Example of a Well-Designed Course in: Statistics

Your Name: **Jill E. Thomley**

Name of Your Institution: **Appalachian State University**

1. Specific Context

- **The subject matter:** Statistics
- **The title of the course:** Statistical Data Analysis I
- **Typical class size:** 30-40
- **Level of the course:** 3000-level course with calculus prerequisite, required for computer science, environmental science, actuarial science, and some of our mathematics majors, an elective for some mathematics majors, statistics minors, and other majors.
- **Mode of delivery:**
  - Face-to-face with flipped design elements, all class meetings are held in a computer lab to facilitate a flexible mix of activities other than lecture.
- **Type of institution:**
  - University

2. General Description of the Course

Catalogue Description: STT 3850. Statistical Data Analysis I (4 credits). F and S. This course provides an overview of modern statistical data analysis. Programming with data, including simulations and bootstrapping, will be an integral part of the course. Techniques for parsing univariate and multivariate data sets will be examined. Coverage of probability, random variables, standard probability distributions and statistical sampling distributions will be sufficient to prepare the student for statistical inference. Inferential topics will include parameter estimation, hypothesis testing for proportions, means and medians, goodness of fit tests, and tests for independence. Standard and computationally intensive regression techniques will also be covered. Prerequisite: MAT 1110.

3. Big Purpose of the Course

The pioneers of statistic data analysis were driven by the need to answer questions and make sense of the increasingly complex and quantitative world around them. Almost a hundred years after R. A. Fisher published what is generally considered to be the first statistical methods textbook, Statistical Methods for Research Workers, we are living in a world where more data is produced every day than in the sum of past centuries: an estimated 2.5 quintillion bytes (that's 2.5 followed by 18 zeros). Virtually every scientific discipline now uses statistics to analyze data and determine the statistical significance of its research findings. Even the humanities fields employ statistics; for example, determining the most likely authorship of written works whose attribution is unknown. We all encounter statistics every day in the news and

popular media, created by others who wish to inform and persuade us. In the 21st century, "big data" drives decisions in businesses from local to global in scope. For example, in 2009, Netflix sponsored a contest that offered a million dollar prize to the individual or team that could statistically improve its ability to predict how much a customer would enjoy a movie based on their previous preferences. Methods used to handle these kinds of complex problems increasingly depend on computational power to create, verify, and implement statistical models. This requires researchers to move beyond many of R.A. Fisher's original analysis methods, which were mostly designed for small samples. Growing computational power and flexible programming languages allow us to implement strategies that exploit the power of both older and newer modes of statistical thinking. Students will learn foundation skills and ways of approaching problems that they can use in their careers and in their everyday lives as they interact with data/statistics.

4. Important Situational Factors/Special Pedagogical Challenge

Situational Factors/Special Challenges: Course Design Response:

1. Diversity in student knowledge about probability and statistical distributions prior to the course.	In prior semesters, students completed graded, online probability quizzes built by the instructor and delivered in our Moodle course management software. Quizzes were repeatable, but many students did them at the last minute or failed to make connections. Now these quizzes are a "gateway" to the graded assignments. Students must earn a minimum score of 80% to continue, but the scores do not count directly toward final grades. Quizzes have unlimited attempts, but there is a set completion deadline. Wuizzes include video tutorials from the J.B. Statistics web site, which facilitates a flipped classroom approach. In-class time is used to allow students to work on the gateway quizzes with their peers. The instructor offers individualized help and support to students to better address diversity in skills. An instructor-created Moodle glossary, which pairs mathematical definitions/ formulas with computer code needed to do probability calculations, is also linked to the quizzes using filters in Moodle. Previous experience shows that students in this course respond well to collaborative opportunities and tend to interact positively and teach one another, not just copy. (Note: these gateway activities are called quizzes because of the way they are programmed using the Moodle to make them self-grading, for instant feedback. They are called "repeatable quizzes" in the syllabus, versus later graded "proficiency problems" that will be used to determine course grades.) www.jbstatistics.com

2. Diversity in student knowledge about programming prior to the course, especially the R language, which is the required software for the course.	During the first two weeks of class, in-class time is be spent doing hands-on activities to develop programming skills such as loading datasets, manipulating vectors, writing loops, and exploring the structure of datasets. Students also learn about R Markdown documents, which will be used to submit graded assignments. Using templates, students begin to build their library of program examples. The commercial web site Data Camp offers an evolving selection of R tutorials (free to students) that introduce up-to-date R programming with statistics concepts. These can be linked directly to Moodle. www.datacamp.com

