Syllabus

Course Information
Advanced Data Analysis (STAT 448) is a very broad analytics course that aims to cover several topics in statistics and data science and to apply these methodologies in the Statistical Analysis System (SAS). The topics in the course include the following: descriptive statistics, data visualization, hypothesis testing, categorical data analysis, ANOVA, linear regression, logistic regression, generalized linear models, PCA, hierarchical cluster analysis, and discriminant analysis. SAS is an analytics program developed on the campus of North Carolina State University and is in use by several companies.

Section 1
Tuesdays & Thursdays 9:30 am - 10:50 am   Mechanical Engineering Bldg 253

Course Website
compass2g.illinois.edu

Instructor: Christopher Kinson   kinson2@illinois.edu   Illini Hall Room 103A
Instructor Office Hours
Tuesdays & Thursdays 11:00 am - 12:00 pm
Other days by appointment only

TA: Robin Tu   robintu2@illinois.edu   Illini Hall Room 104
TA Office Hours
Wednesdays 9:30 am - 11:00 am
Other days by appointment only

Prerequisites
STAT 400 or STAT 409, and STAT 410 (or currently registered)
Familiarity with SAS is helpful, but not required.

Disability Accommodations
To obtain disability-related academic adjustments and/or auxiliary aids, students with disabilities must contact the course instructor <u>and</u> the Disability Resources and Educational Services (DRES) as soon as possible. To contact DRES, you may visit 1207 S. Oak St., Champaign, call 333-4603, e-mail disability@illinois.edu or go to the DRES website.

Academic Integrity
It is expected that all students abide by the campus regulations on academic integrity [http://studentcode.illinois.edu/article1_part4_1-401.html]. Intentional violations of academic integrity can be found at http://studentcode.illinois.edu/article1_part4_1-402.html and include, but are not limited to, copying any part of another student's homework, allowing another student to copy any part of your homework, or submitting a review or summary of a presentation not attended.

Attendance & Laptop Usage
Students are expected to attend class daily and follow along with the examples presented in

class. Students are strongly encouraged to bring laptops to class. Using the laptops for note taking and computational exercises directly related to this course is expected.

Notes, Code, & Data Sets in Class
Abbreviated lecture notes, select SAS code solutions, and the data sets used in class will be found in Compass. You will need to take notes in class.

Required Textbook
A Handbook of Statistical Analyses using SAS, 3rd Edition by G. Der and B. S. Everitt.

Recommended Textbooks

- Learning SAS by Example: A Programmer's Guide by R. Cody

- SAS Statistics by Example (1st Edition) by R. Cody

Required Software

1. SAS Version 9.4, SAS Institute

    - can be purchased/downloaded from the University Webstore
    - free download of SAS for academic use - SAS University
    - free via the cloud - SAS OnDemand for Academics
        - Enroll in the SAS OnDemand course via this link and type the course code as **67d82b91-1abb-49fa-9090-d5a4fb052359**
    - free to use with UIUC login at the following computer labs:
        (a) Scholarly Commons in the Main Library
        (b) ATLAS in Foreign Language Building Room G8
        (c) CITES in the Undergraduate Library (UGL)
        (d) Academic Computing Facility at the ACES Library

2. Word processing software (e.g., Microsoft Word, Google Docs, Latex)

3. PDF reader

Additional Resources

1. SAS Version 9.4 Online Documentation

2. SAS Procedures By Name

3. SAS E-Learning materials from the University Webstore

Tentative Weekly Schedule

Week	Topic	SCT Days	Items Due
1	Syllabus, How to Write HW Reports, and Ch. 1 Intro to SAS		
2	Ch. 2 Data Description and Simple Inference		HW 1
3	Ch. 3 Simple Inference for Categorical Data Sep Quiz on September 13-15, 2018		Sep Quiz
4	Ch. 4 Analysis of Variance I	SCT	HW 2
5	Ch. 5 Analysis of Variance II & Chs. 6-7 Linear Regression		
6	Chs. 6-7 Linear Regression		HW 3
7	Ch. 8 Logistic Regression Oct Quiz on October 11-13, 2018		Oct Quiz
8	Ch. 8, Ch. 9 Generalized Linear Models	SCT	HW 4
9	Ch. 9 & Ch. 16 Principal Components Analysis		
10	Ch. 16 & Ch. 17 Cluster Analysis		HW 5
11	Ch. 17 & Ch. 18 Discriminant Function Analysis Nov Quiz on November 8-10, 2018		Nov Quiz
12	Ch. 18 & Overview of Repeated Measures and Longitudinal Data	SCT	HW 6
13	Fall Break		
14	Chs. 11-14 Longitudinal Data Analysis		
15	How to Write the Final Project Report		
16	Final Project Report Q & A		
	Final Project Report Due Tuesday December 18, 2018 12:00:00 AM		Upload Final Project Report & SAS Code

