

GAISE College Report Appendix

Examples and commentary in this appendix are provided for additional guidance, clarification and illustration of the guidelines in the main report.

(A) Examples of projects and activities

Some activities that could be improved

- (1) Pepsi vs. Coke Activity
- (2) A Central Limit Theorem Activity

Additional examples of activities and projects

- (3) Data Gathering and Analysis: A Class of Projects
- (4) Team constructed questions about relationships
- (5) Comparing Manual Dexterity under Two Conditions

(B) Examples of assessment items

- (1) - (3) Some items with problems and commentary on the flaws
- (4) - (7) Examples showing ways to improve some assessment items
- (8) - (36) Additional examples of good assessment items

(C) Example of using technology

(D) Examples of naked, realistic and real data

(E) Example of a course syllabus

A. Examples of activities and projects

Some desirable characteristics of class activities:

1. *The activity should mimic a real-world situation. It should not seem like “busy work.” For instance, if you use coins or cards to conduct a binomial experiment, explain some real-world binomial experiments that they could represent.*
2. *The class should be involved in some of the decisions about how to conduct the activity. They don't learn much from following a detailed “recipe” of steps.*
3. *The decisions made by the class should require knowledge learned in the class. For instance, if they are designing an experiment they should consider principles of good experimental design learned in class, rather than “intuitively” deciding how to conduct the experiment.*
4. *If possible, the activity should include design, data collection and analysis so that students can see the whole process at work.*
5. *It is sometimes better to have students work in teams to discuss how to design the activity and then reconvene the class to discuss how it will be done, but it is sometimes better to have the class work together for the initial design and other decisions. It depends on how difficult the issues to be discussed are, and whether each team will need to do things in exactly the same way.*
6. *The activity should begin and end with an overview of what is being done and why.*
7. *The activity should be fun!*

Some Activities that could be improved

(1) Pepsi vs. Coke Activity

Today we will test whether Pepsi or Coke tastes better. Divide into groups of 4. Choose one person in your group to be the experimenter. *Note:* If you are not the experimenter, please refrain from looking at the front of the classroom.

- (a) On the table in the front of the classroom are two large soda bottles, one of Pepsi and one of Coke. There are also cups labeled A and B. The experimenter should go to the table and flip a coin. If it's heads, then pour Pepsi into a cup labeled A and Coke into a cup labeled B. If it's tails, pour Pepsi into cup B and Coke into cup A. Remember which is which. Bring them back to your team.
- (b) Have a team member taste both drinks. Record which one they prefer – the one in cup A or the one in cup B.
- (c) The experimenter should now reveal to the team member if it was Coke or Pepsi that was preferred.
- (d) The experimenter should repeat this process for each team member once. Then one of the other team members should give the taste test to the experimenter, so each student will have done it once.
- (e) Come together as a class. Your teacher will ask how many of you preferred Coke.
- (f) Look up the formula in your book for a confidence interval for a proportion. Construct a confidence interval for the proportion of students in the class who prefer Coke.
- (g) Do a hypothesis test for whether either drink was preferred by the class.

Critique: *The test is not double blind. There is no reason why the experimenter can't be blind to which drink is which as well. The person who initially sets up the experiment could cover or remove the labels from the drink containers, and call them drinks 1 and 2. The drinks could then be prepared in advance into cups labeled A and B. The order of presentation should be randomized for each taster.*

(2) Central Limit Theorem Activity

The purpose of this exercise is to verify the Central Limit Theorem. Remember that this Theorem tells us that the mean of a large sample is:

- Approximately bell-shaped
- Has mean equal to the mean of the population
- Has standard deviation equal to the population standard deviation/ \sqrt{n}

Please follow these instructions to verify that the Central Limit Theorem holds.

- (a) Divide into pairs. Each pair should have 1 die.
- (b) Take turns rolling the die, 25 times each, so you will have 50 rolls. Keep track of the number that lands face up each time.
- (c) Draw a histogram of the results. The die faces are equally likely, so the histogram should have a “uniform” shape. Verify that it does.
- (d) Find the mean and standard deviation for the 50 rolls.
- (e) The mean and standard deviation for rolling a single die are 3.5 and 1.708, respectively. Is the mean for your 50 rolls close to 3.5? Is the standard deviation close to 1.708?
- (f) Come together as a class. Draw the theoretical curve that the *mean* of 50 rolls should have. Remember that it's bell-shaped, and has a mean equal to the population mean, so that's 3.5 in this case, and the standard deviation in this case should be $1.708/\sqrt{50} = .24$.
- (g) Have each pair mark their mean for the 50 rolls on the curve. Notice whether or not they seem reasonable, given what is expected using the Central Limit Theorem.

Critique: *This is not a good activity for at least two reasons. First, it has absolutely no real-world motivation and reinforces the myth that statistics is boring and useless. Second, the instructions are too complete. There is no room for exploration on the part of the students; they are simply given a “recipe” to follow.*

How to improve on this activity? The “Cents and the Central Limit Theorem” activity from *Activity Based Statistics* (Scheaffer et al) provides an example for illustrating the Central Limit Theorem that is more aligned with the guidelines. Some other good examples from *Activity Based Statistics*:

- The introduction to hypothesis testing activity (where you draw cards at random from a deck and always get the same color) works well.
- Matching Graphs to Variables generates a lot of discussion and learning.
- Random Rectangles has become a standard, for good reason.
- Randomized Response is not central to the intro course, but it does involve some statistical thinking.

Additional Examples of Activities and Projects

(3) Data Gathering and Analysis: A Class of Projects

The idea for projects like the ones described here comes from Robert Wardrop's *Statistics: Learning in the Presence of Variability* (Dubuque, IA: William C. Brown, 1995). These projects, in turn, are based on a study by cognitive psychologists Kahneman and Tversky.

Consider two versions of the "General's Dilemma:"

Version 1: Threatened by a superior enemy force, the general faces a dilemma. His intelligence officers say his soldiers will be caught in an ambush in which 600 of them will die unless he leads them to safety by one of two available routes. If he takes the first route, 200 soldiers will be saved. If he takes the second, there is a two-thirds chance that 600 soldiers will be saved, and a two-thirds chance that none will be saved. Which route should he take?

Version 2: Threatened by a superior enemy force, the general faces a dilemma. His intelligence officers say his soldiers will be caught in an ambush in which 600 of them will die unless he leads them to safety by one of two available routes. If he takes the first route, 400 soldiers will die. If he takes the second, there is a one-third chance that no soldiers will die, and a two-thirds chance that 600 will die. Which route should he take?

Both versions of the question have the same two answers; both describe the same situation. The two questions differ only in their wording: one speaks of lives lost, the other of lives saved.

A pair of questions of this form leads easily to a simple randomized comparative experiment with the two questions as "treatments." Recruit a set of subjects, sort them into two groups using a random number table, and assign one version of the question to each group. The results can be summarized in a 2x2 table of counts:

Question	Answer	
	A	B
Version 1		
Version 2		

The data can be analyzed by comparing the two proportions using, e.g., Fisher's exact test or the chi-square test with continuity correction.

Exercise Set 1.2 in Wardrop's book lists a large number of variations on this structure, many of them carried out by students. Here are abbreviated versions of just four:

Ask people in a history library whether they find a particular argument from a history book persuasive; the argument was presented with and without a table of supporting data.

Ask women at the student union whether they would accept if approached by a male stranger and invited to have a drink; the male was/was not described as “attractive.”

Ask customers ordering an ice cream cone whether they want a regular or waffle cone; the waffle cone was/was not described as “homemade.”

Ask college students either (1) Would you recommend the counseling service for a friend who was depressed? Or (2) Would you go to the counseling service if you were depressed?

Projects based on two versions of a two-answer question offer a number of advantages:

- (a) Data collection can be completed in a reasonable length of time.
- (b) Randomization ensures that the results will be suitable for formal inference.
- (c) Randomization makes explicit the connection between chance in data gathering and the use of a probability model for analysis.
- (d) The method of analysis is comparatively simple and straightforward.
- (e) The structure (a 2x2 table of counts) is one with very broad applicability.
- (f) Finally, the format is very open-ended, which affords students a wide range of areas of application from which to choose, and offers substantial opportunities for imagination and originality in choosing subjects and the pair of questions.

(4) Sample Project/Activity: Team constructed questions about relationships

(Adapted from Project 2.2, Instructors' Resource Manual, *Mind On Statistics*, Utts and Heckard)

These instructions are for the teacher. Instructions for students are on the “Project 4 Team Form.”

