ISTA 116: Statistical Foundations for the Information Age

Dr. Derek T. Green

Last Revised January 11, 2012

1 Course Information

1.1 Course Description

Understanding uncertainty and variation in modern data: data summarization and description, rules of counting and basic probability, data visualization, graphical data summaries, working with large data sets, prediction of stochastic outputs from quantitative inputs. Operations with statistical computer packages such as R.

1.2 Prerequisites

MATH 109C or MATH 112 or placement beyond the level of College Algebra.

1.3 Course Goals

This course is intended as a first exposure to the nature of uncertainty and variation in data typical of the information age. Students will learn techniques for descriptive statistics, both quantitative and graphical, and will make use of these with real (often large) data sets in lab sessions and in homework. Several weeks are devoted to introducing basic concepts in probability, as a foundation for later study in inferential statistics. The statistical computer package `R` is used for illustration and exploration, as well as to demonstrate the central role of computing in modern data science.

1.4 Locations and Times

Lectures: MW 4:30-5:45
Location: C E Chavez Bldg, Rm 110

Lab: One 2 hr section per week (Tue 12:00-1:50, Tue 4:00-5:50, Wed 12:00-1:50)
Location: Gould-Simpson, Rm 930

1.5 Textbooks

NOTE: All of the books for this course are available online through the links given below.

1.5.1 Required

Statistics Concepts:

Griffiths, D. (2009). *Head First Statistics*. Sebastopol, CA: O'Reilly Media, Inc.
Available online through the University of Arizona bookstore:
`http://proquest.safaribooksonline.com.ezproxy2.library.arizona.edu/book/statistics/9780596527587`

Using `R`:

Owen, W. J. (2010). *The R Guide, ver. 2.5*. Department of Mathematics and Computer Science, University of Richmond, Richmond, VA.
Available for free online at:
`http://cran.r-project.org/doc/contrib/Owen-TheRGuide.pdf`

1.5.2 Recommended Additional `R` Reference

Adler, J. (2010). *R in a Nutshell*. Sebastopol, CA: O'Reilly Media, Inc.
Available online through the University of Arizona bookstore:
`http://proquest.safaribooksonline.com.ezproxy2.library.arizona.edu/book/programming/r/9781449377502`

1.6 Instructor Information

Dr. Derek T. Green
Office: Gould-Simpson 830
Email: dtgreen@email.arizona.edu
Office Hours: Mon 12:30-1:30, Wed 2:00-4:00, and by appointment

1.7 Lab Instructors

Jeremy Wright
wrightjb@email.arizona.edu
Office Hours:
Mon 2:00-3:00 and Tue 2:00-4:00
Location: Gould-Simpson 930

Jeremy Samuelson
jjsamuel@email.arizona.edu
Office Hours:
Mon 3:00-4:00 and Thur 9:30am - 10:30am
Location: Gould-Simpson 228

2 Policies and Grading

2.1 Grading

Grades are based on homework, quizzes, one midterm, and a comprehensive final exam. The grading scheme is as follows:

Component	**Weight**
Lab Assignments	400 pts
In-Class Quizzes	100 pts
Web Quizzes	50 pts
Midterm Exam	200 pts
Final Exam	200 pts
Exam Adjustment	50 pts
Total	1000 pts

Grade	**Point Range**
A	900-1000
B	800-899
C	700-799
D	650-699
E	0-649

Grade Disputes

Disputes about grades on a particular assignment or exam will be entertained for two weeks from the day the assignment is returned, or one week from the date of the final exam, whichever is sooner. These will be resolved by re-grading the entire assignment or exam. Note that this can result in a lower grade in the event that new mistakes are discovered. The final exam will be graded and made available for review by students within 48 hours of its completion, to allow time for any requested regrades.

