# 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 statistical computing software including SAS, R, Python, and Excel. The topics in the course include inference, exploratory data analysis, visualizations, modeling, and machine learning. The expectation is that students will become proficient data analysts using SAS, while being introduced to certain analytical capabilities in R, Python, and Excel. Thus SAS is required software, while the others are recommended. This course aims to push students to investigate problems and employ aspects of the investigative process for data analysis.

**Section 3**
Mondays, Wednesdays & Fridays 3:00 pm - 3:50 pm    Foreign Language Bldg G27

**Course Website**
https://compass2g.illinois.edu/webapps/login/

**Instructor:** Christopher Kinson    kinson2@illinois.edu    Illini Hall Room 103A
**Instructor Office Hours** Tuesdays & Thursdays 4:00 pm - 6:00 pm

**Graders:** Firstname Lastname    netIDs@illinois.edu
**TA Office Hours** 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, Technology, Outside Reading Material, Notes, Code, & Data Sets**
Students are expected to attend class daily. Please follow the Student Codehttps://studentcode.illinois.edu/docs/18.001.FullCodeInside.vf.pdf if you do have absences. Incomplete 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 in order to be successful. Additional

coding guides or reading materials may be provided for learning outside of class. To store your data and save your work either bring removable media (e.g., flash drive) to the class-room or use cloud storage (e.g., Box, Dropbox, Google)

**Recommended Textbooks**

- A Handbook of Statistical Analyses using SAS, 3rd Edition by G. Der and B. S. Everitt
- SAS Statistics by Example by R. Cody
- Python for Data Analysis by W. McKinney
- Using R for Data Analysis and Graphics by J. Maindonald
- Statistical Analysis with Excel For Dummies by R. Schmuller

**Required Software Used in This Course**

1. SAS Version 9.4, SAS Institute
   - can be purchased/downloaded from the University Webstore as "SAS for Windows Teaching and Research License & Media"
   - 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. Presentation software (e.g., Microsoft PowerPoint, Google Slides, Latex)

4. PDF reader

**Other Software Used in This Course**

1. R https://cran.r-project.org/ or RStudio https://www.rstudio.com/

2. Python https://www.python.org/

3. Microsoft Excel

**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 | Items Due |
|------|-------|-----------|
| 1 | Syllabus, How to Write Reports, Intro to Advanced Data Analysis & Intro to SAS | |
| 2 | Asian Dating Data | GR & RLL |
| 3 | Asian Dating Data | GR & RLL |
| 4 | Asian Dating Data | GR & RLL |
| 5 | GP on Analysis of Asian Dating Data | Exam 1 & RLL |
| 6 | Business Loan Data | GR & RLL |
| 7 | Business Loan Data | GR & RLL |
| 8 | Business Loan Data | GR & RLL |
| 9 | GP on Analysis of Business Loan Data | Exam 2 & RLL |
| 10 | Spring Break | |
| 11 | Opioid Overdose Data & Final Project Data Introduction | GR & RLL |
| 12 | Opioid Overdose Data | GR & RLL |
| 13 | Opioid Overdose Data | GR & RLL |
| 14 | GP on Analysis of Opioid Overdose Data | Exam 3 & RLL |
| 15 | Final Project Q & A & Preparation | RLL |
| 16 | Final Project Preparation | |

*Important Dates:*
Spring Break - March 16 - 24, 2019 (no class and no office hours on these days)
Reading Day - Thursday, May 2, 2019 (no class on this day)

**Grading Breakdown**
Reflecting and Learning Logs (RLL): 10%
Group Reports (GR): 40%
Group Presentations (GP): 15%
Exams: 15%
Proposal for Final Project Presentation: 5%
Final Project Presentation: 15%

**Reflecting and Learning Logs (RLL)**
Each week, students will write in the Discussion Board about their experiences as it relates to the course. These logs are intended for you to record your learning strengths and weaknesses, the development of your learning, and strategies for becoming independent learners

who know how to recognize patterns and seek solutions to learning struggles. The logs also serve as a space for you to connect your personal story and experiences to your learning. This might mean exploring your background knowledge, beliefs, and attitudes to bring about self-awareness in your educational journey. Overall, the hope is that you will become empowered and authentic learners. Students will choose 2 sentence starters and complete the sentences with their own thoughts and words. Students should not complete RLLs several weeks in advance. Completing the logs, at the appropriate deadline, will account for 10% of your final grade. More details about the logs will be given in Compass and in class.