Goal: Provide students with experience in formulating a research question, then collecting and describing data to help answer it.

Supplies: (N = number of students; T = number of teams)

- N index cards or slips of paper of each of T colors (or use board space; see below)
- T or 2T overhead transparencies and pens (see Step 3 for the reason for 2T of them)
- T calculators

Students should work in teams of 4 to 6. See the “Sample Project 4 Team Form” below.

Step 1: Each team formulates two categorical variables for which they want to know if there is relationship, such as whether someone is a firstborn (or only) child and whether they prefer indoor or outdoor activities (recent research suggests that firstborns prefer indoor activities and later births prefer outdoor activities); male/female and opinion on something; class (senior, junior, etc) and whether they own a car, etc. To make it easier to finish in time, you may want to restrict them to two categories per variable.

There are two possible methods for collecting data – using index *cards* (or paper) or using the *board*. Each of the next few steps will be described for both methods.

Step 2: *Cards:* Each team is assigned a color, from the T colors of index cards. For instance Team 1 might be blue, Team 2 is pink, and so on. *Board:* Assign each team space on the chalkboard to write their questions.

Step 3: Each team asks the whole class its two questions. *Cards:* The team writes the questions on an overhead transparency and displays them, with each team taking a turn to go to the front of the room. Students write their answers on the index card corresponding to that team's color and the team collects them. For instance, all students in the class write their answers to Team 1's questions on the blue index card, their answers to Team 2's questions on the pink card, and so on. *Board:* A team member writes the questions on the board along with a two-way table where each student can put a hash mark in the appropriate cell.

Step 4: *Cards:* After each team has asked its questions and students have written their answers, the cards are collected and given to the appropriate team. For instance, Team 1 receives all the blue cards. *Board:* All class members go to each segment of the board and put a hash mark in the cell of the table that fits them.

Step 5: Each team tallies, summarizes and prepares a graphical display of the data for their questions. The results are written on an overhead transparency.

Step 6: Each team presents the results to the class.

Step 7: Results can be retained for use when covering chi-square tests for independence if you are willing to pretend that the data are a random sample from a larger population.

NOTE: This can also be done with one categorical and one quantitative variable, and the data retained for use when doing two-sample inference.

PROJECT 4: TEAM FORM

TEAM MEMBERS:

- | | |
|----------|----------|
| 1. _____ | 4. _____ |
| 2. _____ | 5. _____ |
| 3. _____ | 6. _____ |

INSTRUCTIONS:

1. Create two categorical variables for which you think there might be an interesting relationship for class members. If you prefer, you can turn a quantitative variable into a categorical one, such as GPA - high or low (using a cutoff like ≥ 3.0). Each variable should have 2 categories, to make it easier to finish in the allotted time.
2. List the two variables below, designating which is the explanatory variable and which is the response variable, if that makes sense for your situation.

Explanatory variable:

Response variable:

3. Each team will be assigned one segment of the chalk board. One team member is to go to the board and write your two questions. Also, write a “two-way” table on the board in which people will put a “hash mark” into the square that describes them.
4. Everyone will now go to the board and fill in a hash mark in the appropriate box for *each* team’s set of questions.
5. After everyone has gone to the board and filled in all of their data, enter the totals in the table below for your team’s questions. Also enter what the categories are for each variable.

Explanatory Variable	Response Variable		Total
	Category 1:	Category 2:	
Category 1:			
Category 2:			
Total			

6. Create appropriate numerical and graphical summaries to display on your team's overhead transparency. Write a brief summary of your findings below and on the back if needed.
7. A member of each team will present the team’s result to the class, using the overhead transparency.
8. Turn in this sheet and the overhead transparency sheet.

(5) Sample Project/Activity: Comparing Manual Dexterity under Two Conditions

(Adapted from Project 12.2, Instructors' Resource Manual, *Mind On Statistics*, Utts and Heckard)
These instructions are for the teacher. Instructions for students are on the "Project 5 Team Form."

Goal: Provide students with experience in designing, conducting and analyzing an experiment.

Supplies: (N = number of students, T = number of teams)

- T bowls filled with about 30 of each of two distinct colors of dried beans
- $2T$ empty paper cups or bowls
- T stop watches or watches with second hand

NOTE: A variation is to have them do the task with and without wearing a latex glove instead of with the dominant and non-dominant hand. In that case you will need N pairs of latex gloves.

The Story: A company has many workers whose job is to sort two types of small parts. Workers are prone to get repetitive strain injury, so the company wonders if there would be a big loss in productivity if the workers switch hands, sometimes using their dominant hand and sometimes using their non-dominant hand. (Or if you are using latex gloves, the story can be that for health reasons they might want to require gloves.) Therefore, you are going to design, conduct and analyze an experiment making this comparison. Students will be timed to see how long it takes to separate the two colors of beans by moving them from the bowl into the two paper cups, with one color in each cup. A comparison will be done after using dominant and non-dominant hands. An alternative is to time students for a fixed time, like 30 seconds, and see how many beans can be moved in that amount of time.

Step 1: As a class, discuss how the experiment will be done. This could be done in teams first. See below for suggestions.

1. What are the treatments? What are the experimental units?
2. Principles of experimental design to consider are as follows. Use as many of them as possible in designing and conducting this experiment. Discuss why each one is used.
 - a. Blocking or creating matched-pairs
 - b. Randomization of treatments to experimental units, or randomization of order of treatments
 - c. Blinding or double blinding
 - d. Control group
 - e. Placebo
 - f. Learning affect or getting tired
3. What is the parameter of interest?
4. What type of analysis is appropriate – hypothesis test, confidence interval or both?

The class should decide that each student will complete the task once with each hand. Why is this preferable to randomly assigning half of the class to use their dominant hand and the other half to use their non-dominant hand? How will the order be decided? Should it be the same for all students? Will practice be allowed? Is it possible to use a single or double blind procedure?

Step 2: Divide into teams and carry out the experiment.

The Project 5 Team Form shows one way to assign tasks to team members.

Step 3: Descriptive statistics and preparation for inference

Convene the class and create a stemplot of the differences. Discuss whether the necessary conditions for this analysis are met. Were there any outliers? If so, can they be explained? Have someone compute the mean and standard deviation for the differences.

Step 4: Inference

Have teams reconvene. Each team is to find a confidence interval for the mean difference and conduct the hypothesis test.

Step 5: Reconvene the class and discuss conclusions

Suggestions for how to design and analyze the experiment in sample Project 5:

Design issues:

- a. Blocking or creating matched-pairs
Each student should be used as a matched pair, doing the task once with each hand.
- b. Randomization of treatments to experimental units, or randomization of order of treatments
Randomize the order of which hand to use for each student.
- c. Blinding or double blinding
Obviously the student knows which hand is being used, but the time-keeper doesn't need to know.
- d. Control group
Not relevant for this experiment.
- e. Placebo
Not relevant for this experiment.
- f. Learning affect or getting tired
There is likely to be a learning effect, so you may want to build in a few practice rounds. Also, randomizing the order of the two hands for each student will help with this.

One possible design: Have each student flip a coin. Heads, start with dominant hand. Tails, start non-dominant hand. Time them to see how long it takes to separate the beans. The person timing them could be blind to the condition by not watching.

Analysis:

What is the parameter of interest?

Answer: Define the random variable of interest for each person to be a "manual dexterity difference" of

$$d = \text{number of extra seconds required with non-dominant hand} \\ = \text{time with non-dominant hand} - \text{time with dominant hand.}$$

Define μ_d = population mean manual dexterity difference.

What are the null and alternative hypotheses?

$H_0 : \mu_d = 0$ and $H_a: \mu_d > 0$ (faster with dominant hand)

Is a confidence interval appropriate?

Yes, it will provide information about how much faster workers can accomplish the task with their dominant hands. The formula for the confidence interval is

$$\bar{d} \pm t^* \frac{s_d}{\sqrt{n}}$$

where t^* is from the t-table with $df = n-1$, and s_d is the standard deviation of the difference scores.

To carry out the test, compute $t = \frac{\bar{d} - 0}{s_d/\sqrt{n}}$ then compare to the t-table to find the p-value.

