STA 3064 Introduction to Statistical Modeling with SAS
Fall 2020 Course Syllabus

Instructor: Dr. Steven Ramsier
Office: 102B OSB *(Not anticipated to be in the office frequently this semester.)*
Virtual Office Hours (via Zoom): 12:00 Noon to 1:00 PM on Tuesdays
11:30 AM to 12:30 PM on Thursdays.
Email: ramsier@stat.fsu.edu
Phone: 644-3218 (Main Statistics office phone – currently no direct line to the instructor)
Fax: 644-5271

TAs/Graders: Madison Layfield **E-mail:** mil18b@my.fsu.edu
Virtual Help Hour: Wednesdays 10-11 am
Grading: Assignment 1, Case Study, Project B

Mark Zamani **E-mail:** mcz16@my.fsu.edu
Virtual Help Hour: Thursdays 2-3 pm
Grading: Assignments 2, 3, 4, Project A

Class Meeting Times: The class involves asynchronous delivery. You may work on a schedule of your choosing that permits you to meet weekly due dates/times and allows you to work with two or three other students during some weeks in order to complete a project.

Final Period: ***No Final Exam***

Optional Text: Cannon, A.R., Cobb, G.W., et al., *Stat2: Modeling with Regression and ANOVA* (2019), *Second Edition***,** W.H. Freeman and Company.
ISBN-10: 1-319-05407-2
ISBN-13: 978-1-319-05407-6

Reference for Certification Exam: Shreve, J.N., Holland, D.D., *SAS Certification Prep Guide: Statistical Business Analysis Using SAS9* (2018), *First Edition, SAS Institute.*
ISBN-10: 1629603813
ISBN-13: 978-1629603810
Other references will be provided during the course of the semester.

Internet: Online access required for SAS programs and learning management system
Prerequisite: STA3024 or consent of the instructor.
Software: Access to *SAS Studio* (Online, OnDemand version), *SAS University Edition* (Local computer version), or *SAS 9.4* (Windows version available on campus computer labs but differs slightly from the other two versions which will be used in class).

Course Description: This course will cover the following topics utilizing the SAS software: ANOVA, linear modeling, logistic regression, bootstrap sampling, simulation using the data step, and some additional analytic topics.

Students will gain experience in fitting statistical models to data in order to extract information, determine significant factors, and make predictions. Students will be exposed to several popular models including both general linear and generalized linear models in order that they will be able to make informed decisions and appropriate model selection in the future.

Course Objectives: This course is a designated as a *Scholarship in Practice* course.

Scholarship in Practice Objectives. By the end of the course, students will demonstrate the ability to:

- Select, critically evaluate, and apply relevant areas of scholarship to produce an original analysis, project, creative work, performance or other scholarly work that reflects a body of knowledge relevant to the course (*SIP1*);
- Articulate the process of producing a work, from initial plan, to critique, revision, and completion (*SIP2*);
- Critique existing applications of scholarship in order to learn from past success and failures (*SIP3*).

Competencies SIP1, SIP2, and SIP3 will be assessed through the assignments and capstone project for STA 3064 for which a standard rubric will be employed to facilitate a written summary of results.

General Course Objectives. Students who complete this course will be able to:

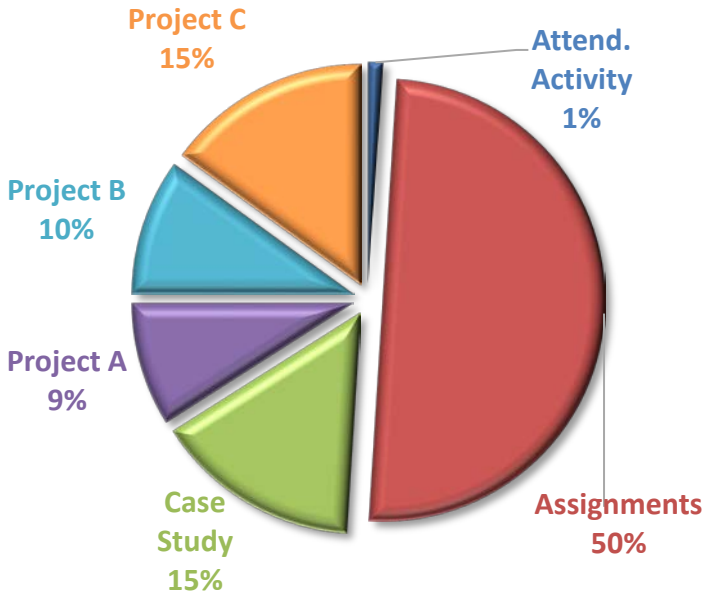
- Choose the appropriate statistical model for given data and situations.
- Analyze differences between population means using the ANOVA and GLM procedures and verify the assumptions.
- Perform pairwise comparisons among groups after finding significant effects in ANOVA.
- Fit a simple linear regression using the REG procedure producing predicted values and confidence intervals.
- Explain the mathematical model for multiple regression.
- Describe model selection options and interpret output to evaluate the fit of several models.
- Perform model diagnostics through residual plots, influential observations identification, and collinearity detection.
- Fit a binary logistic regression model using the LOGISTIC procedure.
- Define and explain the adjusted odds ratio.