Important Dates:
Fall Break - November 17 - 25, 2018 (no class and no office hours on these days)
Reading Day - Thursday, December 13, 2018 (no class on this day)

Grading Breakdown
Participation Assignments: 5%
Homework Assignments: 50%
Quizzes: 15%
Final Project Report and SAS Code: 30%

Participation Assignments
Throughout the semester, students will be tasked with asking questions, answering questions, providing coding examples, making informal presentations, and providing the feedback which are all useful to assess student learning. These assignments will be completed each week, and thus daily attendance is imperative. If you are absent, you will not be able to participate

during class. Participating during class accounts for 5% of your final grade.

Homework Assignments

There are 6 assignments accounting for 50% of your final grade, and the single lowest assignment score will be dropped for undergraduate students. See the **Graduate vs Undergraduate Level Standards** section below for more details. When completing the assignment, read it carefully, and follow the directions part by part. You will have two files (SAS program file and written report file) to turn in for each homework assignment. Save your SAS program file (.sas) and written report file (.pdf) with your name and homework assignment number. The SAS program file should contain your SAS language coding that produces plots, graphs, tables, and other statistical results. The written report file should contain your answers (in words) to the homework questions, which may also require code, plots, graphs, and tables as evidence to support your word answers.

For a student with UIUC NetID doe21 turning in Homework 4, their files would be saved as `HW4_doe21.pdf` and `HW4_doe21.sas` (capitalizing does not matter). Inside of your .sas program file be sure to change all folder references to '`C:\Stat_448`' <u>after</u> running your code. Make sure your homework is neat and readable, containing only relevant results and responses to the questions. You will upload your program file and report file in Compass. **You can submit your program and report file an unlimited number of times prior to the due date. We will grade the most recent submission so long as it is before the deadline and includes both the .pdf and .sas files.** You are expected to complete your homework individually since you are graded as an individual. There is nothing wrong with collaborating with peers to think through the homework problems. Questions about the grading should be directed to the instructor. **Students can dispute their grades or ask for points back up to one week after their original grade has been posted to Compass.** There is no make-up for missed homework.

Late Submissions Policy

Homework and final report assignments submitted up to 2 days after the deadline will be accepted but at an automatic percentage reduction. Your submission will then be graded which may result in additional point deductions due to errors in your solutions. If your homework is submitted:

- $(0, 24]$ hours after the deadline, you will receive a 10% deduction, e.g., your maximum score is 90 out of 100 points

- $(24, 48]$ hours after the deadline, you will receive a 20% deduction, e.g., your maximum score is 80 out of 100 points

- $(48, \infty)$ hours after the deadline, you will receive 100% deduction , e.g., your maximum score is 0 out of 100 points

The exceptions to this policy are outlined in the **Graduate vs Undergraduate Level Standards** section below.

Quizzes

There will be three quizzes each accounting for 5% of the final grade. They will take place

outside of class at a reserved time of your choice within the approved period of quiz days in the College of Engineering Computer Based Testing Facility (CBTF): https://cbtf.engr.illinois.edu. (But first, you should go to https://prairielearn.engr.illinois.edu/ to enroll in our STAT 448 course for the PrairieLearn system. This is the link you need so that you can correctly be added to the list of students allowed to take all of the quizzes.) The CBTF is a computer lab that is open roughly 80 hours each week including evenings and weekends. The computers are Linux based workstations computers with web-browsers to connect to the PrairieLearn servers which is where your quizzes are found. The quizzes are roughly 4 weeks apart. There are 50 minutes allotted for each quiz. See the tentative weekly schedule above for quiz dates and the topics covered.

