) Put 100 in the intervals box and click Recalculate. The applet will count how many of the resulting intervals actually do capture Π. What is this percentage?

(h) Click the Recalculate button a few more times, what value does the "running total" percentage approach?

(i) We are going to use the applet to see how confidence intervals behave as we vary the confidence level and the sample size. Use the applet with the following values to see how the confidence intervals compare. Make sure you use a large "running total" to estimate the percent of intervals that contain Π.

Table 1. Varying the Confidence Level

	Baseline	Comparison	
Π	0.45	0.45	0.45
C-level	95%	90%	99%
n	25	25	25
Margin of Error Compared to Baseline is generally:		Larger Smaller	Larger Smaller
Percent of Intervals that Contain Π is:			

Table 2. Varying the Sample Size

	Baseline	Comparison	
Π	0.45	0.45	0.45
C-level	95%	95%	95%
n	100	50	1000
Margin of Error Compared to Baseline is:		Larger Smaller	Larger Smaller
Percent of Intervals that Contain Π is:			

(j) Write a few sentences about how sample size and confidence level affect the margin of error.

Assessment: From Doug Andrews, Wittenberg College

Assessing CI Interp's Name ___________________

OK, below are your interpretations of the confidence interval that we constructed in class last time. Now it's time for you to assess the quality of each interpretation. Next to each bullet, assign a rating (from 0 to 10) to each interpretation. I'll calculate the correlation between your ratings and my own ratings, and your score for this activity will be based on that correlation. Here goes....

- This confidence interval means that 90% of the students in this stats class have ages in months that fall between 242.4 and 256.3 months. Also, if we were to make an inference about the entire witt population based on this sample, 90% of all witt students would also have ages between 242.4 and 256.3 months.
- We are 90% confident that the mean mu of the ages of the population of all Witt students lies between the values of 242.4 and 256.3, based on a random (kind of) sample of 23 students from stats 127.
- We are 90% confident that, based on this sample, the average age of Wittenberg students is between 242.4 months and 256.3 months
- We can be 90% confident that the true mean age (in months) of all students at Wittenberg falls between the interval (242.4, 256.3).
- This interval says that we're 90% sure that the mean age of people taken from our sample will be between 242.4 and 256.3.
- We are 90% confident that the mean age (in months) for all students at Wittenberg is between 242.4 and 256.3 months.
- All Witt students have an age between 242.4 and 256.3 months with 90% confidence
- 90% of all SRS's with a sample size of 23 from this university would yield a mean somewhere in between 242.4 and 256.3, or in other words if a sample is taken and the mean is not in this range, there is a 10 % chance that it is due to pure luck.
- 90 percent of all samples of student ages (in months) will be between 242.4 and 256.3.
- 90% of Witt students accorrding to this SRS meu=23 fall between 242.4 and 256.3 months

By the way, one group added the following: "However, this does not take into account the fact that this is not an SRS. People who take this stats courses may be a different average age than the entire population of Witt students." How true – and that's probably the more important message from this activity!

Statistical Inference – Hypothesis Testing

Goals:

1. Students will understand that statistical inference is used to assess possible values of a population parameter.
2. Students will be able to test to see if observed data supply sufficient evidence that the population parameter is not as expected.

Key Ideas:

Hypothesis testing consists of

a. specifying a null and alternate hypothesis based on a status quo parameter and an alternate to the status quo
b. collecting data
c. determining how likely the data collected is given the status quo.

We seek evidence to reject or fail to reject the status quo.

Prerequisite skills and concepts: Obtaining a simple random sample or specifying a random assignment, sampling errors and non-sampling errors.

Connection to other statistical concepts:

1. Conditional probability
2. sampling from a population
3. sampling distribution of a statistic

Objectives:

Perform the steps of a hypothesis test and interpret the p-value in the context of the problem

Known Student Difficulties:

1. Alternate hypothesis is research hypothesis
2. Failing to reject null hypothesis is not the same as "accepting the null" We do not have sufficient evidence to reject the null.
3. p-value is a conditional probability – condition null hypothesis is true
4. p-value measures the element of surprise. A small p-value says we are very surprised with our sample results if the null hypothesis is true

Overview of Activity An Introduction to Hypothesis Testing via Roxy Peck's Cookie Game

Students "construct" the steps of a hypothesis test by considering the probability of drawing a red card in new deck of cards.

Related Activity. German Tanks activity (versions of this activity can be found Mathematics Teacher, 1999 vol. 92 issue 8 and *Activity Based Statistics*) can be used to help students understand that all good estimators are not the mean. A CAUSEweb Webinar was presented by Diane Evans on the German Tank problem. A link to the webinar can be found at http://www.causeweb.org/webinar/activity/2009-09/

Assessment

1. AP Statistics exams always have a nuanced inference question.
   Here's the link gets you to the Excel file that classifies each of the free response questions.
   http://apcentral.collegeboard.com/apc/members/courses/teachers_corner/47425.html

2. The ARTIST web site (Assessment Resource Tools for Improving Statistical Thinking) has a good series of confidence interval and hypothesis testing questions. The following is classified as general literacy.