Grade Composition (1000 Points Total):

First Day Attendance Activity	10
4 Assignments (125 pts. each)	500
Case Study	150
Project A (Individual Written Project)	90
Project B (Individual Written Project)	100
Project C (Team Presentation Project)	150
Total	1000

Grade Assignments (Based on total course points earned:

A	930-1000	B-	800-829	D+	670-699
A-	900- 929	C+	770-799	D	630-669
B+	870-899	C	730-769	D-	600-629
B	830-869	C-	700-729	F	0-599



Assignments and Responsibilities

Assignments

The assignments will consist of problems that will be solved using SAS. There will be four (4) assignments given and all assignments are to be turned in on the Friday that they are **due no later than 5:00 PM.** Assignment documents are uploaded via Canvas and **no emailed assignments** will be **accepted**.

Late, unexcused assignments will be penalized as follows: turned in by 5:00 PM the following day (Saturday) – 90% of grade, turned in by 5:00 PM the Sunday after it was due – 75% of grade, thereafter – no credit. Assignments are graded on several components: Correct functions and/or procedures, properly executable, correct results, and correct interpretations. Assignments will be submitted electronically through Canvas.

You are free to discuss the assignment with any of your classmates; however, the activity of students "working together" is not permitted. Your programming, interpretation, and write-up must be done independently. That is, all code, output, and explanations must be generated by you. Your interpretations must be in your own words. Sharing documents and using any portion of another student's (past or present) work, representing it as your own, will result in a score of zero on the assignment.

> **Warning about Using SAS Studio Online:** Access to SAS Studio is done through a web browser and is mostly reliable. However, the program is run on SAS's servers and SAS allocates the resources in order for the program to run smoothly. In the past students have experienced outages and, although these are generally temporary, these can cause students to take longer to complete tasks than would normally be anticipated. Around assignment due dates and times can be especially problematic as several people are attempting to get on the server at once and therefore experience more outages. Understanding this, **a temporary server outage is not a valid excuse to turning in an assignment late**. Good advice is to allow yourself plenty of time to complete your assignments. Please start assignments early to avoid the frustration that a server outage can cause. Trying to complete an assignment at the last minute is a formula for creating extreme stress and potentially adversely affecting your grade.

Case Study: Students will work individually on a project-like assignment in which the data is supplied to the student as well as the application scenario. The case study will be based on regression modelling. As with assignments, you are free to discuss the assignment with any of your classmates; however, your programming, interpretation, and write-up must be done independently and an individual's exclusive work.

Projects: There will be two individual projects and one team project during the course of the term. Each of the projects will focus on a different statistical model. Unlike with the case study, students will be responsible for finding or generating data for which it is appropriate to apply the statistical model for that project.

> **Project A** will consist of applying a simple regression model to data found by the student. Students will first share their project idea with a group of their peers via a discussion board, receive feedback, model and analyze their data, and prepare a brief written report.
>
> **Project B** will consist of applying ANOVA and ANCOVA models to data found by the student. Students will model and analyze their data, and write a report outlining their findings.
>
> **Project C** will consist of a logistic regression model applied to data of interest. Students will work in teams of 3 or 4 and be responsible for finding their own data. Each team will create presentation materials and create a video presentation.

In general, each project consists of finding a data set of interest, employing graphical methods for presenting the data, fitting and refining the appropriate statistical model, assessing the adequacy of the model, generating and running appropriate SAS code, and interpreting the results. Each project data set, to the best of your knowledge, should not have been previously analyzed in the way you plan to use it for your project.

