

# STA 371G: Statistics and Modeling (04730)

## SPRING 2017 SYLLABUS

**Instructor:** Brian Lukoff, Ph.D.

**Email:** [brian.lukoff@utexas.edu](mailto:brian.lukoff@utexas.edu)

**Phone:** (415) 652-8853

**Class Meetings:** MW 3:30-5:00 PM in UTC 4.132

**Office Hours:** MW 5:00-6:00 PM in CBA 3.420

**TA:** Ryan O'Donnell

Will Schievelbein

**Email:** [rodonnell@utexas.edu](mailto:rodonnell@utexas.edu)

[will.schievelbein@utexas.edu](mailto:will.schievelbein@utexas.edu)

**TA:** Hari Prakash

**Email:** [hari.prakash@utexas.edu](mailto:hari.prakash@utexas.edu)

**Office Hours:** M 1:00-3:00 PM, T 5:00-6:00 PM,

W 10:30-11:30 AM & 1:00-2:00 PM, all in CBA 4.304A

## About the course

This course builds on STA 309 and focuses on methods for modeling and analyzing multivariate data. We will explore multiple regression models and their business applications, time-series models, decision analysis and the value of information, and simulation-based methods.

## Course materials

**Textbook:** Rather than a textbook, reading assignments will consist of short lecture notes that you will read before class to prepare for class. You will use Perusall, a social reading platform, to access the reading assignments. There is no cost to use this platform in this class. You can sign up at <http://perusall.com>.

**Student response:** We will be using Learning Catalytics in class to facilitate instant feedback, small group discussion, and other in-class activities. You can purchase access online for \$12 for the semester (there is no cost if you already bought Learning Catalytics for another course this semester). Bring a laptop, smartphone, or tablet to each class—let me know right away if this is not possible for you and we will make other arrangements. You can sign up at <http://learningcatalytics.com>.

**Other materials:** This syllabus, lecture slides, and other materials will be posted on GitHub at <https://github.com/brianlukoff/sta371g>.

## Reading assignments

To make the most efficient use of our time in class together, you'll read the relevant sections of the textbook *before class* using Perusall. Perusall helps you master readings faster, understand the material better, and get more out of the class. To achieve this goal, you will be collaboratively annotating the textbook with others in your class. The help you'll get and provide your classmates (even if you don't know anyone personally) will get you past confusions quickly and will make the process more fun. While you read, you'll receive rapid answers to your questions, help others resolve their questions (which also helps you learn), and advise the instructor how to make class time most productive.

You can start a new annotation thread in Perusall by highlighting text, asking a question, or posting a comment; you can also add a reply or comment to an existing thread. Each thread is like a chat with one or more members of your class, and it happens in real time. Your goals in annotating each reading assignment are to stimulate discussion by posting good questions or comments and to help others by answering their questions.

Research shows that by annotating thoughtfully, you'll learn more and get better grades, so here's what "annotating thoughtfully" means: Effective annotations deeply engage points in the readings, stimulate discussion, offer informative questions or comments, and help others by addressing their questions or confusions. To help you connect with classmates, you can "mention" a classmate in a comment or question to have them notified by email (they'll also see a notification immediately if online), and you'll also be notified when your classmates respond to your questions.

For each assignment we will evaluate the annotations you submit on time. Based on the overall body of your annotations, you will receive a score for each assignment to reflect the effort you have put in.

I will use the questions you pose in Perusall to better focus our class time together: during class, we'll zero in on the material that was most challenging on the reading. This way, we can spend more time in class on what you find challenging and less time on the easy stuff.

## Exams

Exams for this class will be given in the ModLab. You will have access to R during the exam. You may bring one 8.5" by 11" page (both sides) of notes to the first midterm, two pages to the second midterm, and three pages to the final exam. You must bring a picture ID to each test. There will be no make-ups. The final exam will be cumulative and the grade can replace one lower test grade. You must inform me in advance if you are going to miss a test due to observance of a religious holiday or an official university activity.

**Midterm 1:** February 27, 3:30-4:45 PM

**Midterm 2:** April 3, 3:30-4:45 PM

**Final Exam:** May 12, 7-10 PM

## Homework

The homework in this course will likely be significantly different from your first statistics course. Homework problems in STA 371G will usually require more in-depth analysis than problems from STA 309 or another introductory course. For each homework assignment:

- Submit each assignment in Canvas by 11:59 PM the day it is due. You should submit a Word document with the answers to the questions and any R script you created that support the calculations. (This is the electronic version of “showing your work!”)
- Your response to each question should be in paragraph form, using complete sentences. It should read like an argument for why your answer is correct!
- Late homework is not accepted, but the lowest score will be dropped.
- You can work in groups, but each person in your group must write up and submit their results independently.
- Most of the problems on each assignment will be graded for completeness only; to provide you with detailed feedback on your work, we will select a few problems on each assignment to grade for correctness.