3. Negative student attitudes and expectations that make students reluctant to do the course work or appropriately pace themselves. Some of them have put off the course until the last semester they could take it and feel extra pressure not to fail, despite being insecure.	I will use standards-based grading in this course, which focuses on specific goals and assessing competencies or skills, with opportunities for iterative improvement toward skill mastery. A student must achieve some reasonable proficiency in all goal areas to pass the course. I will explicitly state the purpose of each assignment at the top of all my assignments. This will remind students of learning goals and the connection between work and learning. Students often express the idea that class work is mostly about assigning grades versus a process of teaching / learning. For example, one assignment says: PURPOSE: You will learn about... • the basic principles of sampling distributions, and • the sampling distribution of the sample mean (x-bar). In this course, lectures inevitably go over some people's heads while boring others. I will continue to seek out well-crafted statistics videos and use hands-on activities to flip the classroom away from too much lecture, or to break up lectures with reinforcing activities. In past semesters, I have brought in research examples from diverse fields and highlighted the real-world uses of statistics and R. I will continue to seek current examples and create new case studies using my own datasets (e.g., Fitbit data, Myers-Briggs personality data collected from the students) that better represents analyses done in the real world, versus textbook problems that often lack context. I will require students to complete the online human subjects research ethics training required of all researchers at

5. 3-Column Table

Learning Goals	Assessment Activities	Learning Activities
Foundational Knowledge: • Probability: understand and explain probability models and the Law of Large Numbers • Statistical distributions: understand and explain shape, E[X], Var[X], sampling distributions, and the Central Limit Theorem • Strength of statistical evidence: understand the logic of estimation and inference, explain the meaning of p-values in terms of probability, perform the appropriate procedures for a given situation and explain results in context	• Weekly exit tickets automatically scored completed or not by Moodle, and also read by the instructor • iClicker questions used as appropriate in class for formative feedback • Quizzes automatically scored by Moodle • Proficiency problems scored by instructor using a rubric that is available to students prior to the due date, repeatable once	• Textbook readings • Videos (jbstatistics.com and some other sources) • Traditional in-class lectures • Guided data analyses and discussion during class • Online "quizzes" on Moodle; these will be repeatable and students will have to achieve 80% before they may try proficiency problems • Out of class problem solving (proficiency problems); these will be repeatable one time, after initial grading/feedback
Application: • Probability: solve problems using probability models • Distributions: summarize the distribution of data, solve problems about sampling distributions using theory, simulate sampling distributions • Statistical evidence: construct hypotheses, perform appropriate resampling tests, state statistical conclusions, construct bootstrap CIs • Generalizability: interpret statistical conclusions in context (i.e., discuss practical conclusions)	• Weekly exit tickets automatically scored completed or not by Moodle, and also read by the instructor • iClicker questions used as appropriate in class for formative feedback • Quizzes automatically scored by Moodle • Proficiency problems scored by instructor using a rubric that is available to students prior to the due date, repeatable once	• Textbook readings • Videos (jbstatistics.com and other sources) • Traditional in-class lectures • Guided data analyses and discussion during class • Online "quizzes" on Moodle; these will be repeatable and students will have to achieve 80% before they may try proficiency problems • Out of class problem solving (proficiency problems); these will be repeatable one time, after initial grading/feedback

Integration: • Compare theoretical results with resampling results to connect the two modes of exploration • Analyze datasets in multiple ways, depending on the research question • Implement statistical procedures effectively and efficiently in the R language	• Weekly exit tickets automatically scored completed or not by Moodle, and also read by the instructor • iClicker questions used as appropriate in class for formative feedback • Quizzes automatically scored by Moodle • Proficiency problems scored by instructor using a rubric that is available to students prior to the due date, repeatable once	• Textbook readings • Videos (jbstatistics.com and some other sources) • Traditional in-class lectures • Guided data analyses and discussion during class • Online "quizzes" on Moodle; these will be repeatable and students will have to achieve 80% before they may try proficiency problems • Out of class problem solving (proficiency problems); these will be repeatable one time, after initial grading/feedback
Human Dimension: A. Learning about ONE-SELF: Analyze data contributed by self / peers to learn about self, students, instructor, or people beyond our school B. Interacting with OTHERS: Analyze data from many fields to understand the research questions and interests of others, as well as the ethics of data collection	• Weekly exit tickets automatically scored completed or not by Moodle, and also read by the instructor • iClicker questions used as appropriate in class for formative feedback • Online ethics training and plagiarism review (both are pass/fail)	• Textbook readings • Case studies on Moodle for some data sets used frequently in class • Guided collection and analysis of Myers Briggs personality data • Guided collection and analysis of Zener card psychic test data • Guided collection and analysis of anthropometric data • Guided analysis of other datasets from various fields • Discuss data and results in class • Complete online ethics and plagiarism training or review

Caring: Value the role probability and statistics play in the problems and applications of a given field of study or career path	• Responses to exit tickets questions in Moodle • iClicker questions • End of course survey	• Discuss uses and applications of statistics in class • Use real-world data in class • Post relevant links in Moodle
How to Continue Learning: • Use the components of standards-based grading to learn how to reflect on one's own progress n learning • Be able to identify and use free, open-source software like R for learning and development. Also be able to recognize and adapt to changes that such evolving software can make.	• Responses to exit tickets questions in Moodle • iClicker questions • Quiz scored by Moodle • End of course survey	• Discussion on the philosophy of standards-based grading and its implementation • Students use feedback to decide which assignments they want or need to redo to achieve their desired grade (when allowed) • Introduce students to CRAN (R package repository) and discuss methods for peer review/ vetting of packages • Video and quiz on history of the R programming language