PROJECT 5 TEAM FORM

TEAM MEMBERS:

1. _____ 4. _____
 2. _____ 5. _____
 3. _____ 6. _____

INSTRUCTIONS:

You will work in teams. Each team should take a bowl of beans and two empty cups. You are each going to separate the beans by moving them from the bowl to the empty cups, with one color to each cup. You will be timed to see how long it takes. You will each do this twice, once with each hand, with order randomly determined.

1. Designate these jobs. You can trade jobs for each round if you wish.
Coordinator – runs the show.
Randomizer – flips a coin to determine which hand each person will start with, separately for each person.
Time keeper – must have watch with second hand. Times each person for the task.
Recorder – records the results in the table below.
2. Choose who will go first. The *randomizer* tells the person which hand to use first. Each person should complete the task once before moving to the 2nd hand for the first person. That gives everyone a chance to rest between hands.
3. The *time keeper* times the person, while they move the beans **one at a time** from the bowl to the cups, separating colors.
4. The *recorder* notes the time and records it in the table below.
5. Repeat this for each team member.
6. Each person then goes a second time, with the hand *not* used the first time.
7. Calculate the difference for each person.

NAME:	Time for <i>non-dominant</i> hand.	Time for <i>dominant</i> hand.	$d = \text{difference}$ = non-dominant – dominant hand

RESULTS FOR THE CLASS:

Record the data here:

Parameter to be tested and estimated is:

Confidence interval:

Hypothesis test – hypotheses and results:

B. Examples of assessment items

Assessment items to avoid using on tests: True/False, pure computation without a context or interpretation, items with too much data to enter and compute or analyze, items that only test memorization of definitions or formulas.

We first give some examples of assessment items with problems and commentary about the nature of the difficulty

(1) A teacher taught two sections of elementary statistics last semester, each with 25 students, one at 8am and one at 4pm. The means and standard deviations for the final exams were 78 and 8 for the 8am class, and 75 and 10 for the 4pm class. In examining these numbers, it occurred to the teacher that the better students probably sign up for 8am classes instead of 4pm classes. So she decided to test whether or not the mean final exam scores were equal for her two groups of students. State the hypotheses and carry out the test.

Critique: The teacher has all of the population data so there is no need to do statistical inference.

(2) An economist wants to compare the mean salaries for male and female CEOs. He gets a random sample of 10 of each and does a t-test. The resulting p -value is .045.

- (a) State the null and alternative hypotheses.
- (b) Make a statistical conclusion.
- (c) State your conclusion in words that would be understood by someone with no training in statistics.

Critique: The question doesn't address the conditions necessary for a t-test, and with the small sample sizes they are almost surely violated here. Salaries are almost surely skewed.

(3) Which of the following gives the definition of a p -value?

- (a) It's the probability of rejecting the null hypothesis when the null hypothesis is true.
- (b) It's the probability of not rejecting the null hypothesis when the null hypothesis is true.
- (c) It's the probability of observing data as extreme as that observed.
- (d) It's the probability that the null hypothesis is true.

Critique: None of these answers is quite correct. Answers (b) and (d) are clearly wrong; answer (a) is the level of significance and answer (c) would be correct if it continued "... or more extreme, given that the null hypothesis is true."

Examples showing ways to improve some assessment items:

True/false items, even when well written, do not provide much information on student knowledge because there is always a 50% chance of getting the item right without any knowledge of the topic. One current approach is to change the items into forced-choice questions with three or more options. For example,

(4) The size of the standard deviation of a data set depends on where the center is. True or False

Changed to:

(4) Does the size of the standard deviation of a data set depend on where the center is located?

- (a) Yes, the higher the mean, the higher the standard deviation.
- (b) Yes, because you have to know the mean to calculate the standard deviation.
- (c) No, the size of the standard deviation is not affected by the location of the distribution.
- (d) No, because the standard deviation only measures how the values differ from each other, not how they differ from the mean.

(5) A correlation of +1 is stronger than a correlation of -1. True or False

Rewritten as:

(5) A recent article in an educational research journal reports a correlation of +.8 between math achievement and overall math aptitude. It also reports a correlation of -.8 between math achievement and a math anxiety test. Which of the following interpretations is the most correct?

- (a) The correlation of +.8 indicates a stronger relationship than the correlation of -.8
- (b) The correlation of +.8 is just as strong as the correlation of -.8
- (c) It is impossible to tell which correlation is stronger

Context is important for helping students see and deal with statistical ideas in real world situations.

(6) Once it is established that X and Y are highly correlated, what type of study needs to be done in order to establish that a change in X causes a change in Y?

A context is added:

(6) A researcher is studying the relationship between an experimental medicine and T4 lymphocyte cell levels in HIV/AIDS patients. The T4 lymphocytes, a part of the immune system, are found at reduced levels in patients with the HIV infection. Once it is established that the two variables, dosage of medicine and T4 cell levels, are highly correlated, what type of study needs to be done in order to establish that a change in dosage causes a change in T4 cell levels?

- (a) correlational study
- (b) controlled experiment
- (c) prediction study
- (d) survey

Try to avoid repetitious/tedious calculations on exams that may become the focus of the problem for the students at the expense of concepts and interpretations.

(7) A First Year Program course used a final exam that contained a 20 point essay question that asked students to apply Darwinian principles to analyze the process of expansion in major league sports franchises. To check for consistency in grading among the four professors in the course a random sample of six graded essays were selected from each instructor. The scores are summarized in the table below. Construct an ANOVA table to test for a difference in means among the four instructors.

Instructor	Scores					
Affinger	18	11	10	12	15	12
Beaulieu	14	14	11	14	11	14
Cleary	19	20	15	19	19	16
Dean	17	14	17	15	18	15

Critique: The version of the question above requires a fair amount of pounding on the calculator to get the results and never even asks for an interpretation. The revision below still requires some calculation (which can be adjusted depending on the amount of computer output provided) but the calculations can be done relatively efficiently - especially by students who have a good sense of what the computer output is providing.

(7) A First Year Program course ... (same intro as above) ... The scores are summarized in the table below, along with some *Descriptive Statistics* for the entire sample and a portion of the *Oneway ANOVA* output.

Descriptive Statistics						
Variable	N	Mean	Median	TrMean	StDev	SEMean
Score	24	15.000	15.000	15.000	2.919	0.596

One-way Analysis of Variance
 *** ANOVA TABLE OMITTED ***

Instructor	Scores					
Affinger	18	11	10	12	15	12
Beaulieu	14	14	11	14	11	14
Cleary	19	20	15	19	19	16
Dean	17	14	17	15	18	15

Individual 95% CIs For Mean Based on Pooled StDev						
Level	N	Mean	StDev			
Afinger	6	13.000	2.966	-----+-----		
Beaulieu	6	13.000	1.549	(------*-----)		
Cleary	6	18.000	2.000	(------*-----)		
Dean	6	16.000	1.549	(------*-----)		
Pooled StDev = 2.098				12.5	15.0	17.5 20.0

- Unfortunately, we are missing the ANOVA table from the Minitab output. Use the information given above to construct the ANOVA table and conduct a test (5% level) for any significant differences among the average scores assigned by the four instructors. *Be sure to include hypotheses and a conclusion.* If you have trouble getting one part of the table that you need to complete the rest (or the next question), make a reasonable guess or ask for assistance (for a small point fee).
- After completing the ANOVA table, construct a 95% confidence interval for the average score given by Dr. Affinger. Note: Your answer should be consistent with the graphical display given by Minitab.

Some additional examples of good assessment items

(8) Let Y denote the amount a student spends on textbooks for one semester. Suppose Nancy, who is statistically savvy, wants to know how fall, semester 1, and spring, semester 2, compare. In particular, suppose she is interested in the averages μ_1 and μ_2 . You may assume that Nancy has taken several statistics courses and knows a lot about statistics, including how to interpret confidence intervals and hypothesis tests. You have random samples from each semester and are to analyze the data and write a report. You seek advice from 4 persons:

Rudd says "Conduct an $\alpha=.05$ test of $H_0: \mu_1=\mu_2$ vs. $H_A: \mu_1\neq\mu_2$ and tell Nancy whether or not you reject H_0 ."

Linda says "Report a 95% confidence interval for $\mu_1 - \mu_2$."

Steve says "Conduct a test of $H_0: \mu_1=\mu_2$ vs. $H_A: \mu_1\neq\mu_2$ and report to Nancy the p-value from the test."

