It is easy to lie with statistics. It is hard to tell the truth without it.

Andrejs Dunkels

Sta 101.001: Data Analysis and Statistical Inference   Fall 2013

Professor:	Dr. Mine Çetinkaya-Rundel - mine@stat.duke.edu Old Chemistry 213
Teaching Assistants:	Anthony Weishampel - anthony.weishampel@duke.edu George Helman - george.helman@duke.edu Yue Dai - yue.dai@duke.edu Christine Chai - christine.chai@duke.edu
Lecture: **Lab:**	Tuesdays and Thursdays, 3:05 - 4:20, French Sci 2231 Mondays at Old Chem 101 10:05-11:20am, 11:45am-1pm, 1:25-2:40pm, 3:05-4:20pm, 3:40pm - 5:55pm

Required materials:

Textbook	OpenIntro Statistics, Diez, Barr, Çetinkaya-Rundel CreateSpace, 2nd Edition, 2011, ISBN: 978-1478217206 PDF available for free at http://www.openintro.org or paperback copy on Amazon.
Clicker	i>clicker2. ISBN: 1429280476, available at the Duke textbook store or for slightly cheaper on the i>clicker website or on Amazon. For a list of former students selling their used clickers, click here.
Calculator	(Optional) You might need a four function calculator that can do square roots for this class. There is no limitation on the type of calculator you can use.

Course Website: http://stat.duke.edu/courses/Fall13/sta101

Support:

Office hours	Tuesdays 4:30-6:30pm and Wednesdays 1:30-3pm I'm also available to talk after class and by appointment. You're highly encouraged to stop by with any questions or comments about the class, or just to say hi and introduce yourself. Most problem sets will be due on Thursdays, I recommend that you attempt all problems by Wednesday so that you can come to office hours with questions.
SECC	Sunday - Thursday 4pm - 9pm (Old Chemistry 211A) The statistics education center has upper level statistics students available to help you. For more information and a schedule see http://stat.duke.edu/courses/resources-students.

Exams: Midterm: Tuesday, October 8

Final: Friday, December 13 (2-5pm)

Course goals & objectives:

This course introduces students to the discipline of statistics as a science of understanding and analyzing data. Throughout the semester, students will learn how to effectively make use of data in the face of uncertainty: how to collect data, how to analyze data, and how to use data to make inferences and conclusions about real world phenomena.

The course goals are as follows:

1. Recognize the importance of data collection, identify limitations in data collection methods, and determine how they affect the scope of inference.

2. Use statistical software to summarize data numerically and visually, and to perform data analysis.

3. Have a conceptual understanding of the unified nature of statistical inference.

4. Apply estimation and testing methods to analyze single variables or the relationship between two variables in order to understand natural phenomena and make data-based decisions.

5. Model numerical response variables using a single explanatory variable or multiple explanatory variables in order to investigate relationships between variables.

6. Interpret results correctly, effectively, and in context without relying on statistical jargon.

7. Critique data-based claims and evaluate data-based decisions.

8. Complete two research projects: one that employs simple statistical inference and another that employs more advanced modeling techniques.

Tips for success:

1. Complete the reading before a new unit begins, and then review again after the unit is over.

2. Be an active participant during lectures and labs.

3. Ask questions - during class or office hours, or by email. Ask me, the TAs, and your classmates.

4. Do the problem sets - start early and make sure you attempt and understand all questions.

5. Start your projects early and and allow adequate time to complete them.

6. Give yourself plenty of time time to prepare a good cheat sheet for exams. This requires going through the material and taking the time to review the concepts that you're not comfortable with.

7. Do not procrastinate - don't let a week go by with unanswered questions as it will just make the following week's material even more difficult to follow.

Course structure:

The course is divided into seven learning units. For each unit a set of learning objectives and required and suggested readings, videos, etc. will be posted on the course website. You are expected to complete the readings and familiarize yourselves with the learning objectives prior to beginning the unit. We will begin the unit with a readiness assessment: multiple choice questions that you answer using your clickers at the beginning of class. You will then re-take this assessment as a team. The rest of the class time will be split between lecture, during which you'll get to interact and participate via clickers, and application exercises that you'll complete in your teams. Each unit will wrap up with a performance assessment.

Grading:

Your final grade will be comprised of the following.

Clicker questions: 5%	Project 1: 10%
Problem sets: 7.5%	Project 2: 10%
Labs: 7.5%	Midterm: 15%
Readiness assessments: 12.5%	Final: 25%
Performance assessments: 2.5%	Peer evaluations: 5%

Cumulative numerical averages of 90 - 100 are guaranteed at least an A-, 80 - 89 at least a B-, and 70 - 79 at least a C-. These ranges may be lowered, but they will not be raised, e.g., if everyone has averages in the 90s, everyone gets at least an A-. The exact ranges for letter grades will be determined after the final exam. The more evidence there is that the class has mastered the material, the more generous the curve will be.