"Curve" Policy

The instructor may decide at the end of the term to lower the thresholds stated above for all students. Any decision to do so will be based solely on aggregate class performance, and not on circumstances pertaining to any individual. Once any grade disputes have been resolved for individual assignments, and the final thresholds have been set for all students, the total number of points earned will automatically determine the final grade. **No negotiations about individual students' letter**

grades will be entertained once final grades are assigned, except as permitted by the policy stated above.

2.2 Homework Assignments

There will be 8 homework assignments, consisting of a mixture of paper-and-pencil and computational problems. Individual assignments are worth 50 points. Assignments are graded for correctness; however, genuine effort is nearly always worth partial credit. Homework will typically be assigned during lecture on a Monday and will be due two weeks later on Monday before midnight (see individual assignments for actual due dates).

Collaboration Policy

Students may discuss homework assignments with the instructor, TAs and each other, however, **each student must submit her or his own solutions**. Any fellow student collaborators must be identified along with the nature of the collaboration (e.g., "we worked together", "she helped me", "I helped him"), and any asymmetric collaboration should be limited to hints or explanation of concepts, and not detailed walkthroughs of particular solutions.

Late Policy

Assignments are due at the end of the first lab section after the announced due date, unless otherwise specified. Permission for an extension must be granted by the lab instructor in advance of the deadline in order to receive full credit for a late assignment. The lab instructors are under no obligation to grant such an extension.

Calculation of Final Homework Grade

Each student's lowest homework grade will be replaced by her/his highest homework grade, to produce a final score out of 400.

2.3 Quizzes

In-Class Quizzes

Several unnanounced quizzes will be given at the beginning of lectures throughout the semester, each consisting of one or two short and straightforward questions designed to demonstrate that students are paying attention in class and doing the required reading. Students who do so should find the quizzes easy.

The five highest quiz scores will be used to calculate the final quiz grade. **No make-up quizzes will be given for any reason**.

Web Quizzes

Web quizzes will occasionally be assigned via d2l. These are intended to provide practice with problems more like those that will appear on exams (and hence can be done without the computer). Students are encouraged to make at least their first attempt without any reference materials. These are graded for completion only.

2.4 Exams

One in-class exam will be given during the lecture on March 7^{th}. A comprehensive final exam will be given on Wednesday May 9^{th} from 3:30 P.M. to 5:30 P.M. The format of the exams will be mixed, and may include multiple choice, short verbal answer, and "math problems". Each exam is graded out of 200 pts. Each student's higher exam grade will be rescaled to 250 points.

Permitted Materials

Students will be allowed to use one double-sided $8\frac{1}{2} \times 11$" page of notes for the midterm, and two pages for the final. A simple, self-contained hand calculator may also be used. **No other electronic devices of any kind, including cell phones, iPods, laptops, etc., will be permitted during the exams**.

Make-up Policy

Exams may only be made up in case of a serious, unanticipated emergency. The student or an appointed proxy must contact the instructor in advance of the exam. Clearly worded documentation, from a doctor, dean, etc., demonstrating physical inability to take the exam at the appointed time, will be required to grant a make-up appointment. **No make-ups will be granted for personal reasons, including travel or personal hardship**. Any make-up exams that are granted must be scheduled for after the original time.

2.5 University Policies

Classroom Behavior

Students are expected to behave respectfully toward each other and to the instructor and TAs. Disrespectful behavior includes the use of cell phones or other electronic devices in the classroom during class hours.

The Arizona Board of Regents Student Code of Conduct is here: `http://dos.web.arizona.edu/uapolicies/scc5308abcd.html#sccphilosophy`

ABOR Policy 5-308, prohibits threats of physical harm to any member of the University community, including to oneself. See: `http://policy.web.arizona.edu/~policy/threaten.shtml`.

Special Needs and Accommodations

Students who need special accommodation or services should contact the

Disability Resources Center
1224 East Lowell Street, Tucson, AZ 85721
(520) 621-3268
FAX (520) 621-9423
email: `uadrc@email.arizona.edu`
web: `http://drc.arizona.edu/`.