Gloria says "Compare \bar{y}_1 to \bar{y}_2 . If $\bar{y}_1 > \bar{y}_2$ then test $H_0: \mu_1=\mu_2$ vs. $H_A: \mu_1>\mu_2$ using $\alpha=.05$ and tell Nancy whether or not you reject H_0 . If $\bar{y}_1 < \bar{y}_2$ then test $H_0: \mu_1=\mu_2$ vs. $H_A: \mu_1<\mu_2$ using $\alpha=.05$ and tell Nancy whether or not you reject H_0 ."

Rank the 4 pieces of advice from worst to best and explain why you rank them as you do. That is, explain what makes one better than another.

(9) Researchers took random samples of subjects from two populations and applied a Wilcoxon-Mann-Whitney test to the data; the P-value for the test, using a non-directional alternative, was .06. For each of the following, say whether the statement is True or False and say why.

- There is a 6% chance that the two population distributions really are the same.
- If the two population distributions really are the same, then a difference between the two samples as extreme as the difference that these researchers observed would only happen 6% of the time.
- If a new study were done that compared the two populations, there is a 6% probability that H_0 would be rejected again.
- If $\alpha = .05$ and a directional alternative were used, and the data departed from H_0 in the direction specified by the alternative hypothesis, then H_0 would be rejected.

(10) An article on the CNN web page on Monday

(<http://www.cnn.com/HEALTH/9612/16/faith.healing/index.html>) begins with the sentence "Family doctors overwhelmingly believe that religious faith can help patients heal, according to a survey released Monday." Later the article states "Medical researchers say the benefits of religion may be as simple as helping the immune system by reducing stress" and Dr. Harold Koenig is reported to

say that "people who regularly attend church have half the rate of depression of infrequent churchgoers."

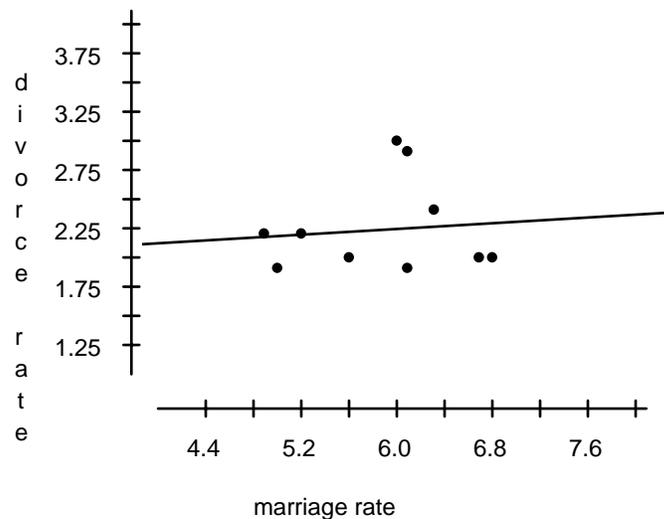
Use the language of statistics to critique the statement by Dr. Koenig and the claim, suggested by the article, that religious faith and practice help people fight depression. You will want to select some of the following words in your critique: observational study, experiment, blind, double-blind, precision, bias, sample, spurious, confounding, causation, association, random, valid, reliable.

(11) Francisco Franco (Class of '98) weighed 100 Hershey's Kisses (with almonds). He found that the sample average was 4.80 grams and the SD was .28 grams. In the context of this setting, explain what is meant by the sampling distribution of an average.

(12) A gardener wishes to compare the yields of three types of pea seeds - type A, type B, and type C. She randomly divides the type A seeds into three groups and plants some in the east part of her garden, some in the central part of the garden, and some in the west part of the garden. Then she does the same with the type B seeds and with the type C seeds.

- (a) What kind of experimental design is the gardener using?
- (b) Why is this kind of design used in this situation? (Explain in the context of the situation.)

(13) The following scatterplot shows how divorce rate, y , and marriage rate, x , are related for a collection of 10 countries. The regression line has been added to the plot.



- (a) The U.S. is not one of the 10 points in the original collection of countries. It happens that the U.S. has a higher marriage rate than any of the 10 countries. Moreover, the divorce rate for the U.S. is higher than one would expect, given the pattern of the other countries. How would adding the U.S. to the data set affect the regression line? Why?

- (b) Think about the scatterplot and regression line after the U.S. has been added to the data set. Provide a sketch of the residual plot. Label the axes and identify the U.S. on your plot with a triangle.

(14) Researchers wanted to compare two drugs, formoterol and salbutamol, in aerosol solution, to a placebo for the treatment of patients who suffer from exercise-induced asthma. Patients were to take a drug or the placebo, do some exercise, and then have their "forced expiratory volume" measured. There were 30 subjects available. (Based on A.N. Tsoy, et al., *European Respiratory Journal* 3 (1990): 235; via Berry, *Statistics: A Bayesian Perspective*.)

- (a) Should this be an experiment or an observational study? Why?
(b) Within the context of this setting, what is the placebo effect?
(c) Briefly explain how to set up a randomized blocks design (RBD) here.
(d) How would an RBD be a helpful? That is, what is the main advantage of using a RBD in a setting like this?

(15) I noticed that 8 students from the 114 class attended the review session prior to the second exam (in April). The average score among those 8 students was lower than the average for the 21 students who did not attend the review session. Suppose I want to use this information in a study of the effectiveness of review sessions.

- (a) What kind of study is this: observational or experimental? Why?
(b) What kind(s) of sampling error(s) or bias(es) might be of concern here?
(c) (Hypothetical) I gave the data for the 8 who attended and for the 21 who did not attend to my friend George. George used the data to conduct a hypothesis test. Does a hypothesis test make sense? If so, what is H_0 ? If not, why not?

(16) For each of the following three settings, state the type of analysis you would conduct (e.g., one-sample t-test, regression, Chi-square test of independence, Chi-square goodness-of-fit test, etc.) if you had all of the raw data and specify the roles of the variable(s) on which you would perform the analysis, but do not actually carry out the analysis.

- (a) Elizabeth Larntz (Class of '02) measured the effect of exercise on pulse for each of 13 students. She measured pulse before and after exercise (doing 30 jumping jacks) and found that the average change was 55.1 and the SD of the changes was 18.4. How would you analyze the data?
(b) Three HIV treatments were tested for their effectiveness in preventing progression of HIV in children. Of 276 children given drug A, 259 lived and 17 died. Of 281 children given drug B, 274 lived and 7 died. Of 274 children given drug C, 264 lived and 10 died. How would you analyze the data?
(c) A researcher was interested in the relationship between forearm length and height. He measured the forearm lengths and heights of each of 16 women. How would you analyze these data?

(17) I had Data Desk construct parallel dotplots of the data from four samples. I then conducted a test of $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ and rejected H_0 at the $\alpha = .05$ level. I also tested $H_0: \mu_1 = \mu_2 = \mu_3$ and rejected H_0 at the $\alpha = .05$ level. However, when I tested $H_0: \mu_2 = \mu_3$ using $\alpha = .05$ I did not reject H_0 . Likewise, when I tested $H_0: \mu_1 = \mu_4$ using $\alpha = .05$ I did not reject H_0 .

- (a) Your job is to sketch a graph of the parallel dotplots of the data. That is, based on what I told you about the tests you should have an idea of how the data look. Use that idea to draw a graph. Indicate the sample means with triangles that you add to the dotplots.
- (b) It is possible to get data with the same sample means that you graphed in part (a), but for which the hypothesis $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ is not rejected at the $\alpha = .05$ level. Provide a graph of this situation. That is, keep the same sample means (triangles) you had from part (a), but show how the data would have been different if H_0 were not to be rejected.

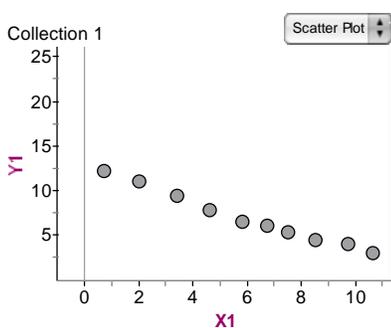
(18) Atley Chock (Class of '02) collected data on a random sample of 12 breakfast cereals. He recorded x = fiber (in grams/ounce) and y = price (in cents/ounce). A scatterplot of the data shows a linear relationship. The fitted regression model is

$$\hat{y} = 17.42 + 0.62X$$

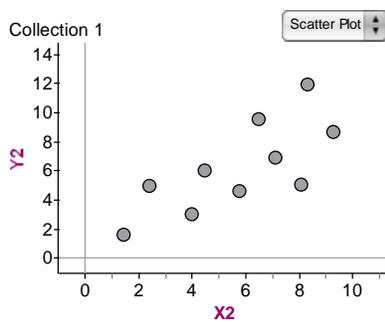
The sample correlation coefficient, r , is .23. The SE of b_1 is .81. Also, $s_{y|x} = 3.1$.