Grading Procedures

Standards-based grading is a philosophy and pedagogical method that rethinks the relationship between work, learning, and course objectives. Rather than tracking a student's percentage score for a sequence of assignments and totaling them into a final percentage grade, students are assessed based on proficiency with regard to a set of course learning objectives (standards). Standards-based grading identifies student difficulties and allows multiple chances to meet or exceed each standard. Research on this pedagogy has shown increases in student learning and satisfaction. A shift to a student-centered, standards-based system in course work is potentially beneficial for both formative and summative assessment of student learning. It will still produce a final grade (summative assessment), but along the way this method has the potential to more clearly show students and instructors the specific course content areas in which students need to improve (formative assessment). Designing or redesigning assignments within the framework of a standards-based system that explicitly connects outcomes to course objectives seemed to me like an opportunity to provide more targeted feedback to students while also using my time and energy an instructor in a more effective way.

Assessment Methods:

Proficiency Problems ~ Your final grade will be based primarily on your work on three proficiency problems in five areas: (1) probability distributions, (2) normality and outliers, (3) hypothesis testing for means and proportions, (4) hypothesis testing for tables of counts, and (5) sampling distributions. The problems will be graded using a rubric with four levels (No Attempt, Statistics Novice, Statistics Apprentice, and Statistics Master), which will be specifically defined for each area. You will have two attempts at each problem, so you can improve your proficiency based on instructor feedback.

Example Problem and Rubric:

Psychologists at the University of Virginia studied a potential link between personality and geography through a series of observational studies and experiments. In results they published in 2015, the researchers stated that more people would pick the beach than the mountains for a vacation, but that introverts prefer the mountains more than extraverts. Some of the students represented in the PersonalitySTT3850 dataset were asked the question, "If you could choose one place to go on vacation, where would you go?" The possible choices were "the beach" and "the mountains". You will use our data to conduct a statistical hypothesis test to answer the question of whether introverts are more likely than extroverts to prefer the mountains.

A) Load the dataset and name it DATA. Use the `summary` function to explore its contents.

B) Subset the data so that it only includes cases with a valid answer to the question, "If you could choose one place to go on vacation, where would you go?" Name the subset TESTDATA. Then create a contingency table to show the proportion of people in each group (Extravert or Introvert) who chose mountains versus beach.

C) Write an appropriate null and alternative hypothesis for the research question.

D) Compute the test statistic from the sample data. Print out the result.

E) Conduct a permutation resampling test to compute a p-value for your test statistic. Print out the result.

F) If we set our acceptable chance of Type I error to 5%, do you reject Ho based on your test?

G) Given on your statistical conclusion in (G), what would be the answer to the research question?

H) Give two reasons why this dataset might not be a representative sample for testing this research question.

Proficiency Goal: Perform a permutation resampling hypothesis test for means and proportions using R and generalize the results in the problem context.		
Stat Novice (1)	Stat Apprentice (2)	Stat Master! (3)
• Two or more components in the "Stat Master" column are incorrect, but student made a reasonable attempt at solving the problem • R code functions incorrectly or not at all. • Note: if the student cannot code the problem in R, they will likely score at this level	• One of components 1-5 in the "Stat Master" column is incorrect, but the overall logic of the test is correct. For example, the wrong test statistic is calculated, but sufficient simulations are generated and p-value for the incorrect test stat is correctly produced and the logical conclusion is drawn • The test is implemented correctly, but discussion (item 6) is missing or not correct/sufficient. • R code functions correctly but has extra or redundant code (e.g., student is using template with no mods to remove lines of code used for teaching, change names of variables, etc.)	1. The null and alternative hypotheses are correctly stated 2. The test statistic is correctly calculated from the data 3. Sufficient resamples have been generated for validity 4. The p-value is correctly calculated from the results of the simulation 5. The appropriate statistical decision is made, given α 6. The practical significance of the problem is discussed in the context of the original research problem/ question 7. R code is correct, efficient, and lacks extraneous code; matches the problem being solved.

Repeatable Quizzes ~ You will complete several online quizzes in Moodle. Your attempts are unlimited, so you can go back and try the problems you missed again. These learning activities are intended to encourage practice and mastery of certain skills, to serve as a gateway to the Proficiency Problems. Repeatable quizzes will be open book/notes/internet and students will be allowed to collaborate, within certain guidelines. You must pass certain quizzes with an 80% or greater to submit related Proficiency Problems for grading. Scores on repeatable quizzes do not count directly toward your final grade.

Exit Tickets ~ Each week you will complete a short reflection and work plan on Moodle, addressing: brief feedback course material you understood and what you are confused about, as well as what learning activities you intend to engage in between Thursday's class and the subsequent Tuesday class meeting. These will be used to help determine +/- on final course grades (e.g., B- or B+ versus B).

