STAT 201 and 301 have been redesigned to fit into the Statistics Department’s Pathways to Understanding and Mastery of Statistics (PUMAS) framework. This framework allows students of any background to begin taking Statistics courses at a level appropriate to them, and continue on to any level they desire.

Many students will decide between entering with the statistics curriculum by starting with STAT 201 or with STAT 301a,b,c. The STAT 201 and STAT301a,b,c curriculum share a core of material that comprises about 2/3rds of the STAT 301 curriculum. The topics include:

- The role of data and statistics in scientific research
- Descriptive statistics
- Elementary probability and the normal distribution
- Sampling distributions and the Central Limit Theorem
- Confidence Intervals
- Hypothesis Testing
- Correlation and simple linear regression

STAT201 covers these topics at a slower pace than STAT301a,b,c,… because it integrates substantial review of mathematical skills with the material and it is taught with weekly small recitation sections.

Our STAT 201 and 301 redesigns reflect the advice and principles laid out in the American Statistical Association’s 2016 Guidelines for Assessment and Instruction in Statistics (GAISE) recommendations. The GAISE report can be found here:


Here are the recommendations:

1. Teach statistical thinking.
   - Teach statistics as an investigative process of problem-solving and decision making.
   - Give students experience with multivariable thinking.
2. Focus on conceptual understanding.
3. Integrate real data with a context and purpose.
4. Foster active learning.
5. Use technology to explore concepts and analyze data.
6. Use assessments to improve and evaluate student learning.

GAISE also gives suggestions for topics that may be omitted from an intro stat course (p. 23). These are:

1. Probability theory
2. Constructing plots by hand
3. Basic statistics
4. Drills with z-, t-, chi-square, and F-tables
5. Advanced training on a statistical software program

We present a summary of how the recommendations are addressed.

1. **Recommendation 1: Teach statistical thinking**
   Here, GAISE recommends teaching statistics as an investigative process, as opposed using a “cookbook” type approach. In the past, STAT 201 / 301 taught “step by step” procedures, using small toy examples disconnected from one another. In the re-design, we have moved away
from the step-by-step approach and to a big picture approach. For example, we use a handful of real data sets from real studies and continually come back to them as the class progresses, showing how different methods can be used to explore different aspects of the same data. We have also structured the course so that each new topic builds upon the previous one, so that students see the interconnectedness of topics rather than seeing them as compartmentalized. GAISE also recommends teaching multivariate thinking, which we do early on at a conceptual level with examples in which adding a new variable affects the perceived relationship between two others (a la Simpson’s Paradox). We then revisit this in regression analysis, where students analyze large datasets containing many variables.

2. **Recommendation 2: Focus on conceptual understanding**

   This is the largest change we have made in our course redevelopment. The course notes emphasize the purpose and logic behind statistical methods. Course assignments ask students to answer questions about why methods are used. We have de-emphasized formulas and calculations, and for the formulas we do teach, we dedicate time to teaching what each part of the formula is doing and why. We have eliminated all “mysterious” formulas, in which students were asked to put numbers into formulas without knowing how they worked. The GAISE recommendations suggest that teachers:
   
   * View the primary goal as to discover and apply concepts.
   * Focus on students’ understanding of key concepts, illustrated by a few techniques, rather than covering a multitude of techniques with minimal focus on underlying ideas.
   * Pare down content of an introductory course to focus on core concepts in more depth.
   * Perform most computations using technology to allow greater emphasis on understanding concepts and interpreting results.
   * Although the language of mathematics provides compact expression of key ideas, use formulas that enhance the understanding of concepts, and avoid computations that are divorced from understanding.

   All of these suggestions were followed in our 201 / 301 redevelopment. For example, we cover fewer mathematical varieties of hypothesis testing, in favor of spending more time on the conceptual framework of hypothesis testing. We teach basic inferential methods for means but not for proportions, which cuts in half the number of formulas and minor details that students need to worry about (e.g. “is this a z-test or a t-test?”). We have removed chi-square testing, which involves lots of mysterious calculations. We have de-emphasized the computational components of ANOVA and regression. We have eliminated statistical tables, so that students are no longer looking up critical values for making 95% CIs. And most computations are performed using software rather than pencil and paper.

3. **Recommendation 3: Integrate real data with a context and purpose**

   The older versions of STAT 201 and 301 used many small, often fictitious data sets. This was necessary when students were being made to calculate statistics by hand. The new versions of the courses use real data sets, including one from an influenza vaccine study performed here at CSU. Large portions of assignments involve interpreting statistical results in the context of the scientific questions that motivated collecting the data. For example, students use ANOVA to analyze data from a study on morphine tolerance in rats, and once they have the results they must answer detailed questions that relate the results back to the motivating question of “what is the nature of morphine tolerance?” We go far beyond the traditional “interpretations” of declaring statistical results “significant” or “not significant” and then being done. We avoid letting students recite boilerplate phrases and challenge them to think deeply about how the results of a data analysis should and should not be used to inform scientific questions.
Three of the specific GAISE suggestions that we followed are:

- Use class-generated data to formulate statistical questions and plan uses for the data before developing the questionnaire and collecting the data. For example, ask questions likely to produce different shaped histograms, or use interesting categorical variables to investigate relationships. It is important that data gathered from students in class does not contain information that could be embarrassing to students and that students’ privacy is maintained.

