

STAT411-2 - Methods of Data Analysis I

Spring 2018

9am MWF JABS 407

3:10pm W ABB 234

Instructor: Dr. Al Parker, Barnard(EPS)304, 994-5145, parker@math.montana.edu.

Office Hours: Mondays and Fridays, 12pm-1pm; other times by appointment.

Course Web Page: <http://www.math.montana.edu/parker/courses/STAT411>

Text: Ramsey & Schafer (2013). *The Statistical Sleuth: A Course in Methods of Data Analysis (3rd Edition)*, Duxbury Press. Book website: www.statisticalsleuth.com. All of the data presented in the book is available in the R package *Sleuth3*.

Software: The FREE and powerful software R will be used throughout the semester for case studies, homeworks, project (STAT511), and exam questions. All of the following links can be found on the course website. R's homepage: <https://www.r-project.org/>. Download R: <http://cran.r-project.org/>. RStudio: <https://www.rstudio.com/products/rstudio/download>. Help with R: the book *An Intro to R* <http://cran.fhcrc.org/doc/manuals/r-release/R-intro.pdf> and the web page <http://www.math.csi.cuny.edu/Statistics/R/simpleR>.

Prerequisite for STAT411: STAT 217 or STAT 332; **Prerequisite for STAT511:** graduate standing and either STAT216 or STAT401

Topics: randomization, permutation, t , and F tests; ANOVA and linear contrasts; multiple regression. STAT411 provides an in-depth examination of these topics, focuses on communicating statistical results to others in a clear and concise manner, and emphasizes the appropriate scope of inference for a study. By the end of the course, you will be able to critically evaluate the inferential statements made in research papers. In the text, we will cover Chapters 1-3, 5-12.

Participation: You are expected to actively participate in class by verbally asking questions (e.g., from the required reading or from the last homework) and by verbally answering questions posed to you. Questions posed to you in class are designed to (1) assist in your learning rather than to demonstrate what you have learned; (2) reinforce concepts; and/or (3) motivate new topic(s). If you prefer to “pass” on a question, no worries, just respond “pass”.

Grading is based on about 650 points (STAT411), 715 points (STAT511):

- Midterm exam on Friday, March 2: 200 points ($\approx 30\%$)
- Comprehensive Final Exam on Thursday, May 3, 6:00–7:50pm: 200 points ($\approx 30\%$)
- Homeworks: about 25 points each for a total of about 250 points ($\approx 40\%$)
- STAT511 project: 65 points ($\approx 10\%$)

Letter grades will be assigned according to the following percentages:

F 0-59, D 60-64, D+ 65-69, C 70-74, C+ 75-79, B 80-84, B+ 85-89, A 90-100.

Exams: All exams are in class, worked individually, and are “closed book.” Calculators are allowed, cell phones are not. If an extreme circumstance occurs (e.g., illness or previously scheduled activities vital to academic program), notify me prior to the exam and provide appropriate documentation. Otherwise, there are no make-up exams.

Homeworks: Homework will be assigned roughly each week and usually due on Mondays. To receive full credit, homeworks must be turned in when due. To have late work considered for partial credit, you must notify me ahead of time. Data analysis problems must be submitted in the format described in the *Writing a Statistical Report* guidance. These reports will be graded for content as well as for the organization and quality of the write-up. Of particular note:

- Reports must be typed and no longer than 2 pages.
- Your report should have 4 sections with a title for each section: Introduction, Statistical procedures used, Summary of statistical findings, Scope of inference.
- Results and statements in response to specific questions in a report must be substantiated by calculations, a logical argument, and/or computer code and output. If you have a question about which questions require justification, then ask.
- The R code and output used to complete a homework are to be included in an Appendix in your report. Never just paste in computer code to answer a question. Answer each question using proper English sentences, then reference the appropriate computer output if appropriate.
- Any Table or Figure that you generate must be included in the Appendix, labeled and referenced from the body of your report.

Labs: Labs will be held every week at 3:10pm on Wednesdays in ABB 234 to help you use R. A lab may pertain to a current homework or it may involve work on a separate problem set. Attending labs is a required part of the course. Labs may be collected and graded and counted towards the homework part of your grade.

Learning outcomes: Upon completing this course, a student will be able to: explain the fundamental concepts of statistical inference; explain the connections among the randomization, permutation, sampling, and t distributions; recognize the connections between study design and appropriate statistical analysis and scope of inference; demonstrate understanding of a basic linear model using the reference-level parameterization and the meaning of the parameters of the model; demonstrate ability to appropriately tie a question of interest to a model parameter; demonstrate understanding of the assumptions underlying common models and methods, and how to critique and check appropriateness of those assumptions; demonstrate a basic understanding of using statistical software to perform data analysis; demonstrate a basic understanding of using computer software to investigate statistical concepts through simulation; demonstrate ability to write statistical reports including justification for assumptions and decisions made, a concise summary of findings interpreting results in the context of the problem, and a written description of the scope of inference of the findings; demonstrate ability to critically evaluate conclusions made from a statistical analysis; demonstrate ability to recognize common mistakes made in statistical analysis, interpretation, and inference in general.

Help: When working on homeworks, you may provide help to and/or receive help from any of your fellow classmates. What a great way to learn! Each student must hand in an *independent* write-up of each homework unless the homework explicitly states otherwise. I will always make time to help answer your questions. Feel free to email questions. Email is the best way to set up a time to meet with me outside of office hours. If you email me with a question about R, paste in any relevant R-code and R-output.

Disabilities: If you have a documented disability for which you may be requesting an accommodation, contact me and Disabled Student Services as soon as possible.