Ethics ~ Ethics are an important part of responsible research. Much of the data we work with was collected from human subjects. You will complete training required of any anyone at ASU who engages in human subjects research. You will also review anti-plagiarism guidelines. You must receive "Credit" for these two assignments to pass the course, regardless of your other grades.

The minimal requirements for an A in this course are:

- Achieve "Statistics Master" level on two problems in each Proficiency Group
- Achieve "Statistics Apprentice" level on one problem in each Proficiency Group
- Complete Plagiarism Goblins and Human Subjects Research Ethics
- Complete 9 of 15 Exit Tickets (< 9 = A-)

The minimal requirements for a B in this course are:

- Achieve "Statistics Master" level on one problem in each Proficiency Group
- Achieve "Statistics Apprentice" level on one problem in each Proficiency Group
- Achieve "Statistics Novice" level on one problem in each Proficiency Group
- Complete Plagiarism Goblins and Human Subjects Research Ethics
- Complete 9 of 15 Exit Tickets (> 12 = B+, < 9 = B-)

The minimal requirements for a C in this course are:

- Achieve "Statistics Apprentice" level on two problems in each Proficiency Group
- Achieve "Statistics Novice" level on one problem in each Proficiency Group
- Complete Plagiarism Goblins and Human Subjects Research Ethics
- Complete 9 of 15 Exit Tickets (> 12 = C+, < 9 = C-)

The minimal requirements for a D in this course are:

- Achieve "Statistics Apprentice" level on one problem in each proficiency group
- Achieve "Statistics Novice" level on two problems in each proficiency group
- Complete Plagiarism Goblins and Human Subjects Research Ethics
- Complete 9 of 15 Exit Tickets (> 12 = D+, < 9 = D-)

6. Weekly Schedule

This is a four-hour class that is sometimes held three days per week (two 75 minute class periods and one 50 minute class period) and sometimes held four days per week (four 50 minute class periods). Therefore the schedule is given in terms of overall weekly topics. With regard to quizzes, it was suggested that students begin the quizzes in the weeks indicated, but they were not due until just before the relevant proficiency problems. Some weeks do not have the full 200 minutes of class time due to start/stop times of the semester and fall break, such as Week 15.

Week:	
	Content: course introduction, MBTI data collection activity, vocabulary of data and datasets, R script and R Markdown basics, reading data into R, creating vectors and matrices in R, textbook Chapter 1 and supplemental readings and/or videos such as Kelly Black's R Tutorial (http://www.cyclismo.org/tutorial/R/index.html) Assessments: Plagiarism Goblins, History of S and R quiz, Introduction to the R Language quiz, iClicker questions, Exit Ticket #01
	Content: Zener Card data collection activity, permutations and combinations, probability rules and distributions, random variables, expected value/variance, basic operations in R, basic probability calculations in R, supplemental readings and/or videos Assessments: Discrete and Continuous Random Variables quiz, Expected Value and Variance quiz, Probability/Combinations/ Permutations quiz, iClicker questions, Exit Ticket #02
	Content: Bernoulli trials, Binomial/Hypergeometric/Geometric/Negative Binomial distributions, "Are You a Psychic?" activity, R functions for discrete probability distributions, graphing discrete pmf and cdf functions, supplemental readings and/or videos Assessments: Binomial quiz, Geometric/Negative Binomial quiz, Hypergeometric quiz, iClicker questions, Exit Ticket #03
	Content: in-class work time for quizzes, reviewing R functions and R Markdown documents, and Discrete Distribution proficiency problems (due before classes the following week) Assessments: Discrete Distribution proficiency problems, Exit Ticket #14
	Content: Uniform, Exponential, and Normal distributions, R functions for continuous probability distributions, supplemental readings and/or videos Assessments: Uniform quiz, Exponential quiz, Normal quiz, iClicker questions, Exit Ticket #15
	Content: exploratory data analysis for categorical and quantitative data (summary statistics and graphs), creating summaries and graphs in R, textbook Chapter 2 and supplemental readings and/or videos, Assessments: resubmission of Discrete Distribution proficiency problems, iClicker questions, Exit Ticket #16
	Content: boxplots and determining outliers, normal probability and empirical cumulative distribution plots, skewness, kurtosis, creating summaries and graphs in R, textbook Chapter 2 and supplemental readings and/or videos Assessments: Descriptive Statistics quiz, iClicker questions, Exit Ticket #17
	Content: in-class work time for quizzes and Normality and Outliers proficiency problems (due before class the following week) Assessments: Normality and Outliers proficiency problems, iClicker questions, Exit Ticket #18