Work load:

You are expected to put in 4-6 hours of work outside of class. Some of you will do well with less time than this, and some of you will need more.

Teams:

To construct highly functional teams of learners, you are asked to complete a short survey as well as a pre-test to gauge your previous exposure to statistics and statistical literacy. You will complete these during the first lab.

You will be assigned to teams of 4-5 students based on the results of the survey and the pre-test. Once team assignments have been made there is no option for changing teams, other than under extraordinary circumstances. You will work in these teams during labs and application exercises as well as on your second project. In addition, your team members will be your first point of contact in this class. You should always sit with your team, in lab and in lecture.

You are encouraged to study with your team members and other classmates. But remember that anything that is not explicitly a team assignment must be your own work.

Lectures:

Lecture slides will be posted by noon the day of the lecture, so that you can print them before lecture and bring them with you to class. Note that these slides are not intended to be exhaustive and will be a poor substitute for taking your own notes during the course of the lectures. You are responsible for all the material covered in all components of the course, not just the lectures. Please ask questions in class, office-hours or by e-mail if you are struggling (or just curious), do not wait until just before an exam when it may be too late.

Discussion Forum on Sakai:

Any non-personal questions related to the material covered in class, problem sets, labs, projects, etc. should be posted on the Discussion Forum on Sakai. Before posting a new question please make sure to check if your question has already been answered. The TAs and myself will be answering questions on the forum daily and all students are expected to answer questions as well.

Clicker questions (5%):

Throughout the lectures you will use clickers to answer conceptual questions and for data collection/class surveys. In order to receive credit for the clicker questions you must register your clicker at http://iclicker.com/support/registeryourclicker. In the Student ID field enter your Net ID, and in the Remote ID field enter the alphanumeric code printed below the barcode on the back of your clicker.

These questions are on material introduced in class that day, and you get credit for clicking in, regardless of whether you have the correct answer. To get credit for the day you must respond to at least 75% of the questions. The objective of these questions is to help make you an active participant and gauge the class' mastery of the material.

You are required to bring your clicker to every lecture and it is your responsibility to show up to class on time. Most importantly, it is your responsibility to come to class. I realize that occasionally you may be late, forget your clicker, or need to miss class. *Up to three* unexcused late arrivals or absences will not affect your clicker grade. If one person is simultaneously using two or more clickers, the all owners of the clickers will receive a 0 for an overall clicker grade, and will be reported to the Office of Student Conduct.

Problem sets (7.5%):

These will be assigned (approximately) weekly on the course webpage and will be comprised of problems from the textbook. Each assignment will list roughly five to seven problems from the book to be turned in for grading, and roughly 10 practice problems. You do not need to turn in the practice problems, and the solutions can be found in the back of the book.

The objective of the problem sets is to help you develop a more in-depth understanding of the material and help you prepare for exams and projects. They will be graded out of 20 points and grading will be based on completeness as well as accuracy. In order to receive credit you must *show all your work. Lowest score will be dropped.*

You are welcomed, and encouraged, to work with each other on the problems, but *you must turn in your own work.* If you copy someone else's work, both parties will receive a 0 for the problem set grade as well as being reported to the Office of Student Conduct. Work submitted on Sakai will be checked for instances of plagiarism prior to being graded.

You will turn in your problem sets on Sakai by *typing* your answers in the assignment form. You might want to also maintain a copy of your work in a word processor. All assignments will be time stamped and late work will be penalized based on this time stamp (see late work policy below).

Labs (7.5%):

The objective of the lab is to give you hands on experience with data analysis using modern statistical software. The labs will also provide you with tools you will need to complete the projects successfully. We will use a statistical analysis package called RStudio, which is a front end for the R statistical language. You can use RStudio on any computer with a web browser. To get an RStudio account you need to provide me with your Gmail address.[1]

During labs you will work in teams. Each week one team member will be responsible of typing up the responses to the Exercises and the On Your Own part of the lab, while the entire team engages in discussion and comes up with the necessary steps to answer these questions. You should be able to complete your lab report by the end of the lab period and turn it in to your TA. If you need additional time you will generally have until the following Monday (check the course schedule for exact due dates). *Lowest score will be dropped.* If you *do not attend a particular lab section*, you are *not eligible for credit on that lab.*

---

[1]If you don't have a gmail account, you can create one at https://www.google.com/accounts/NewAccount.