Grade Complaints: Address your work in question first to the TA responsible for grading it (may be a different person for assignments and projects). Provide **a clear, brief, written explanation** of why you think you deserve additional credit. The written statement must be provided **within one week** after the work is graded and available to the class in general. All grade disputes must be resolved by the last day of normal classes (before finals week).

Tentative Course Outline:

Week of	Topics	Assignments/Project
Aug. 24	Intro / Stat. Modeling, Simple Regression	First Day Attend. Due 8/26
Aug. 31	Prediction, Transformations, Outliers	
Sep. 7	Regression Inference, Multiple Regression	Assignment #1 due 9/11
Sep. 14	Multicolinearity, Nested Models, Predictor Selection	Project A Idea Posts 9/16 & 9/18
Sep. 21	Subsets, Categorical Coding, Bootstrap	**Project A due 9/25**
Sep. 28	One-Way ANOVA, Post-Hoc Testing	Case Study Due 10/2
Oct. 5	Two-Way ANOVA. Interaction, Eq. Variance Test	
Oct. 12	Comparisons, Contrasts, Reg w/Indicators, ANCOVA	Assignment #2 due 10/16
Oct. 19	Project B Set-Up, Logit Function	
Oct. 26	Binary Log. Reg., Odds, Inference, Empirical Log. Plot	**Project B due 10/30**
Nov. 2	Log. Reg. Test Stat., Multiple Logistic Regression	Assignment #3 due 11/6
Nov. 9	Log. Reg, Model Assessment, ROC Curves	
Nov. 16	Gain/Lift, 2-Way Tables, Categorical Predictors	Assignment #4 due 11/20
Nov. 23	Validating /Scoring Models, Thanksgiving Break	
Nov. 30	Project C Prep., Proj. C Video Presentation	**Project C Video due 12/4**
	Dec. 4: Project C Materials Upload	**Project C Upload due 12/4**

Certificate in *SAS Programming and Data Analysis:*
This course satisfies one of the four courses required for the SAS Programming and Data certificate jointly sponsored by FSU and the SAS Institute. **Students** interested in the program **must apply to the program before** the end of the semester in which **the second course in the program is taken**. In addition, a portfolio is required to be submitted in the last semester of program and a representative assignment and/or project from this course must be included. For more details see http://stat.fsu.edu/sas-certificate.

Liberal Studies for the 21st Century:
The Liberal Studies for the 21st Century Program at Florida State University builds an educational foundation that will enable FSU graduates to thrive both intellectually and materially and to support themselves, their families, and their communities through a broad and critical engagement with the world in which they live and work. Liberal Studies thus offers a transformative experience. This course has been approved as meeting the requirements for Scholarship in Practice and thus is designed to help you become a flexible thinker; a productive member of society; and an independent learner.

Academic Honor System:
The Florida State University Academic Honor Policy outlines the University's expectations for the integrity of students' academic work, the procedures for resolving alleged violations of those expectations, and the rights and responsibilities of students and faculty members throughout the process. Students are responsible for reading the Academic Honor Policy and for living up to their pledge to ". . . be honest and truthful and . . . [to] strive for personal and institutional integrity at Florida State University." (Florida State University Academic Honor Policy, found at http://fda.fsu.edu/Academics/Academic-Honor-Policy.)

Americans with Disabilities Act:
Students with disabilities needing academic accommodation should:
(1) register with and provide documentation to the Student Disability Resource Center; and
(2) bring a letter to the instructor indicating the need for accommodation and what type.

Please note that instructors are not allowed to provide classroom accommodation to a student until appropriate verification from the Student Disability Resource Center has been provided.

This syllabus and other class materials are available in alternative format upon request.

For more information about services available to FSU students with disabilities, contact the:

Student Disability Resource Center
874 Traditions Way
108 Student Services Building
Florida State University
Tallahassee, FL 32306-4167
(850) 644-9566 (voice)
(850) 644-8504 (TDD)
sdrc@admin.fsu.edu
http://www.disabilitycenter.fsu.edu/

Syllabus Change Policy:
Except for changes that substantially affect implementation of the evaluation (grading) statement, this syllabus is a guide for the course and is subject to change with advance notice.