	Content: introduction to concepts of hypothesis testing, permutation test for the difference between two means, permutation test for the difference between two proportions, conducting permutation tests in R, comparing permutation tests for two means with traditional two-sample t-tests, assumptions of permutation versus traditional tests, textbook Chapter 3 and supplemental readings and/or videos Assessments: Introduction to Hypothesis Testing quiz, iClicker questions, Exit Ticket #09
	Content: test of independence and test of homogeneity using permutation resampling and traditional chi-square tests, textbook Chapter 3 and supplemental readings and/or videos Assessments: Hypothesis Testing Errors and Significance quiz, resubmission of Normality and Outliers problems, iClicker questions, Exit Ticket #10
	Content: test of goodness of fit for categorical data and the normal distribution using simulation and traditional chi-square tests, Benford distribution, textbook Chapter 3 and supplemental readings and/or videos Assessments: Human Subjects Research Ethics, iClicker questions, Exit Ticket #11
	Content: in-class work time for quizzes and Hypothesis Testing proficiency problems (due before class the following week) Assessments: Hypothesis Testing proficiency problems , iClicker questions, Exit Ticket #12
	Content: sampling distributions, finding expected value and variance for sums and linear transformations of random variables analytically and via simulation, sampling distribution of the mean, Central Limit Theorem, textbook Chapter 4 and supplemental readings and/or videos Assessments: Introduction to Sampling Distributions quiz, Sampling Distributions, Part 1 quiz, iClicker questions, Exit Ticket #13
	Content: Normal approximation to the Binomial distribution, continuity correction, central limit theorem for proportions, solving problem analytically and via simulation, textbook Chapter 4 and supplemental readings and/or videos Assessments: Sampling Distributions, Part 2 quiz, Sampling Distributions, Part 3 quiz, resubmission of Hypothesis Testing problems, iClicker questions, Exit Ticket #14
	Content: in-class work time for quizzes and Sampling Distribution proficiency problems (due the last day of class) Assessments: Sampling Distribution proficiency problems, Exit Ticket #15
	FINAL EXAM: Work on resubmission of Sampling Distribution proficiency problem or other optional resubmissions (two problem from any previous set of proficiency problems)

- Teaching Strategy

  STANDARDS-BASED GRADING: As described above. This can be time intensive, but for me it feels like a better use of my personal energy resources than the traditional homework/quiz/test model.

  TRANSPARENCY: I can rely on the book for topic sequencing, but no book can give students a reason to be excited about class. I am open with students about my own struggles. We discuss the reason for certain types of assignments and pedagogical reasons for why do what I do in class. This was particularly important when implementing standards based grading, which was a new model for most students.

SHARED EXPERIENCES: I hated my first statistics class. Later as a psychology graduate student, I had to learn to program on a mainframe, which I found tedious. While working on research at a VA hospital, I had an epiphany about statistics while collaborating with a professional statistician. It changed my life and career. Now I LOVE statistics, so I know that a major reversal in attitude is possible. I believe everyone needs to know some probability and statistics to navigate our data-filled world. Even social media like Facebook and Twitter is filled with research and statistics that can be used to connect to students.

TALK THE TALK: Vocabulary is critical to properly using statistics and communicating results. Every discipline has a language where words may change meaning. Learning to "speak like a statistician" is as important as being able to do specific tests. I speak to students in the vocabulary from the start, tell them why it's important, and do not let students get away with misusing terms, even early on.

POOL KNOWLEDGE: I believe every student in my course is capable, even if they do not. I also assert that the only real way to learn statistics is to DO statistics. For me, collaborative work has been effective for pooling experiences and peer-teaching knowledge and skills—I call it "crowd sourcing". Even weak students have some piece of knowledge they can contribute; talking about the material is as important for long-term learning as reading, writing, watching, and doing. However, it also requires trusting students to not try to get a free ride, which can be difficult.

REAL DATA: I use at least some real-life data and include students in data collection (e.g., determining their Myers-Briggs personality types and taking anthropometric measurements). Feeling connected to data helps students connect to analyses and discussions of concepts, methods, and results, including reliability and validity. Each dataset is solving a mystery and telling a story. I want students to be invested in uncovering the connections between research questions and answers. This is one difference between statistics and many mathematics classes: the answer is only important in the context of the data and the process that produced it.

7. Evidence of Impact (optional)

The final grade distribution is shown below, as compared to the same course in Fall 2015 and Fall 2014. In 2015, repeatable quizzes counted directly toward students' final grades and students took traditionally-scored exams. In 2014, I used traditional homework and exams. I have taught the course in other terms, but spring semester (and definitely summer) courses may differ from fall in ways other than content.

	A, A-	B+, B, B-	C+, C, C-	D+, D, D-	F	Drops
2016 (n = 80)	51	10	4	0	6	9
2015 (n = 74)	32	18	12	6	4	2
2014 (n = 72)	25	21	15	4	4	3

About 72% of Fall 2016 students course an A, versus 44% and 36% in previous fall semesters. Some instructors might consider this proportion of A's unacceptably high and attribute the results solely to the fact that students got multiple attempts. I also had more experience teaching this course. However, if the premise of grades is that they reflect learning, then arguably greater learning occurred among the 2016 group than the previous two, at least among C or better students. While the D/F/Drop rate in 2016 was similar to previous semesters (19% vs. 16% in 2014 and 15% in 2015), there were more F's and drops, with no D's. It seems there was a shift toward the extremes versus the middle of the scale, though without more knowledge about why students dropped it is not clear whether this was caused by the course redesign or external factors. Further investigation is needed; I discuss this more in Section 8.