- Use subsets of variables in different parts of the course but integrate the same data sets throughout. (Example: Use side-by-side boxplots to compare two groups, then use twosample t-tests on the same data. Use histograms to investigate shape, and then later in the course to verify conditions for hypothesis tests. Encourage students to explore how multiple variables in the data set relate to one another.)

- Minimize the use of hypothetical data sets to illustrate a particular point or to assess a specific concept.

We demonstrate regression using data collected from students, in which they guess the age of people in photographs. We use the same data sets for teaching statistical summaries, graphics, confidence intervals, hypothesis tests, and regression. And we avoid using fake “toy” data sets.

4. **Recommendation 4: Foster active learning**

The STAT 201 and 301 curricula now include a number of active learning elements. There are frequent iClicker questions included throughout lecture. Students are encouraged to take a minute or two to discuss these questions with those around them. Some of these questions are meant to give students feedback on their own understanding (e.g. review type questions), but many are meant to get students thinking about the concepts to be explored. We give many iClicker questions in which there is no “correct” answer, and arguments can be made for many options. We also use iClicker questions to have students guess what the result of a statistical analysis will be before it is performed. This encourages them to think through the details of how a method works without seeing the answer first.

We have developed full class activities that take up an entire class session. In one of these, students roll dice and calculate means and then share their values to construct histograms that demonstrate the Central Limit Theorem. In another, students get into small groups and take “samples” from a bag of numbers, then use these samples to construct 95% confidence intervals. The groups share their intervals and we create a large forest plot showing all the intervals and seeing how they vary. Students are then asked to guess the population mean of the bag of numbers.

5. **Recommendation 5: Use technology to explore concepts and analyze data**

STAT 201 and 301 have incorporated technology in two major ways:

- All analyses are performed using JMP, which is a popular and powerful data analysis package that is also user-friendly.

- Important concepts are demonstrated using visual simulation apps, most of which we created in-house.

Before the 201 / 301 redevelopment, students would calculate statistics either by hand or using a TI-83 or -84 calculator. Now they use JMP, which is a Graphical User Interface based data analysis platform. We chose JMP because a) it is something real researchers use, b) it is the most user friendly software package we are aware of, and c) it gives students a much more realistic data analysis experience than they would get using a pencil and paper or a hand calculator.

The GAISE suggestions for teachers on this recommendation include:
• Perform routine computations using technology to allow greater emphasis on interpretation of results.
• View the primary goal as discovering concepts rather than covering methods.
• Implement computer-intensive methods to find p-values and de-emphasize t-, normal and other probability tables. Analyze large, real, data sets.
• Perform simulations to illustrate abstract concepts.

These suggestions have all been implemented in STAT 201 and 301. Students spend much less time calculating numbers and more time interpreting them. Students use a visual interactive app to find p-values, rather than looking them up on a table of finding them using functions in a calculator. This app graphically represents the p-value as the area under a distribution, which helps students conceptualize what p-values quantify. And students are shown a wide variety of statistical topics in a visual manner, using a mix of apps that we developed ourselves and apps developed at other universities and made freely available.

We want to note also that students do not pay extra for any of this. The Statistics Department pays for the JMP site license at CSU, and all students are given a one year license free of charge.

6. **Recommendation 6: Use assessments to improve and evaluate student learning**

STAT 201 and 301 now use iClicker questions to give students frequent feedback on course topics. We have also developed an extensive set of electronic homework assignments that give students multiple attempts at solving a variety of problem types.

GAISE does suggest some things here which we are not able to implement, such as collaborative projects and written assignments. We do not have the resources to effectively grade and give feedback on these types of assignments. We have, however, made it easy for students to work collaboratively on their homework assignments by establishing the Statistics Success Center, which has been well attended.

7. **Omitted topics**

Here, we highlight the aspects of previous versions of STAT 201 / 301 that we have done away with, as per the GAISE recommendations:

• Probability theory has been almost completely taken out. The courses cover how probability is defined and interpreted, and how the concept of probability is integral to the use of statistical analysis. But students are not taught mathematical probability rules, or asked to solve mathematical probability problems.
• With the exception of the class activities, students are not asked to construct plots by hand. Software is used for this.
• Basic statistics (means, medians, histograms) are still covered, but in STAT 301 they are covered quickly. STAT 201 spends a little bit more time on basic topics.
• The use of z-, t-, chi-square, and F- tables has been eliminated completely.

8. **Other updates**

Here are some additional updates we have made to the STAT 201 / 301 curricula:

• We took the old course notes and split them into streamlined sets of lecture slides and supplemental documents. These courses are not taught out of textbooks; students’ primary resource is the course notes. In the past, we posted thorough slides that were covered during lecture. Many students found these to be too wordy and dense, and they encouraged instructors to just “read the slides”, which makes for a boring class period. Now, the slides outline the important topics for lecture and set up discussion questions, and extra details and examples have been moved to supplemental documents that students can read outside of class.
• We have created videos of worked examples that students can view outside of class. One common request from students is “more examples”, but we have limited time for this during class. So we have developed additional worked examples along with videos that go through the solutions. We are also in the process of creating shorter topics videos that students can view to get a quick overview of important topics.

• We have developed large sets of exam practice problems, including some quite challenging problems that let students test themselves on their understanding prior to taking exams. We make these available a week before the exam, and then post solutions a few days later, encouraging students to attempt the problems before seeing solutions. We dedicate some class time for students to work through the more challenging problems in groups, and we hold 2.5 hour long evening review sessions before each exam, which have been well attended.