The policies of the CBTF are the policies of this course, and academic integrity infractions related to the CBTF are infractions in this course. If you have accommodations identified by the DRES for exams, please take your Letter of Accommodation (LOA) to the CBTF proctors in person before you make your first exam reservation. The proctors will advise you as to whether the CBTF provides your accommodations or whether you will need to make other arrangements with your instructor. Any problem with testing in the CBTF **must** be reported to CBTF staff at the time the problem occurs. If you do not inform a proctor of a problem during the test then you **forfeit** all rights to redress. See the https://cbtf.engr.illinois.edu for more information about the testing facility and quiz scheduling.

SCT; Short Course Thursday
Since the quizzes are out-of-class assessments, the University suggests equal time away from class. Hence, on SCTs, class will be cut short to 30 minute sessions, since the quizzes are 50 minutes, and our regular class is 80 minutes. Attendance is still mandatory on SCTs, and the 30 minute session may prove vital to the overall learning experience.

Final Project Proposal and Report
The focus of the final exam, which is a final project consisting of a written report and SAS program, is applying data analysis approaches we have covered during the semester to prepare a single thoroughly cleaned and validated dataset. All students will write a 5-10 page report on the analysis of a data set. The large goal is to get students to showcase their analytical and writing skills in a "white paper" format. By a certain deadline, all students must submit a proposal for the final project, which includes a topic title, question(s) you want to answer, data set with description, and analysis technique(s). The analysis techniques could include any method from ANOVAs to discriminant function analysis. Submitting the proposal allows for feedback and guidance from the instructor about the student's ideas and overall direction. Students are expected to upload their final reports and the associated SAS code by Tuesday December 18, 2018 12:00:00 AM in Compass. The final report will count for 30% of the final grade. The **Late Submissions Policy** applies to the final report. More details about the final project will be announced later in the semester.

Graduate vs Undergraduate Level Standards
Graduate students will have two distinct differences in standards of performance. The first is that graduate students taking this course do not get to drop any homework assignments.

Undergraduate students taking this course will be allowed to drop their single lowest HW grade as mentioned in the **Homework Assignments** section above. The second is that graduate students taking this course are only allowed up to 1 day for the late submissions policy. In other words, graduate students are allowed to submit their assignments $(0, 24]$ hours after the deadline, thus receiving a 10% deduction, e.g., your maximum score is 90 out of 100 points. Undergraduate students are given up to 2 days as mentioned in the **Late Submissions Policy** section above.

Final Grade Percentage & Letter Grade

Lower bound	Upper bound	Letter Grade
0.967	1.000	A+
0.933	0.966	A
0.900	0.932	A-
0.867	0.900	B+
0.833	0.866	B
0.800	0.832	B-
0.767	0.800	C+
0.733	0.766	C
0.700	0.732	C-
0.667	0.700	D+
0.633	0.666	D
0.600	0.632	D-
0.000	0.600	F

Questions You May Have
If it's a general question that all students would benefit from, please ask it in class or post it in the Discussion Board on Compass. If it's more specific to your performance and learning, please send an email to the Instructor. All course emails should be sent from your official illinois.edu email address and contain a subject line which begins with "STAT 448" for quickest response.

For Your Safety
We have been asked by Public Safety (police.illinois.edu/safe) to share the following information in case of weather or security emergencies.
    Emergency Response Recommendations
    Video on Emergency Response

Sexual Misconduct Policy and Reporting
The University of Illinois is committed to combating sexual misconduct. Faculty and staff members are required to report any instances of sexual misconduct to the University's Title IX and Disability Office. In turn, an individual with the Title IX and Disability Office will provide information about rights and options, including accommodations, support services, the campus disciplinary process, and law enforcement options.

A list of the designated University employees who, as counselors, confidential advisors, and medical professionals, do not have this reporting responsibility and can maintain confidentiality, can be found at wecare.illinois.edu/resources/students/#confidential. Other information about resources and reporting is available at wecare.illinois.edu.

The Last Word
The instructor reserves the right to make any changes he considers academically advisable. Any changes will be announced in class and on Compass. It is your responsibility to attend the class and keep track of the proceedings.