- (a) Find r^2 and interpret r^2 in the context of this problem.
- (b) Suppose that a cereal has 2.63 grams of fiber/ounce and costs 17.3 cents/ounce. What is the residual for this cereal?
- (c) Interpret the value of $s_{y|x}$ in the context of this problem. That is, what does it mean to say that $s_{y|x} = 3.1$?
- (d) In the context of this problem explain what is meant by "the regression effect."

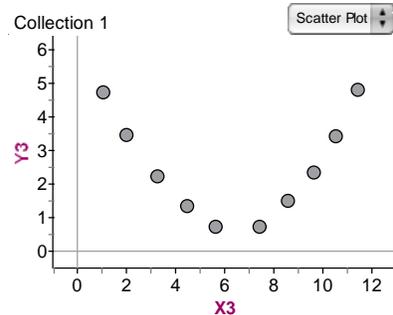
(19) Give a **rough** estimate of the sample correlation for the data in each of the scatterplots below.



$r \approx$ _____



$r \approx$ _____



$r \approx$ _____

(20) A matched pairs experiment compares the taste a regular cheese pizza of Pizza Joe's to Domino's. Each subject tastes two unmarked pieces of pizza, one of each type, in random order and states which he or she prefers. Of the 50 subjects who participate in the study, 21 prefer Pizza Joe's.

- (a) Find a 96% confidence interval for the proportion of the population who prefers Pizza Joe's to Domino's.
- (b) How large a sample is required if the desired margin of error of the confidence interval is 4%?

(21) It was claimed that 1 out of 5 cardiologists takes an aspirin a day to prevent hardening of the arteries. Suppose that the claim is true. If 1500 cardiologists are selected at random, what is the probability that at least 275 of the 1500 take an aspirin a day?

(22) Identify whether a scatterplot would or would not be an appropriate visual summary of the relationship between the variables. In each case, explain your reasoning.

- (a) Blood pressure and age
- (b) Region of country and opinion about stronger gun control laws
- (c) Verbal SAT and math SAT score
- (d) Handspan and gender (male or female)

(23) The paragraphs that follow each describe a situation which calls for some type of statistical analysis. For each you should:

- (i) Give the **name** of an appropriate statistical procedure to apply (from the list below). You may use the same procedure more than once and some questions might have more than one correct answer.
- (ii) In some problems you will also be given a p-value. Use it to reach a conclusion *for that specific problem*. Be sure to say something more than just Reject H_0 or Fail to Reject H_0 . (Assume a 5% significance level)

Some statistical procedures you might choose:

- | | |
|---|---------------------------------|
| Confidence interval (for a mean, p, ...) | Normal distribution |
| Determining sample size | Correlation |
| Test for a mean | Simple linear regression |
| Test for proportion | Multiple regression |
| Difference in means (paired data) | Two-way table (Chi-square test) |
| Difference in means (two independent samples) | ANOVA for difference in means |
| Difference in proportions | Two-way ANOVA for means |

(a) Anthropologists have found two burial mounds in the same region. They know that several different tribes lived in the region and that the tribes have been classified according to different lengths of skulls. They measure a random sample of skulls found in each burial mound and wish to determine if the two mounds were made by different tribes. (p-value = 0.0082)

(b) The Hawaiian Planters Association is developing three new strains of pineapple (call them A, B, and C) to yield pulp with higher sugar content. Twenty plants of each variety (60 plants in all) are randomly distributed into a two acre field. After harvesting, the resulting pineapples are measured for sugar content and the yields are recorded for each strain. Are there significant differences in average sugar content between the three strains? (p-value = 0.987)

(c) Researchers were commissioned by the Violence In Children's Television Investigative Monitors (VICTIM) to study the frequency of depictions of violent acts in Saturday morning TV fare. They selected a random sample of 40 shows which aired during this time period over a twelve week period. Suppose that 28 of the 40 shows in the sample were judged to contain scenes depicting overtly violent acts. How should they use this information to make a statement about the population of all Saturday morning TV shows?

(d) The Career Planning Office is interested in seniors' plans and how they might relate to their majors. A large number of students are surveyed and classified according to their MAJOR (Natural Science, Social Science, Humanities) and FUTURE plans (Graduate School, Job, Undecided). Are the type of major and future plans related? (p-value = 0.047)

(e) Sophomore Magazine asked a random sample of 15 year olds if they were sexually active (yes or no). They would like to see if there is a difference in the responses between boys and girls. (p-value = 0.029)

(f) Every week during the Vietnam War, a body count (number of enemy killed) was reported by each army unit. The last digits of these numbers should be fairly random. However, suspicions arose that the counts might have been fabricated. To test this, a large random sample of body count figures was examined and the frequency with which the last digit was a 0 or a 5 was recorded. Psychologists have shown that people making up their own random numbers will use these digits less often than random chance would suggest (i.e. 103 sounds like a more "real" count than 100). If the data were authentic counts, the proportion of numbers ending in 0 or 5 should be about 0.20. (p-value=0.002)

(g) In one of his adventures, Sherlock Holmes found footprints made by the criminal at the scene of a crime and measured the distance between them. After sampling many people, measuring their height and length of stride, he confidently announced that he could predict the height of the suspect. How?

(24) How accurate are radon detectors of a type sold to homeowners? To answer this question, university researchers placed 12 detectors in a chamber that exposed them to 105 picocuries per liter of radon. The detector readings found are below. A printout of the descriptive statistics from Minitab follows.

91.9 97.8 111.4 122.3 105.4 95.0
 103.8 99.6 96.6 119.3 104.8 101.7

Variable	N	Mean	Median	TrMean	StDev	SE Mean
readings	12	104.13	102.75	103.54	9.40	2.71

Variable	Minimum	Maximum	Q1	Q3
readings	91.90	122.30	96.90	109.90

- (a) Is there convincing evidence that the mean 20 readings of all detectors of this type differs from the true value of 105? Perform the appropriate hypothesis test with $\alpha = .05$.
- (b) What is the Type I error associated with this problem?
- (c) What is the Type II error associated with this problem?
- (d) What is the probability of a type II error if the reading of the detectors is too low by 5 picocuries (really 100 when it should read 105)?

(25) According a Food and Drug Administration (FDA) study, a cup of coffee contains an average of 115 mg of caffeine, with the amount per cup ranging from 60 to 180 mg depending on the brewing method. Suppose you want to repeat the FDA experiment to obtain an estimate of the mean caffeine content to within 5 mg with 95% using your favorite brewing method. In problems such as this, we can estimate the standard deviation of the population to be $\frac{1}{4}$ of the range. How many cups of coffee must you brew?

(26) An advertisement claims that by applying a particular drug, hair is restored to bald headed men. Outline the design of an experiment that you would use to examine this claim. Assume that you have money to use 20 bald men in this experiment.

(27) A study of iron deficiency among infants compared samples of infants following different feeding regimens. One group contained breast-fed infants, while the children in another group were fed a standard baby formula without any iron supplements. Here are the summary results on blood hemoglobin levels at 12 months of age:

Group	n	\bar{x}	s
Breast-fed	23	13.3	1.7
Formula	19	12.4	1.8

Assume that the blood hemoglobin levels in children (both breast-fed and formula fed) are normally distributed. Do a significance test to determine the statistical significance of the observed difference.

(28) Which implies a stronger linear relationship, a correlation of $+0.4$ or a correlation of -0.6 ? Briefly explain your choice.

(29) A group of physicians subjected the polygraph to the same careful testing given to medical diagnostic test. They found that if 1000 people were subjected to the polygraph and 500 told the truth and 500 lied, the polygraph would indicate that approximately 185 of the truth tellers were liars and 120 of the liars were truth-tellers. In the application of the polygraph test, an individual is presumed to be a truth-teller until indicated that s/he is a liar. What is a Type I error in the context of this problem? What is the probability of a Type I error in the context of this problem? What is a Type II error in the context of this problem? What is the probability of a Type II error in the context of this problem?