Fall 2016 students completed a final reflection at the end of the course. Some results are presented below (n = 67 for most questions). Note that these data only represent students who completed the course, not drops or similar cases. Thus, they are undoubtedly a bit biased toward those who were successful.

Questions about the Repeatable Quizzes	1= definitely no	2 = mostly no	3 = neutral	4 = mostly yes	5 = definitely yes	Average
The quizzes helped me prepare for the proficiency problems.	-----	3%	7%	52%	37%	4.24
Students should get extra credit for good quiz scores.	12%	22%	22%	28%	15%	3.12
Quiz scores should count toward students' final grades.	30%	27%	30%	12%	1%	2.28
Quizzes should have a lower "pass" threshold than 80%.	37%	40%	16%	6%	-----	1.91
Quizzes should be optional (for people who want practice).	47%	35%	8%	11%	-----	1.82

Quizzes should have a higher "pass" threshold than 80%.	45%	37%	15%	3%	-----	1.76

Students seem to agree that gateway quizzes were useful in preparing for graded proficiency problems, that 80% was a reasonable threshold score, and that quizzes should not be optional. Some students would have liked the quizzes to count toward their final grade, possibly as extra credit, but they generally seemed to embrace the concept of using them as practice and formative feedback.

Questions about the Proficiency Problems	1= definitely no	2 = mostly no	3 = neutral	4 = mostly yes	5 = definitely yes	Average
Being able to revise and resubmit the problems improved my learning.	-----	-----	1%	24%	75%	4.73
The feedback I received was sufficient to make corrections.	-----	-----	4%	51%	45%	4.40
The feedback I received for problems was clear.	3%	3%	7%	37%	49%	4.27
I could see the connections between the different sets of proficiency problems.	-----	1%	10%	54%	34%	4.21
I often felt unsure about what I was being asked to do.	3%	46%	26%	25%	-----	2.73
I would work harder at the problems the first time if I knew I only got one chance at them.	16%	36%	16%	25%	6%	2.69
Being able to revise and resubmit the problems just rewards slackers.	48%	39%	7%	4%	1%	1.73
I would rather have had more typical in-class exams in this course.	67%	22%	7%	3%	1%	1.52

Overall, students had a strong preference for the proficiency problems versus typical tests and seemed to feel they learned more by being able to revise and resubmit, though of course they may be influenced by the higher grades versus true learning. They also tended to say that the feedback they received about their initial attempts was clear and sufficient. Some felt unsure about what they were being asked to do, and there were mixed feelings about whether they worked as hard as they could on the first submissions, knowing they had a chance to resubmit.

To gain more insight into their thought processes, I asked the students, "If someone you knew was going to take this class in the future, give three concrete and specific recommendations you would make to help them succeed, based on your experience

in the course." Most of the answers revolved around themes such as doing quizzes early and taking advantage of the chance to work together. For example:

- 1. DO THE GATEWAY QUIZZES EARLY 2. Take the time to get comfortable and practice with R, especially if you're new to programming languages. Subsetting is specifically critical to succeeding. 3. Take paper notes on the in-class lectures. The notes written on the board are important and not easily copied directly into an RMD or other file quickly in the time-frame of the discussion.

- Absolutely use the glossary provided from the beginning of class to the end. This helps you pass the quizzes and to work out proficiency problems. Get ahead on quizzes, do not fall behind. The quizzes are easy to forget. Always watch the videos before the quizzes, it can save time in the long run. Find a group of people to work with. There's a lot of information to retain and know. Hearing two, or even three people explain things all differently, helps you learn how to perform the problems.

- 1. It is important to go to class. Tempting as it may be to just work on the quizzes and proficiency problems from home, lectures cover details and examples that you won't get from reading the textbook or looking at [Moodle]. It's also important to attend class on workdays so that you can request help from the professor. 2. Put full effort into the proficiency problems on the first try. It may be tempting to blow it off since there is a resubmission, but it's important to actually do the work the first time. 3. Put full effort into the quizzes. They are the best way to prepare for the proficiency problems, so it's important to completely understand how to solve all the problems on the quizzes before moving on to the problems.

- First of all, I'd tell them to take the class with you, since you're the only teacher that teaches 3850 with standards based grading. Here are the three pieces of advice I'd give them to succeed: 1. The class is what you make it. We're all here to learn, so don't take the class lightly and go the easy route through it; take advantage of grading system and use it to really learn the material. It's quite rare that we're able to learn iteratively. 2. Find a friend, or a group, that also takes the class seriously, and collaborate together. The benefits of working with like-minded individuals far outweigh the cons. You're able to bounce ideas off one another, and you'll never learn something nearly as well as you would until you start teaching it to others. 3. Use the resources you're given. There are numerous R-scripts, presentations, glossary entries, etc. It's easy to forget, or not even know about them, but they are immensely helpful, and are often directly related to the material covered in the class.