Readiness assessments (12.5%):

Individual and team readiness assessments will be given at the beginning of a unit. These are multiple choice assessments comprised of conceptual questions addressing the learning objectives of the new unit. You are not expected to master all topics in the unit ahead of time, but you are responsible for completing the reading assignment, understanding how the material fits in the greater framework of the course, and acquire a conceptual understanding of the learning objectives. As described above, you will first take the individual readiness assessment using your clickers, and then re-take the same assessment in teams. Your performance on both assessments will factor into your final grade. *Lowest score will be dropped.*

Performance assessments: (2.5%)

Performance assessments will be given at the end of a unit. These are very similar to the readiness assessments in format, however you will be taking them outside of class on Sakai. You will need to complete the assessment in one sitting within a 24-hour period from when it is posted, and they will be due at 11:55pm on the listed due date. Outstanding performance will require mastery of all topics in the unit. *Lowest score will be dropped.*

Projects (10% × 2):

The objective of the projects is to give you independent applied research experience using real data and statistical methods. Project 1 will be completed *individually*. For a parameter of interest to you, you will describe the relevant data, compute a confidence interval and conduct a hypothesis test, and summarize your findings in a report. Project 2 will be completed *in teams*. This will be a multiple regression project. You will turn in a write up summarizing your findings as well as present your findings during the last lab session. Note that you *must complete both projects and score at least 30% of the points on each project in order to pass this class.*

Midterm (15%) and final (25%):

The midterm and final exam dates are listed on the first page of the syllabus. Exam dates cannot be changed and no make-up exams will be given. If you can't take the exams on these dates you should drop this class. *You can't pass this class if you do not take the final exam.*

You are allowed to bring one sheet of notes ("cheat sheet") to the midterm and the final. This sheet must be no larger than $8\frac{1}{2}$" × 11", and *must be prepared by you.* You may use both sides of the sheet.

Peer evaluations (5%):

Throughout the semester you will be asked to complete six peer evaluations. These will be used to ensure that all team members contribute to the success of the group and to address any potential issues early on. If you feel that there are issues within your team, you are encouraged to discuss it with your team members and to bring it to my or your TA's attention.

Email:

I will regularly send announcements by email, please make sure to check your email daily. While email is the quickest way to reach me outside of class, it is often more efficient to answer statistical questions in person. Note that all material related questions should be posted on the appropriate Sakai forum, and email should be reserved for personal communication.

Other learning resources:

Aside from the TAs and the professor's office hours, you can also make use of the Academic Resource Center.

For more information, see http://web.duke.edu/arc.

Students with disabilities:

Students with disabilities who believe they may need accommodations in this class are encouraged to contact the Student Disability Access Office at (919) 668-1267 as soon as possible to better ensure that such accommodations can be made.

Academic integrity:

Duke University is a community dedicated to scholarship, leadership, and service and to the principles of honesty, fairness, respect, and accountability. Citizens of this community commit to reflect upon and uphold these principles in all academic and non-academic endeavors, and to protect and promote a culture of integrity. Cheating on exams and quizzes, plagiarism and copying others' work on homework assignments and projects, lying about an illness or absence and other forms of academic dishonesty are a breach of trust with classmates and faculty, violate the Duke Community Standard, and will not be tolerated. Such incidences will result in a 0 grade for all parties involved as well as being reported to the Office of Student Conduct. Additionally, there may be penalties to your final class grade. Please review the Duke's Academic Dishonesty policies at http://www.studentaffairs.duke.edu/conduct/resources/academicdishonesty.

Excused Absences:

Students who miss graded work due to a scheduled varsity trip, religious holiday or short-term illness should fill out an online NOVAP, RHoliday or short-term illness form respectively and will be given the grade of their final exam for those items. If you will be absent the day a graded work is due, you are still responsible for completing the assignment and it is your responsibility to make arrangements with me in a timely fashion. Those with a personal emergency or bereavement should seek a Dean's Excuse; check with your academic dean for details.

Policies:

- Late work policy for problem sets and labs reports:

  - work submitted during class: lose all points
  - next day: lose 30% of points
  - after class on due date: lose 10% of points
  - later than next day: lose all points

- Late work policy for projects: 10% off for each day (24-hour period) late.

- There will be no make-up for clicker questions, individual and team readiness assessments, labs, problem sets, projects, or exams. If the midterm exam must be missed, absence must be officially excused *in advance*, in which case the missing score will be imputed using the final exam score. Missed midterms not excused in advance will receive a grade of 0. The final exam must be taken at the stated time.

- Regrade requests must be made *within one week* of when the assignment is returned, and must be submitted in writing. These will be honored if points were tallied incorrectly, or if you feel your answer is correct but it was marked wrong. No regrade will be made to alter the number of points deducted for a mistake. There will be no grade changes after the final exam.

- Clickers may not be shared, and the clicker registered to a person may only be used by that person. Failure to abide by this will result in a 0 clicker grade for everyone involved.

- Use of disallowed materials (including any form of communication with classmates or looking at a classmate's work) during exams will not be tolerated. This will result in a 0 on the exam for all students involved, possible failure of the course, and will be reported to the Office of Student Conduct. If you have any questions about whether something is or is not allowed, ask me beforehand.