(30) Audiologists have recently developed a rehabilitation program for hearing-impaired patients in a Canadian program for senior citizens. A simple random sample of the 30 residents of a particular senior citizens home and the seniors were diagnosed for degree and type of sensorineural hearing loss which was coded as follows: 1 = hear within normal limits, 2 = high-frequency hearing loss, 3 = mild loss, 4 = mild-to-moderate loss, 5 = moderate loss, 6 = moderate-to-severe loss, and 7 = severe-to-profound loss. The data are as follows:

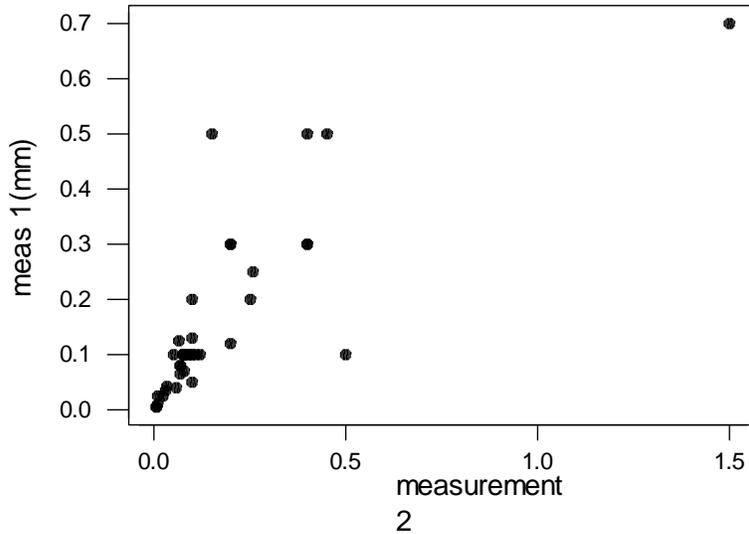
6 7 1 1 2 6 4 6 4 2 5 2 5 1 5
4 6 6 5 5 5 2 5 3 6 4 6 6 4 2

- Create a boxplot of the data.
- Give a good description of the data.
- Find a 95% confidence interval for the mean hearing loss of senior citizens in this Canadian program. The mean and standard deviation of the above data are 4.2 and 1.808 respectively. Interpret the interval.

(31) A utility company was interested in knowing if agricultural customers would use less electricity during peak hours if their rates were different during those hours. Customers were randomly assigned to continue to get standard rates or to receive the time-of-day structure. Special meters were attached that recorded usage during peak and off-peak hours; the technician who read the meter did not know what rate structure each customer had.

- Is this an observation study or experiment? Defend your answer.
- What is the explanatory and response variable?
- List a potential confounding variable in this work.
- Is this a matched pair design? Defend your answer.

(32) At the beginning of the semester, we measured the width of a page in our statistics book. Below is the scatterplot of first measurement vs. the second measurement.



- (a) Describe the distribution.
- (b) Estimate the correlation with and without the circled point.

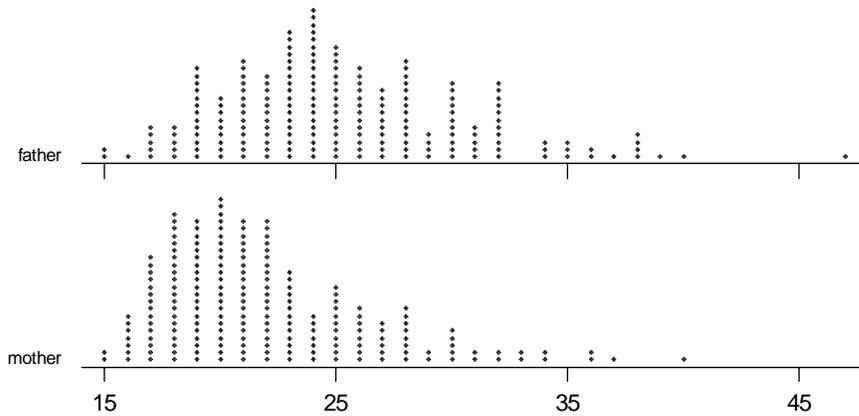
(33) A study in the Journal of Leisure Research investigated the relationship between academic performance and leisure activities. Each in a sample of 159 high school students was asked to state how many leisure activities they participated in weekly. From the list, activities that involved reading, writing or arithmetic were labeled “academic leisure activities.” Some of the results are as follows:

	mean	standard deviation
GPA	2.96	0.71
Number of leisure activities	12.38	5.07
Number of academic leisure activities	2.77	1.97

Based on these numbers (and knowing that the GPA is a value between 0 and 4 and number of activities can not be negative) discuss the potential skewness of each of the above variables.

(34) Events A and B are disjoint. Discuss whether or not A and B can be independent.

(35) A sample of 200 mothers and a sample of 200 fathers were taken. The age of the mother when she had her first child and the age of the father when he had his first child were recorded. Below are the dotplots



- Describe the data for the mother's age.
- Describe the data for the father's age.
- Compare the distributions.
- A suggestion is made to check the correlation between the ages if we wish to compare the two populations. Is this a good suggestion? Why or why not?

C. Example of using technology

This example starts with a real world situation, has students do a physical simulation using cards and then brings in computer technology to automate the simulation

A technology-based simulation to examine the effectiveness of treatments for cocaine addiction

A study on the treatment of cocaine addiction described the results of an experiment comparing two drugs for helping addicts stay off cocaine (D.M. Barnes, “Breaking the Cycle of Cocaine Addiction”, Science, Vol. 241, 1988, pp. 1029-1030). A group of 48 cocaine addicts who were seeking treatment were randomly divided into two groups of 24. One group was treated with a new drug called *desipramine*, while the other group was given *lithium*. The results are summarized in the table below where we consider patients who do not relapse as successfully treated.

	No Relapse	Relapse
Desipramine	14	10
Lithium	6	18

While we observe that desipramine was more successful than lithium in this particular experiment, can we conclude that the improvement is statistically significant? i.e. Would we expect to see such a large difference if the drugs were equally effective and it was just the random assignment process that happened to get so many more successful cases in the desipramine group? We will address this question through simulation, first using a physical demonstration based on shuffling cards, then with a computer simulation that allows us to see the differences for many random assignments of the addicts to the treatment groups.

Physical simulation

Take a deck of 54 playing cards (including two jokers) and remove 6 of the black cards (spades or clubs). The remaining deck should match the subjects in the cocaine experiment with all of the red cards and the jokers representing patients who relapsed and the 20 black cards representing patients who were treated successfully. If we shuffle the deck and deal out two piles of 24 cards each, we will simulate the assignment of addicts to the two treatment groups when the success does not depend on which drug they take. Do so and fill in the 2-way table with the “success” (black cards) and “relapse” (red/jokers) counts for each group.

	No Relapse	Relapse
Desipramine		
Lithium		

Note that once you know one number in the table, you can fill in the rest, since you know there are 24 in each treatment group and 20 will not relapse, while 28 will relapse (that is why we sometimes say there is just one degree of freedom in the 2x2 table). To keep things simple then, we can just keep track of one count, such as the number of “no relapse” in the desipramine group.

Shuffle all the cards again, deal 24 for the desipramine group and count the number of black cards.
Number of “no relapse” in desipramine group = _____

Pool the results for your class (counting # of black cards in each random group of 24 cards assigned to the “desipramine” group) in a dotplot. How often was the number black cards as large (or larger) than the 14 cases that were observed in the actual experiment?

The p-value of the original data is the proportion, assuming both drugs are equally effective, of random assignments that have 14 or more “no relapse” cases going to the desipramine group. Estimate this proportion using the data in your class dotplot.

Computer Simulation

To get a more accurate estimate of the proportion of random assignments that put 14 or more no relapse cases into the desipramine group, we’ll turn to a computer simulation.

Start with a dataset (provided online) consisting of 2 columns and 48 rows. The first column (Treatment) has the value “desipramine” in the first 24 rows and “lithium” in the remaining 24 rows. The second column (Result) has the values “no relapse” and “relapse” to match the data in the original 2x2 table from the cocaine experiment.

Have the computer permute the values in the “Result” column to represent a new random assignment of subjects to the treatment groups where the outcome does not depend on which drug was taken. Count the number of “no relapse” cases in the desipramine treatment group and have the result stored somewhere. Automate this process to repeat 1000 times*.

Look at a histogram or dotplot of the distribution of counts for the 1000 simulations. Does it seem unusual to have as many as 14 “no relapse” cases in the desipramine group?