Most students had never been exposed to standards-based grading, so I asked, "Was the standards-based grading system used in this course confusing to you? If yes, how do you think it could it be explained more clearly to future students? In either case, what do you think is one pro and one con of this system as compared to more commonly-used points based systems you may have experienced in other courses?" Responses were mostly positive, with a few mixed opinions about the high bar set for an A and the general lack of the kind of partial credit. For example:

- No, the system seemed fairly straightforward, it just seemed a little unbalanced, considering a tough start to the semester can make making a decent grade difficult if not impossible. One pro is that the outline for what you need to do is clear for each grade, but a con to that is that there is no lenience and you must meet the hard criteria.

- I don't think it was confusing, I only think it was mildly ridiculous at times. I would miss one thing, and to fix it all I'd have to do is change one line of code and I would go from a 2 to 3. Maybe make it a 5 pt scale instead of a 3 pt scale. Would be much more forgiving and make me feel less like I made a 66% on my assignment.

- Pro: I think it's more easily understood and requires less mathematical finagling than a traditional points- or percentage-based grading system. Con: I feel like the hard cap on grades (i.e. "You cannot get an A if you don't get two 3's on every problem set") is very stress-inducing and makes it easy to feel "doomed" in the course early on. Since I ended up with two 2's and a 3 on my first set of proficiencies, I spent most of the rest of the semester worried about "blowing" my two extra resubmissions, which I did when I had a period of intense illness during the third set of proficiency problems and failed to earn the proper credits on my resubmissions. I knew I was "locked out" of a better grade based on the rubric, and that directly impacted my motivation to succeed in the course. In a points/percentage system in this situation, I would say, "I need to do better"; with this system, I found myself saying, "I had better not do any worse."

- I think the total amount of points gained from the problems should translate to a grade, because if you were to get all 3's on one problem, then a 3 and two 2's on another, does that translate to an A after averaging? I did like how I was not so focused on my grade because I knew that the only grades were the proficiency problems, so I had all of my grades readily in front of me.

- The system was not confusing. If students are confused, perhaps a table or graphic could help them visualize it better. I loved this system. It allowed me to learn without stress or test anxiety. I feel I learned much more in this course than I would have with more traditional exams. The ability to redo problems did not make me lazy. It enabled me to actually learn from my mistakes, which most traditional classes to not allow. A con would be that you have to do very well on every topic in order to get a good grade. In a traditional course, if you do poorly on one exam, it would be averaged in with the others and hopefully balance. I wish more professors would use similar grading systems.

(Note: This last answer was slightly edited for length; I eliminated some specific examples or points that may not be as useful to a broader audience.)

I do not think the standards based grading used in this course was confusing, although I was initially curious as to what the proficiency problems entailed, since our grade relied so heavily on them. I quite enjoyed this style of teaching / learning, and I wish that other classes followed suit. I'd prefer to talk about more than one pro or con, so here goes:

- PROS 1. There is a sense of accomplishment when completing the required quizzes that unlocked a set of proficiency problems. 2. The gating system for each proficiency problem clearly defined what knowledge we were expected to demonstrate, and forced us to practice it until we knew the material well. This is a huge pro, since compared to other grading systems where quizzes can potentially hurt you, they are only beneficial here. 3. The encouraged collaboration with open resources models the real world, and gives us valuable experience collaborating with others. It removes the need to cram (which leads to a higher retention rate of knowledge). There's nothing like debating answers to problems, since to prove your rhetoric, you need to have evidence to back it up. This lends itself to obtaining a deeper understanding of the material. 4. We're treated as mature adults, and our learning becomes our responsibility. This system is highly dependent on the individual, which is very refreshing, and is how things are in the real world. 5. Lastly, iterative based learning is just how people learn. Trial and error has been engrained in basic behavior forever. It removes the stress behind quizzes and exams due to having multiple submissions, but on the other hand, we still have to know the material to do well. I don't think I've learned material in college as thoroughly as I have in this class, and that says a lot, since I'm very, very interested in some major specific courses.
- CONS 1. The system can be taken advantage of to get good grades with little effort. Although this is risky, and would be done by people with poor character, but a person could supposedly not submit, or do poor work on their first time around, then copy answers from those that got perfect scores. The feedback that's provided explains exactly where that specific individual made their mistake(s), which is invaluable, but can also be abused by sharing that information with people that put no effort into the problems… 2. We talked about this in class and a possible solution, but a system like this imposes quite a bit of grading on your part. I think having some way to automate grading, through GitHub or some other means (it's definitely possible!), would speed

things up quite a bit for you. I felt bad that you had to spend so much of your time grading, writing feedback, etc. This is something that I'm interested in, from a programmer's perspective. 3. This was probably my fault, but the last set of quizzes and prof problems seemed way more difficult to me than they should have been. I didn't catch everything during the lectures, and so I spent a ridiculously long amount of time trying different approaches to solving problems on the quizzes. (This was the good part of iterative based learning, as I was eventually able to find the solutions to said problems on my own)...

