STAT 5302
Intermediate Data Analysis II

lecture:	MWF 11:30am–12:25pm in Cockins Hall 232
instructor:	Vincent Q. Vu (vqv at stat osu edu)
office:	Cockins Hall 325
office hours:	Wed 1:00pm-3:00pm, or by appointment
web:	Class schedule, assignments, and course announcements will be posted on Carmen (carmen.osu.edu)
prerequisites:	STAT 5301 or permission of the instructor

1 Overview

STAT 5302 is the second course in a two-semester sequence in Intermediate Data Analysis (STAT 5301–5302). We assume that students are familiar with organizing and summarizing data, the nature of relationships between variables, sampling distributions and the underlying rationale for hypothesis tests and confidence intervals. We also assume that students are comfortable with a variety of models and inferential procedures. Specifically, the material in 5302 relies heavily on the additive model (see the early part of the text for a description of this model) and one-way ANOVA. The course will cover simple linear regression, multiple linear regression, and two-way (and multi-way) ANOVA. For each of the common statistical methods covered in the course, we will focus on generation of appropriate models for data, estimation of the model parameters and their inference, and model diagnostics. Applications of the methods will be illustrated with data analysis.

2 GE goals and expected learning outcomes

STAT 5302 satisfies the general education (GE) requirement in Data Analysis.

Goals: *Students develop skills in drawing conclusions and critically evaluating results based on data.*

Expected learning outcomes: *Students understand basic concepts of statistics and probability, comprehend methods needed to analyze and critically evaluate statistical arguments, and recognize the importance of statistical ideas.*

3 Course materials & computing

Homework problems, reading assignments, and lecture material will be drawn from the required textbook:

- *The Statistical Sleuth: A Course in Methods of Data Analysis* (Third Edition) by F.L. Ramsey and D.W. Shafer

You will be required to do some basic statistical analyses on the computer using the R software environment for your assignments. I recommend interacting with R via the RStudio IDE (integrated development environment)—RStudio requires R to be installed. Both R and RStudio are free, open-source software and can be downloaded from the following websites.

- R (www.r-project.org)
- RStudio (www.rstudio.com)

Additional information on R and RStudio will be distributed on the course website.

4 Tentative course schedule

The following is a tentative schedule of topics. Students are expected to read the indicated chapters from the textbook concurrently with lectures. Further details will be announced on the course website.

Week	Date	Topic	Reading
1	1/8–	Introduction; simple linear regression model; least squares regression estimation	7.1–7.3
2	1/15–	Inferential tools for regression; estimation of the distribution the response at some value of the explanatory variable; prediction of a future response	7.4
*	1/15	**Martin Luther King Jr. Day (no class)**	
3	1/22–	Assumptions of linear regression; graphical tools for model assessment	8.1–8.3, 8.6.3
4	1/29–	Interpretation after log transformations; ANOVA for regression; R^2; regression vs one-way ANOVA	8.4–8.6
5	2/5–	Multiple linear regression model; interpretation of regression coefficients; specially constructed explanatory variables	9.1–9.3
6	2/12–	Least squares estimation for multiple linear regression; strategies for data analysis; graphical methods for exploration	9.4–9.5
7	2/19–	Inferences about regression coefficients; predictions	10.1–10.2, 10.4.3
8	2/26–	F-tests and ANOVA for multiple linear regression; adjusted R^2; related issues	10.3–10.4
*	3/2	**Midterm Exam**	
9	3/5–	Model checking with residual plots; model refinements	11.1–11.2, 11.5

Week	Date	Topic	Reading
*	3/12–	**Spring Break (no class)**	
10	3/19–	Influential observations; case-influence statistics; weighted least squares	11.3–11.4, 11.6
11	3/25–	Issues relating to many explanatory variables; sequential variable-selection techniques;	12.1–12.3
12	4/2–	All subsets regression; model selection criteria; Bayesian model selection/averaging; related issues	12.4–12.7
13	4/9–	Two-way ANOVA; additive/nonadditive models for two-way tables; alternative parameterizations for two-way ANOVA	13.1–13.2, 13.5
14	4/16–4/23	Testing for factor effects; examining interactions; sequential sum of squares; tests for block effects; multiple comparisons	13.3–13.5
*	4/27	**Final Exam (12:00pm – 1:45pm)**	

5 Coursework & Grading

There will be homework, a midterm exam, and a final exam.

- 30% Homework (approximately bi-weekly; posted on Carmen; due at the beginning of class)
- 30% Midterm exam (March 2, 2017)
- 40% Final exam (April 27, 2017; 12:00pm – 1:45pm)

Late homework will not be accepted. No make-up exams will be given. The final exam will be cumulative, but will emphasize recent material. Exam rules will be announced in class two weeks before the exam date.

6 Academic misconduct

It is the responsibility of the Committee on Academic Misconduct to investigate or establish procedures for the investigation of all reported cases of student academic misconduct. The term "academic misconduct" includes all forms of student academic misconduct wherever committed; illustrated by, but not limited to, cases of plagiarism and dishonest practices in connection with examinations. Instructors shall report all instances of alleged academic misconduct to the committee (Faculty Rule 3335-5-487). For additional information, see the Code of Student Conduct http://studentlife.osu.edu/csc/.

7 Disability services

The University strives to make all learning experiences as accessible as possible. If you anticipate or experience academic barriers based on your disability (including mental health, chronic or tempo-

rary medical conditions), please let me know immediately so that we can privately discuss options. To establish reasonable accommodations, I may request that you register with Student Life Disability Services. After registration, make arrangements with me as soon as possible to discuss your accommodations so that they may be implemented in a timely fashion. SLDS contact information: slds@osu.edu; 614-292-3307; slds.osu.edu; 098 Baker Hall, 113 W. 12th Avenue.