Count the number of simulations that have 14 or more successes in the desipramine group (either from the graph if feasible or by sorting the simulated counts column) and divide by 1000 to get another approximation of the p-value for the original data.

Does it seem reasonable that the larger number (14) of successful cases appeared in the desipramine group by random chance or would it be more appropriate to conclude that desipramine probably works better than lithium at treating cocaine addiction?

**Some technology alternatives:* The most difficult step here is to automate the simulations to record the counts for many random assignments. Some packages, such as Fathom, have easy to use tools designed for exactly such purposes. Others, such as Minitab, allow a bit of programming through macros which can be built in advance and repeated in a loop. A somewhat less enlightening simulation could be accomplished with a stat package that allows generation of random data from a hypergeometric distribution, although students would then lose the connection to the physical randomizations. Finally, an ambitious instructor could construct (or possibly find on the web) an applet to perform the required simulations and collect the results.

D. Examples of naked, realistic and real

(1) Naked data (not recommended)

Find the least squares line for the data below. Use it to predict Y when X=5.

X	1	2	3	4	6	8
Y	3	4	6	7	14	20

Critique: Made-up data with no context (not recommended). The problem is purely computational with no possibility of meaningful interpretation.

(2) Realistic data (better, but still not the best)

The data below show the number of customers in each of six tables at a restaurant and the size of the tip left at each table at the end of the meal. Use the data to find a least squares line for predicting the size of the tip from the number of diners at the table. Use your result to predict the size of the tip at a table that has five diners.

Diners	1	2	3	4	6	8
Tip	\$3	\$4	\$6	\$7	\$14	\$20

Critique: A context has been added which makes the exercise more appealing and shows students a practical use of statistics.

(3) Real data (recommended)

The data below show the quiz scores (out of 20) and the grades on the midterm exam (out of 100) for a sample of eight students who took this course last semester. Use these data to find a least squares line for predicting the midterm score from the quiz score.

Assuming that the quiz and midterm are of equal difficulty this semester and the same linear relationship applies this year, what is the predicted grade on the midterm for a student who got a score of 17 on the quiz?

Quiz	20	15	13	18	18	20	14	16
Midterm	92	72	72	95	88	98	65	77

Critique: Data are from a real situation that should be of interest to students taking the course.

E. Example of a course syllabus

The syllabus below comes from a course where the instructor(s) have made a conscious effort to reflect the guidelines including cut some previously covered topics, go into more depth on key concepts, focus on collecting and producing data, integrate technology, alternative methods of assessment, real data sets and class activities.

STAT 1000 Fall 2004 Basic and Applied Statistics Syllabus

Intended Audience: Undergraduate students who have not studied statistics, and who have had a high school algebra course.

Course Goals/Objectives:

Students will learn the basics of descriptive and inferential statistics, so that they may develop statistical literacy and reasoning, and be able to carry out statistical investigations. By the end of this course students should be able to:

- Explain the “big picture” of statistical investigations.
- Understand the core statistical ideas
- Understand and critique articles and news stories that use statistics
- Experience and understand process of statistical investigations by “doing” statistics
- Understand why statistics cannot prove conclusions but can suggest them.
- Value what statistics can do for us, and not just think that statistics can lie.

Course Materials:

Textbook: *Statistics in Action*, Ann Watkins, Richard L. Scheaffer, and George W. Cobb, bundled with *Fathom Dynamic Statistics™* CD. Additional materials will be available for printing and bringing to class via WebCT.

Instructional format:

This is **NOT** a class where you come each day, listen, watch, and take notes! The primary method for learning new statistical concepts and methods will be by reading the textbook, working out problems from the textbook, and participating in class activities, discussions, and demonstrations. Many of these activities will include using *Fathom*, **state-of-the-art statistical software** designed to help students learn statistics and **eliminate much of the “math”** and number crunching, so they can focus on what statistics really mean and how we use them.

Small group and large group activities will be used to apply and deepen students’ understanding. Real data sets will be used during each class to help students develop statistical thinking and learn how to analyze and interpret data.

It is essential that students attend class each day and if they have to miss a class, should make every attempt to make up the work by obtaining notes from students and copies of the materials from WebCT. However it is almost impossible to learn as much from an activity that was carried out and discussed in class. Also, some form of assessment will be used in each class period: a pop quiz, minute paper, or short task.

Course requirements:

Attend class each day and participate in small group activities and large group discussions. Read about 20 pp each week in the text book. Write out solutions to assigned problems in the text and bring to class. Complete assessments as listed below.

Assessments:

Assessments can be expected every day, some scheduled and some unscheduled.

- 30%: written research project report. Students are expected to demonstrate their learning by completing a research project and turning in a written paper. Project milestones will provide pacing and feedback in completing a high quality project
- 30% Three in-class tests on the following topics (see attached policies concerning missing a test)
 - Test # 1: Design of experiments
 - Test # 2: Descriptive statistics and the normal distribution
 - Test # 3: Analyzing bivariate relationships
- 20% Final practical exam: students are provided a data set and software to answer a set of statistical questions.
- 10%: Critiques:
 - One graph critique
 - One critique of a statistical article
- 10% In class assessments which will include:
 - Homework related quizzes to encourage attendance and completion of homework problems
 - Minute papers to discuss course concepts and provide a vehicle for informal communication
 - A take home task that completes an in-class activity
 - A meaningful paragraph or similar writing task

Grading:

Percentage Cutoff	Grade	Percentage Cutoff	Grade	Percentage Cutoff	Grade
92.5%	A	80.5%	B-	59.5%	D
89.5%	A-	76.5%	C+	Below 59.5%	F
86.5%	B+	72.5%	C		
82.5%	B	69.5%	C-		

Stat 1000 Course Outline: Basic and Applied Statistics

Session	Date	Topic	Lesson Goals / Objectives	Read assigned sections / pages	HW to complete before class; Assignment due dates
1	9/7	Overview/Introduction to the course	Introduce ourselves, set the stage for the course		<ul style="list-style-type: none"> • Install <i>Fathom</i> on “home computer” • Complete “Walkthrough Guide,” pp 1-10
2	9/9	A Case Study in: Data Exploration Introduction to <i>Fathom</i> TM and inference	Introduce role of context, explore data, introduce the logic of inference, and use simulation for inference	Read Sections 1.1 and 1.2	<ul style="list-style-type: none"> • 1.1: D2., D7, P1, E1, E2 • 1.2: Activity 1.1, D12, D13
3	9/14	Why take samples and how not to	<ul style="list-style-type: none"> • Learn the basic vocabulary of sampling and surveys • Learn reasons for using samples • Recognize common instances of selection and response bias 	Read Section 4.1	<ul style="list-style-type: none"> • 4.1: E1, E7, E12
4	9/16	Randomizing: Playing it safe by taking chances	Learn and understand: <ul style="list-style-type: none"> • Why we rely on chance to pick a sample • Definition of a SRS • How to recognize and implement probability samples including: stratified, cluster, multistage, systematic 	Read Section 4.2:	<ul style="list-style-type: none"> • 4.2: P9, E14, E18, E19
5	9/21	Experiments and Inference about Cause	Learn: <ul style="list-style-type: none"> • Characteristics of well defined experiment • Difference between an experiment and observational study • Instances of confounding • Randomizing treatments protects against confounding • Build the underpinnings of inference 	Read Section 4.3	<ul style="list-style-type: none"> • 4.3: E21, E23, E25, E26 • Project Milestone: Introduction—Pose two research questions