Test of significance Q1182

Which of the following is true?

a. It is impossible to prove the null hypothesis.
b. It is always possible to prove the null hypothesis.
c. It is possible to prove the null hypothesis under certain conditions

Activity - Hypothesis Testing Students develop the hypothesis testing process.
http://www.amstat.org/publications/jse/v2n1/eckert.html

Materials needed cookies, two new decks of cards

Procedure summarized

- A new deck of cards is opened in front of the students and shuffled.
- Students are asked if they would like to participate in the process of drawing cards to receive a cookie. If so, they should queue up by the instructor.
- Students draw cards one at a time, replacing the card in the deck after noting the color. The deck is reshuffled after each draw. If a red (black) card is drawn, the student receives a cookie. If not, the student sits down. Usually after about five or six draws the remaining students "give up" and sit down before drawing a card.
- The deck of cards is put aside and students never know what is in the deck.

The parallel to hypothesis testing

- Hypotheses are written
  Prior to the first draw, students expect the status quo that the proportion of red (black) cards is equal to 50%. The alternate (prior to the first draw) is that the proportion of red (black) cards is not equal to 50%.
  $H_o : p = 0.5$
  $H_a : p \neq 0.5$

- Evidence is collected
  Students draw a card from the deck, recording the color. (Card is replaced; deck is reshuffled.) A success is drawing the red card and getting the cookie. A failure is drawing a black card and thus not getting a cookie.

- p-value is calculated
  The probability of success using the expected proportion of 50% red is p = 0.5, the null hypothesis. The probability of failure is 1-p = 0.5. We observe a series of failures.
  one failure p-value = 0.5
  two failures p-value = 0.5 * 0.5 = 0.25
  three failures p-value = 0.5 * 0.5 * 0.5 = 0.125
  four failures p-value = 0.5 * 0.5 * 0.5 * 0.5 = 0.0625
  five failures p-value = 0.5 * 0.5 * 0.5 * 0.5 * 0.5 = 0.03125
  six failures p-value = 0.5 * 0.5 * 0.5 * 0.5 *0.5 * 0.5 = 0.015625

- Decision is made
  We decide that it is unlikely that the deck has 50% red cards based on the observed evidence

- Population parameter is never known
  Deck is put away and students don't know if they've made the correct decision.

Annotated List of Resources for Teaching Introductory Statistics

General Resources

- CAUSE (Consortium for the Advancement of Undergraduate Statistics Education): http://www.causeweb.org/ is a national organization whose mission is to support and advance undergraduate statistics education, in four target areas: resources, professional development, outreach and research. The web site has links to material and resources in each of these areas. Resources can be searched by pedagogical type or statistical techniques. Upcoming workshops and professional development opportunities are listed. Information on USCOTS – the US Conference on Teaching Statistics (May 28-30, 2015 can be found here.
- GAISE (Guidelines for Assessment and Instruction in Statistics Education) guidelines: http://www.amstat.org/education/gaise/ Reports for preK-12 and College level and commentary on these documents.
- Common Core State Standards http://www.corestandards.org/Math/Content/HSS/introduction/
- American Statistical Association https://www.amstat.org The ASA has a tab on its homepage linking to numerous resources for statistics education. In addition, there is a section (group similar to a SIGMAA) devoted to statistics education.
- This is Statistics: http://thisisstatistics.org/ ASA's site to promote the study of statistics. Links to examples of how people who work in careers in statistics have fun, make a difference, satisfy curiosity, and make money.

Lesson Plan and Video Resources

- Against All Odds, Inside Statistics: http://www.learner.org/resources/series65.html A video instructional series on statistics for college and high school classroom and adult learners. The series consists of 32 videos of length 6 to 14 minutes along with coordinated guides for students and faculty.
- Statistics Education Web (STEW): http://www.amstat.org/education/stew/ for K-12. Lesson plans for teachers. Some of this material is appropriate for a conceptual statistics course. It could also be used in pre-service and in-service teacher training.

Data Set Resources

- Data and Story Library: http://lib.stat.cmu.edu/DASL/ online library of and stories that illustrate the use of basic statistics methods. Link was broken as of December 16th. Statisticians are working with CMU to reopen link.
- Electronic Encyclopedia of Statistical Examples and Exercises, available with Moore and McCabe: http://www.whfreeman.com/eesee/eesee.html These sets of statistical stories are classified by subject (e.g. Animal and Plants, Political Science, Law and Demography) or statistical topic (e.g. Descriptive Statistics, Regression. The stories are referenced in some Freeman texts; however, they are freely available. Assessment questions are included with each story.
- Journal of Statistics Education Datasets and Stories column: http://www.amstat.org/publications/jse/ This peer reviewed journal includes an archive of datasets and their associated stories. The datasets are complex and, generally, each can be used for several different types of analysis.