You must register and request that the Center or DRC send official notification of your accommodations needs as soon as possible. Please plan to meet with the

instructor by appointment or during office hours to discuss accommodations and how the course requirements and activities may impact your ability to fully participate. The need for accommodations must be documented by the appropriate office.

Student Code of Academic Integrity

Students are encouraged to share intellectual views and discuss freely the principles and applications of course materials. However, graded work/exercises must be the product of independent effort unless otherwise instructed. Students are expected to adhere to the UA Code of Academic Integrity as described in the UA General Catalog. See: `http://dos.web.arizona.edu/uapolicies/`.

Confidentiality of Student Records

See `http://www.registrar.arizona.edu/ferpa/default.htm`

Subject to Change Statement

Information contained in this syllabus, other than the grade and absence policy, may be subject to change by the instructor, with advance notice.

3 Tentative Schedule of Lecture Topics and Readings

HF: Head First Statistics
RG: The R Guide

Topic	**Approximate Dates**	**Assigned Readings**	**Add'l Lab Readings**
Syllabus/Introduction	Jan. 11		
Visualizing Information	Jan. 18, 23	HF: xxxiii-xxxv HF(1): 1-43,644	RG: 1.1-1.8, 2.1 RG: 3.3
Measuring Central Tendency	Jan. 25, 30	HF(2): 45-81	
Measuring Variability & Spread	Feb. 1, 6, 8	HF(3): 83-125	RG: 5.1, 5.2
Correlation & Regression	Feb. 13, 15, 20	HF(15): 605-640	
Calculating Probabilities	Feb. 22, 27, 29	HF(4): 127-195	
Catch-up and Review	Mar. 5		
MIDTERM EXAM	***Mar. 7***		
Discrete Probability Distributions	Mar. 19, 21	HF(5): 197-240	
Permutations & Combinations	Mar. 26	HF(6): 241-268	
Discrete Random Variables The Bernoulli, Binomial and Discrete Uniform Distributions	Mar. 28, Apr. 2	HF(7): 269-324	RG: 6.1, 6.3.1
Using The Normal Distribution	Apr. 4, 9, 11	HF(8,9):325-359	
Using Statistical Sampling	Apr. 16, 18, 23	HF(10): 415-439	RG: 6.4
Sampling Distributions The Central Limit Theorem Simulation and Bootstrapping to Estimate Sampling Distributions	Apr. 25, 30	HF(11):441-466	RG: 6.1-6.3, 6.6
Catch-up and Review	May 2		
FINAL EXAM	***May 9***		

4 Tentative Schedule of Lab Topics

Dates	Topics
Jan. 17-18	Intro and `R` Fundamentals
Jan. 24-25	`R` Fundamentals Cont'd, Tables, Plotting Bar Charts, Pie Charts
Jan. 31, Feb. 1	Summary Statistics, Stem and Leaf Plots, Strip Charts, Histograms, Density Curves
Feb. 7-8	Visualizing Central Tendency and Variability, Transformations, Boxplots
Feb. 14-15	Contingency Tables, Margin Tables, Conditional Proportion Tables
Feb. 21-22	Overlaid Densities, Multiple Boxplots, Summary Statistics by Group, Scatterplots, Correlation
Feb. 28-29	Linear Regression with `lm()` Prediction, Residual Plots
Mar. 6-7	Midterm Review
Mar. 13-14	*Spring Break*
Mar. 20-21	Probability Basics
Mar. 27-28	Sampling from a Population With and Without Replacement using `sample()`
Apr. 3-4	Computing Binomial Probabilities and Quantiles in `R`
Apr. 10-11	Normal Probabilities and Quantiles
Apr. 17-18	Sampling from a Distribution
Apr. 24-25	Repeated Sampling, Visualizing Distributions of Sample Statistics
May 1-2	Approximating Sampling Distributions with the CLT and with Simulation
May 1-2	Bootstrapping, Catch Up