## Group Reports (GR)

Students are expected to attend class daily, take notes, and collaborate on group assignments each week during class. Roughly each week, students will be assigned into groups of 4 to answer certain questions and formalize the answers in a report document that includes plots, tables, and corresponding analyses. More details about the format of the report will be discussed in class.

In order for a student to receive credit for the GR, that student must be present and actively participating in class for that group assignment. Group members will receive the same grade for the reports unless a member is absent. Any student who is absent on days when groups are working on the reports will receive a 0 for the report. Follow the Student Code regarding absences. The group reports will be submitted by the end of class (3:50 pm) and graded for analytical thoroughness, grammar, and overall quality. Feedback will be provided via in-person tutorials with either the instructor during office hours, or the graders by appointment. At least one group member, called a "representative", must be present at the feedback tutorials in order for the group to receive points for the revision. If a group's representative cannot make it to the tutorials sessions, then virtual accommodations should be scheduled with the instructor. Once receiving feedback, groups can submit their revised reports (up to one week after original deadline) for half of the credit that was deducted from the original report. Group members who were absent originally, but who come to the tutorials and contribute to the revision, may receive three-fourths of the newly scored group report grade. These group reports account for 40% of your final grade. Late submissions of group reports will be accepted but with a strong penalty; see the **Late Submissions Policy**. There is no make-up for missed group reports. To receive feedback, groups must submit their reports. If a group is satisfied with their original grade and does not wish to receive additional points, then they need not attend feedback tutorials.

## Group Presentations (GP)

There are 3 regular group presentations during the semester. They will take place during class after groups have had enough time to investigate and analyze various datasets. See the tentative weekly schedule above for a rough time frame. Each group is given exactly 8 minutes to present with up to 2 minutes of feedback from the instructor. Each student should have a speaking role during the presentation that clearly indicates their contribution to the group work. If a group member is absent from the GP and cannot make virtual accommodations to present, then that student will receive a 0 for their presentation score. However, the student may still receive credit for the presentation materials that their group contributed. The presentations will be in a slideshow format with more details given in class.

The presentations account for 15% of your final grade. There is no make-up for missed group presentations.

**Exams**
There will be three exams each accounting for 5% of the final grade. They will take place in class and distributed on paper as written exams. The exams are roughly 4 weeks apart. There are 50 minutes allotted for each exam. See the tentative weekly schedule above for exam dates.

**Final Project Proposal and Presentation**
The focus of the final exam (final project) is a group-based oral presentation with a poster. The poster presentation will take place on Thursday May 9, 2018 9:30 am - 11:00 am. Students and faculty will visit your posters and your group will discuss your analysis project and field any questions they have. There will also be judges who will rate your presentations and posters. The group with the highest ratings will receive up to 2 bonus points added to their final grade. The final project's data set will be on natality or the incidence of birth in the US. Your group will carry out an investigation and data analysis based on your own questions and claims. 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, analysis technique(s), data set with description, and software you will use. Software used for the project can be either SAS, R or Python. Submitting the proposal allows for feedback and guidance from the instructor about the group's ideas and overall direction. By a certain deadline, groups will submit their digital posters to the Instructor so that the posters may be printed that the Instructor will pay for. The proposal accounts for 5%, while the presentation accounts for 15% of your final grade. The **Late Submissions Policy** does apply to the proposals. Late submissions of the digital posters will not be accepted. More details about the final project will be announced later in the semester.

**Late Submissions Policy**
Group reports and final project proposals submitted up to 1 day 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 issues in your report. If submitted:

- $(0, 24]$ hours after the deadline, you will receive a 25% deduction, e.g., your maximum score is 75 out of 100 points

**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 | 0.899 | B+ |
| 0.833 | 0.866 | B |
| 0.800 | 0.832 | B- |
| 0.767 | 0.799 | C+ |
| 0.733 | 0.766 | C |
| 0.700 | 0.732 | C- |
| 0.667 | 0.699 | D+ |
| 0.633 | 0.666 | D |
| 0.600 | 0.632 | D- |
| 0.000 | 0.599 | F |

**For Your Safety**

We have been asked by Public Safety (https://police.illinois.edu/emergency-preparedness/run-hide-fight/) 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 https://wecare.illinois.edu/resources/students/#confidential. Other information about resources and reporting is available at https://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.