- CAUSE data resources: https://www.causeweb.org/cwis/b37/index.php?P=BrowseResources&FieldId=27 Link to datasets cataloged by editors of CAUSEweb.
- Journal of American Medical Association with abstracts of articles: http://jama.ama-assn.org/ Abstracts are freely available. The abstracts can be used to help students understand various study designs including double and single blind, placebo and control groups. Explanations of subject recruitment are often included in the abstract which facilitates discussion of allowable generalizations.
- Government websites – Census, Labor, etc., www.usa.gov/statistics Data are not easily accessed but with some effort a student can get substantial data from this portal.

Activity Resources

- CAUSEweb: http://www.causeweb.org/ search by statistical topic or pedagogical method – A go to site for many activities. Several activities marked with a ** have been reviewed. Some activities have instructor guides.
- *Activity Based Statistics: Student Guide*, 2nd Edition, 2004. R.L. Scheaffer, A.E. Watkins, J.A. Witmer, M. Gnanadesikan. Revised by T. Erickson. Key College Publishing, ISBN 1-930190-72-7. Instructor Guide which outlines how to use the activities is available. One of the original collections of activities. Several of the classic examples (e.g. random rectangles, central limit theorem for dates on coins) come from this resource.
- Stat2Labs by Shonda Kuiper: http://web.grinnell.edu/individuals/kuipers/stat2labs. Activities/Games for statistical learning. Guided activities which allow students to create and test hypotheses. While this site is intended for use in a second course, aspects of it can be used to design simple experiments which can be tested in a first course.

Useful applets

- Web Interface for Statistics Education (WISE): http://wise.cgu.edu/ contains tutorials geared toward student understanding and applets which are embedded in the tutorials. The applets can be used without the tutorials. The tutorials are very useful in explaining common misunderstandings of statistical concepts. The tutorial on power and sample size is particularly useful.
- Rice University Simulations and Demonstrations: http://onlinestatbook.com/stat_sim/ contains many of the "standard" applets for teaching statistical concepts. The sampling distribution is particularly useful (http://onlinestatbook.com/stat_sim/sampling_dist/index.html)
- Statistical Calculators at UCLA (SOCR): http://calculators.stat.ucla.edu/
- Type I and II error effects allows changing difference between means, sample size, alpha level: http://www.amstat.org/publications/jse/v11n3/java/Hypothesis/ This article in JSE explains Type I and Type II errors and has an associated applet.
- Correlation guessing game: http://www.istics.net/Correlations. John Marden at University of Illinois Urbana-Champaign has produced a neat Java applet for practicing the connection between scatterplots and correlation. Email correlations@stat.istics.net to set up your own school-specific leaderboard.

- Rossman/Chance Applet Collection: http://www.rossmanchance.com/applets/ A nice collection of applets, including the Reeses Pieces applet used in this workshop.

Books Several of these books are classics in teaching statistics. We list them here to give you an idea of where to look for some low-tech ideas for teaching statistics.

- Publications by MAA (can be found at www.maa.org under tab entitled "Publications"
    1. *Statistics for the Twenty-First Century*, Florence and Sheldon Gordon, eds., MAA Notes 26, 1992. Includes "Low-Tech Ideas for Teaching Statistics," "Annotated Bibliography of Real-World Datasets Useful for Teaching Applied Statistics," and two classic articles "Towards Lean and Lively Courses in Statistics" by Robert Hogg and "Teaching Statistics as a Respectable Subject" by David Moore.
    2. *Teaching Statistics: Resources for Undergraduate Instructors*, Thomas Moore, ed., MAA Notes 52, 2000. Includes sections on Teaching with Data, Established Projects in Active Learning, Textbooks, Technology and Assessment
    3. *Innovations in Teaching Statistics*, Joan Garfield, ed., MAA Notes 65, 2005. Addresses statistics teachers in different environments including those teaching in mathematics departments, two-year community colleges, statistics departments and in other disciplines.
- *Teaching Statistics: a bag of tricks* by Andrew Gelman and Deborah Nolan, Oxford University Press, 2002

AP Stuff The AP Statistics community is both active and intentional in providing good resources to the AP teachers. These resources represent two places to start looking at available materials.

- http://apcentral.collegeboard.com/apc/public/courses/teachers_corner/2151.html
- Ruth Carver's Statistics Website: http://mathforum.org/ruth/stat.html

Journals/Magazines

- Journal of Statistics Education (JSE): http://www.amstat.org/publications/jse/ (free to all)
- Statistics Education Research Journal (SERJ): http://www.stat.auckland.ac.nz/~iase/publications.php?show=serj
- Significance magazine, jointly published by the ASA and the Royal Statistical Society. http://www.statslife.org.uk/significance
- Teaching Statistics: http://www.teachingstatistics.co.uk/
- Chance: http://www.amstat.org/publications/chance/

Fun Stuff

- Fun resources on CAUSEweb: http://www.causeweb.org/resources/fun/agreement.php.
- Some comics: http://xkcd.com/882/, http://xkcd.com/539/, http://xkcd.com/552/, http://xkcd.com/1132/, http://xkcd.com/605/