Session	Date	Topic	Lesson Goals / Objectives	Read assigned sections / pages	HW to complete before class; Assignment due dates
6	9/23	Designing experiments to reduce variability	<ul style="list-style-type: none"> • Cause and effect requires randomized experiment • Distinguish variability within vs. between treatments • Understand why one reduces variability within treatments • Differentiate randomized, matched pairs, and randomized block designs • Learn advantages and disadvantages of each type of design and when it is appropriate to use them in practice 	Read Section 4.4	<ul style="list-style-type: none"> • 4.4: P28, E27, E30, E36, E48
7	9/28	Test #1: Design of Experiments			
8	9/30	Exploring distributions and graphical displays for distributions	Describe (univariate) data as a distribution, Recognize and interpret graphs (dot plot, stemplot, histogram, bar graph)	Read Section 2.1; Read 2.2 (exclude tennis ball activity)	<ul style="list-style-type: none"> • 2.1: P4, P5, E1, E2, E10 • 2.2: P8, D11, E16, E19
9	10/5	Measures of center	Understand and interpret mean, median, and mode and influence of outliers on these measures	Read 2.3 Part 1: pp 53-56	<ul style="list-style-type: none"> • 2.3: D18, D19a, P18, P19, P21 • Project Milestone: Methods—Data Sample & Organization
10	10/7	Measures of spread : range and IQR, boxplots	Understand the concept of variability and spread. Understand IQR and Range. Interpret boxplots. Comparing descriptions and graphs	Read 2.3 Part 2: pp 57-64	<ul style="list-style-type: none"> • 2.3: P22, P24, P25, P26.
11	10/12	Measures of spread : standard deviation	Understand and use Standard deviation as a measure of spread	Read 2.3: pp 64-72, 74	<ul style="list-style-type: none"> • 2.3: D26, D27, P28, P29
12	10/14	The normal distribution	Employ the normal (and standard normal) distribution as a model ; use standard deviations and z-scores to measure variation from the mean	Read Section 2.4	<ul style="list-style-type: none"> • 2.4: E46, E49, E53, E74

Session	Date	Topic	Lesson Goals / Objectives	Read assigned sections / pages	HW to complete before class; Assignment due dates
13	10/19	Reasoning about variability	Integrate measures of variability (i.e., spread): IQR, standard deviation, range	Reread all of chapter 2	<ul style="list-style-type: none"> Ch. 2: E23, E24, E31, E32 Matching graphs to statistics
14	10/21	Test #2: Descriptive statistics and the Normal distribution			<ul style="list-style-type: none"> Project Milestone: Descriptive Statistics
15	10/26	Scatterplots	<p>Understand bivariate relationships</p> <ul style="list-style-type: none"> Understand nature of bivariate data Describe shape (form), trend (direction), and variation (strength) in a scatterplot and interpret in context of the data Use <i>Fathom</i>TM to create a scatterplot Answer contextual questions using information from a scatterplot Know scatterplots are appropriate graphs to answer questions about the relationship between two quantitative variables Understand how lurking variables affect the relationship between two variables 	Read Section 3.1	<ul style="list-style-type: none"> 3.1: P2, E2, E4, E5, E7
16	10/28	Getting a line on a pattern	<ul style="list-style-type: none"> Use movable line to predict y given x 	Read Section 3.2	<ul style="list-style-type: none"> 3.2: P5, E11, E16, E22 Graph Critique due
17	11/2	Correlation: Strength of a Linear Trend	<ul style="list-style-type: none"> Estimate correlation from a scatterplot Understand correlation should not be computed from nonlinear data Understand a high correlation does not imply that the data are linear Be aware of lurking variables and correlation does not imply causation 	Read Section 3.3	<ul style="list-style-type: none"> 3.3: P10, E33, E37 Mid-term feedback
18	11/4	Bivariate Wrap Up			
19	11/9	Test #3: Analyzing Bivariate relationships			<ul style="list-style-type: none"> Project Milestone: Bivariate Analysis

Session	Date	Topic	Lesson Goals / Objectives	Read assigned sections / pages	HW to complete before class; Assignment due dates
20	11/11	Sampling from a population	<ul style="list-style-type: none"> Understand basic Ideas of Sampling; sampling proportions 	Read 5.1: pp 268-270	<ul style="list-style-type: none"> All use the random number table
21	11/16	Generating Sampling Distributions	<ul style="list-style-type: none"> Be able to Generate sampling distributions for a variety of statistics, observe the predictable pattern, contrast sample distributions with sampling distributions Use the simulation process model to explain sampling distributions 	Read Section 5.2	<ul style="list-style-type: none"> 5.2: E7, E9, E11, E15
22	11/18	Sampling distribution of sample mean	<ul style="list-style-type: none"> Use simulations to illustrate and explain the Central Limit Theorem, apply the CLT to different contexts 	Read Section 5.3	<ul style="list-style-type: none"> 5.3: E16, E20, E22, E24, E25
23	11/23	Probability using Simulation and experiments	<p>Using simulations and experiments for inference:</p> <ul style="list-style-type: none"> Use experiments to predict probabilities Conduct simulation of the experiment Sketch a simulation process model 	Read Section 6.1: pp 327-337	<ul style="list-style-type: none"> 6.1: P4, P5, E2, E4
24	11/25	<i>Thanksgiving Holiday</i>			
25	11/30	Toward a confidence interval (CI) for the mean	<ul style="list-style-type: none"> Review simulation model activity Find a CI from a sample from fixed populations Understand CI as reasonably likely sample means Interpret a CI for a mean, understand confidence level Understand relationship between capture rate and confidence level 	Read Section 9.1 (focus on concepts, not formulas)	<ul style="list-style-type: none"> 9.1: P1, E3, E5, E7

Session	Date	Topic	Lesson Goals / Objectives	Read assigned sections / pages	HW to complete before class; Assignment due dates
26	12/2	Toward a significance test for the mean (1-sample test)	<ul style="list-style-type: none"> Understand logic of significance test Identify and perform four steps in a significance test for a mean. Understand 1-sample test terms: Statistical significance, Null hypothesis, Alternative hypothesis Interpret a P-value Compare a confidence interval to a two-tailed hypothesis test. 	Read Section 9.2	<ul style="list-style-type: none"> <i>Against all Odds</i> #20: significance tests 9.2: E11, E12, E13, E14
27	12/7	When you estimate sigma: t distribution	<ul style="list-style-type: none"> Differentiate the true standard error (SE) vs. estimated SE Check conditions Significance tests for the mean Differentiating P-value tests versus fixed level tests 	Read Section 9.3	<ul style="list-style-type: none"> 9.3: E15, E17, E18
28	12/9	Inference for the difference of 2 means (2-sample test)	<ul style="list-style-type: none"> Understand CI and test of significance to compare two means Construct CI for mean difference Perform 4 steps in significance test for mean difference Deepen understanding of comparing means in terms of: CI, capture rate, statistical significance, P-value, and one-tailed test and two-tailed test 	Read Section 9.5	<ul style="list-style-type: none"> 9.5: E26, E29, E32 Article Critique due
29	12/14	Review and wrap up on inference			<ul style="list-style-type: none"> Milestone: Inferential Statistics
Final Project	12/17	Final Project is due no later than 12:30pm Friday, 12/17/04 Deliver in hardcopy to 325 Peik Hall or 330 Burton Hall.			<ul style="list-style-type: none"> Final paper with summary and conclusions.
Final Exam	12/22	Final Practical Exam, 325 Peik Hall: 1:30 – 3:30pm			

More on Grading

- **Tests:** Tests will consist of a variety of questions (multiple-choice, open-ended, etc.) designed to test your ability to apply the knowledge you gain by working on homework problems and participating in class activities and discussion. You may always use your calculator, Fathom, and your note-card during the free-response section of the tests. However, only a pen or pencil will be permissible during the multiple-choice section.

- **Making-up a test:** In general tests are not to be made up. Exceptions may be granted in cases of illness or emergency. If you cannot be in class on the day of the test, it is your responsibility to notify me **before** the test. If a make-up is granted it will be at the discretion of the instructor. If you fail to make-up the test at the scheduled time, you will not be able to make it up at all. If you cannot notify me before missing a test, you must provide documentation explaining your absence for the instructor to determine if an exception should be granted.

- **Homework assignments**

As a student of statistics, working through homework problems is an important piece in building a complete understanding of the concepts, as well as allowing you to practice doing statistics. Only by trying to apply the concepts can you be sure that you really understand them. Homework assignments should be regarded as a genuine “learning experience.” We urge you to form study groups to work on these problems and master the concepts. You should, however, be sure that the effort is truly collaborative. The best strategy for completing the assignment is to begin tackling the questions alone, then discussing with others, and finally writing up your answers by yourself. Feel free to consult the teaching assistant and instructors when you are stuck – but try not to ask for more help than you need to get started.

Homework will be assigned but not collected or graded. Many of the solutions to the homework problems are given in the back of the textbook. In addition, instructors and the teaching assistant will have a copy of the solutions manual available in their office.

Homework assignments can also be emailed to the teaching assistant for a brief perusal to make sure you are on the right track; however thorough explanations and help will not be provided via email. It will be your responsibility to work through the assigned problems and get help on those you do not understand. **Some of the exam and quiz questions will be very similar if not identical to homework problems.**