In closing, I'd just like to say that I was hugely impressed with the amount of effort you put into teaching this course. The glossary, rescripts, presentations, etc., were very valuable, and I appreciate the time you put into compiling those. It was the first standards based grading I've ever participated in, and it was a real pleasure. I don't think I've ever gained such an understanding of a class up until this point.

The last student notes a significant challenge of this type of course redesign: there is a great deal of grading, though not much more than a traditional course that uses multiple homework assignments and tests. With fewer students in my sections, my feedback might have been more effective. There is a significant time sink in setting up a course like this, including creating quizzes and specific goal-tailored problems.

Finally, I asked the students whether they changed the way they worked on quizzes and problems as the course progressed, since the style was new to most of them. In general, students reported that they learned to pace themselves better through the semester, with less procrastination and more peer collaboration. They also generally made better use of the many resources provided. In their own words…

- At the beginning i tried to just get through the quizzes and reach the 80% mark but as the class went on I tried to actually learn the material to perform better on proficiency problems.

- Towards the end of the course as the material got more difficult, I treated the quizzes as more of a trial and error process than as an actual evaluation of my ability. I felt like this helped me understand the material better because I wasn't afraid to try different things.

- In the beginning, I worked on the quizzes by myself, which took me longer to get them done, which put me behind. Halfway through the semester, I began working with a group of classmates which helped me understand the material faster.

- In the beginning, I tried to keep up with the quizzes as they were due on the calendar, but I ended up succumbing to the "no due date" and "work at your own pace" mantra that's present in the class. I realized pretty quickly after the first series of proficiency problems that waiting to the last minute to complete quizzes is a huge mistake, as it adds a large amount of stress and work if you're trying to get all 3's the first time around. As such, I changed the way I worked on the quizzes and focused on pacing myself in a way that gave me enough time to work on proficiency problems. In addition, I also began saving all the work for each quiz in an R file, so that if I needed to reference how I completed a specific problem, I could go back and find my answer.

- At the end of the course I would start the proficiency problems much earlier in order to fully benefit from the in class work days. I started to realize further into the class how important work days were to succeed.

- I also started working on the proficiency problems in groups. I stopped being so worried about the first submission since I knew I would be corrected and had the opportunity to resubmit. I tried to work on them more ahead of time as well to have a better chance of looking back later and getting them correct.

- I felt like as I learned more about R and what was being asked became clearer I was able to get into them faster. I approached the problems with a more organized mindset later in the semester, especially after I got comfortable using the glossary and in-class scripts. The feedback given would always alter my approach as well for the resubs. Sometimes it would show me maybe my logic was wrong and that I needed to think about the problem differently. I would carry that over into the next set of problems and wonder how to be more efficient.

8. Most Exciting Aspect of the Re-Designed Course for Me

I found many aspects of this process exciting. It began when a colleague and I were accepted for a course redesign workshop at our university. Although we intended to focus on different courses, we had common goals about the use of standards-based grading and this spurred deep and interesting conversations about how to put these into practices. Interacting with peers from many disciplines during the workshop and receiving feedback from the course leaders, I gained new insights into a course I'd been teaching for a while. I was able to refine and use in new ways some aspects of the course I already liked and had invested a great deal of time in developing, and which students had told me were effective in the past (e.g., Moodle quizzes). I was also able to make concrete some of the nebulous ideas I had considered in the past about how to make my course better but never had time to put into practice. This mode of thinking has percolated into my other courses.

Approaching the primary graded course assessment as sets of proficiency problems graded with rubrics, rather than traditional exams, fits very well with my preference (and professional guidelines) for case studies using real data and a holistic approach to statistical problem solving. As I continue to build a library of proficiency problems for this course, I'll be able to have students attempt new (albeit similar) problems to demonstrate proficiency rather than correct their original work, which is more in keeping with the way standards-based grading is typically implemented.

I was perhaps most pleased by the enthusiastic student response to this redesign—at least among those who remained in the course. Given student diversity and the fact that the course is programming-intensive, I believe encouraging active learning with some degree of self-pacing really is the best solution. I do have some concerns about the students who dropped, because their decision may or may not have been related to the course design. I'm working on ways to reach struggling students and encourage engagement sooner—for example, as some of my students suggested, I plan to have less open-ended deadlines on gateway quizzes to help them better pace themselves. I also plan to offer my pacing guide in HTML format versus on the Moodle calendar, since it is easier to see the whole course at once and more friendly on mobile devices, which so many students use.

Overall, the energy in the classroom while students worked on the gateway quizzes and proficiency problems was high and they were interacting with one another in positive ways—not copying answers from one another, but rather comparing code and solutions to mutually teach, learn, and solve problems. My subjective perception of the experience, backed up by their own comments, is that students didn't treat the proficiency problems as a "brain dump" the way I've seen at times with exams, but rather as another chance for engagement, learning, and improvement.

9. My Contact Information

My name and institution: Jill E. Thomley, Appalachian State University
My email address: thomleyje@appstate.edu