## Team projects

There will be one team project to be completed for this class, where you will apply some of the techniques of this course to real-world data of your choosing. Normally, all members of a group will receive the same grade for the project, but group members that do not fully participate may have their grade reduced.

## Grading

Your course grade will be calculated as follows:

5%	Learning Catalytics	Based on your participation on Learning Catalytics questions in class (whether you get the questions correct won't affect your score; answer 75% or more of the questions to get full credit).
5%	Perusall	Based on the thoughtfulness and effort of your work on the reading assignments.
15%	Homework	
15%	Team project	
20%	Midterm 1	
20%	Midterm 2	
20%	Final Exam	If your final exam score is <i>higher</i> than your Midterm 1 or 2 score, I will overwrite the lower midterm grade with your final exam score.

Grades will be assigned to comply with the Guidelines for Grading in McCombs Undergraduate Classes.

## Contacting the instructor

I have office hours after each class (5:00-6:00 PM) in CBA 3.420, and by appointment. Please come to office hours to ask questions, get additional help, or discuss any course-related issues. I will do my best to reply to all email inquiries within 24 hours. Be aware that we cannot discuss specific grades over email (we can only discuss grades in person).

## Computing

The practice of statistics requires extensive numerical calculations. We will use R, a powerful, state-of-the-art statistical software platform, for statistical computing in this course. R is frequently used by practitioners and being able to use R to do statistical analysis is a valuable (and marketable!) skill. No prior knowledge of R is needed; we'll teach you all you need to know.

You may find it helpful to have a laptop in class to use R for data analysis, participate using Learning Catalytics, and/or follow along on class notes. If you don't bring a laptop to class, you will need to bring some other device that has a calculator and can access the web (e.g., a smartphone or tablet).

## Students with Disabilities

Students with disabilities may request appropriate academic accommodations from the Division of Diversity and Community Engagement, Services for Students with Disabilities, 512-471-6259, <http://www.utexas.edu/diversity/ddce/ssd/>.

## Religious Holy Days

By UT Austin policy, you must notify me of your pending absence at least fourteen days prior to the date of observance of a religious holy day. If you must miss a class, an examination, a work assignment, or a project in order to observe a religious holy day, you will be given an opportunity to complete the missed work within a reasonable time after the absence.

## Policy on Scholastic Dishonesty

The McCombs School of Business has no tolerance for acts of scholastic dishonesty. The responsibilities of both students and faculty with regard to scholastic dishonesty are described in detail in the BBA Program's Statement on Scholastic Dishonesty at <http://www.mcombs.utexas.edu/BBA/Code-of-Ethics.aspx>. By teaching this course, I have agreed to observe all faculty responsibilities described in that document. By enrolling in this class, you have agreed to observe all student responsibilities described in that document. If the application of the Statement on Scholastic Dishonesty to this class or its assignments is unclear in any way, it is your responsibility to ask me for clarification. Students who violate University rules on scholastic dishonesty are subject to disciplinary penalties, including the possibility of failure in the course and/or dismissal from the University. Since dishonesty harms the individual, all students, the integrity of the University, and the value of our academic brand, policies on scholastic dishonesty will be strictly enforced. You should refer to the Student Judicial Services website at <http://deanofstudents.utexas.edu/sjs/> to access the official University policies and procedures on scholastic dishonesty as well as further elaboration on what constitutes scholastic dishonesty.

Examples of scholastic dishonesty in this course include copying or collaborating during assessments, discussing or divulging the contents of a quiz or exam with another student who will take the test, use of assignment solutions from another student or semester, and attempting to gain credit for course participation or quizzes while not actually in class.

## Campus Safety

Please note the following recommendations regarding emergency evacuation from the Office of Campus Safety and Security, 512-471-5767, <http://www.utexas.edu/safety>:

- Occupants of buildings on The University of Texas at Austin campus are required to evacuate buildings when a fire alarm is activated. Alarm activation or announcement requires exiting and assembling outside.
- Familiarize yourself with all exit doors of each classroom and building you may occupy. Remember that the nearest exit door may not be the one you used when entering the building.
- Students requiring assistance in evacuation should inform the instructor in writing during the first week of class.

- In the event of an evacuation, follow the instruction of faculty or class instructors.
- Do not re-enter a building unless given instructions by the following: Austin Fire Department, The University of Texas at Austin Police Department, or Fire Prevention Services office.
- Behavior Concerns Advice Line (BCAL): 512-232-5050
- Further information regarding emergency evacuation routes and emergency procedures can be found at: <http://www.utexas.edu/emergency>.

## Course schedule

This schedule is subject to change as necessary.

MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY
Jan 16	17	18 <b>Lecture 1</b> Introduction to analytics	19	20
23 <b>Lecture 2</b> Review of distributions and estimation	24	25 <b>Lecture 3</b>	26	27
30 <b>Lecture 4</b> Simple regression	31	Feb 1 <b>Lecture 5</b>	2	3
6 <b>Lecture 6</b> Multiple regression	7	8 <b>Lecture 7</b>	9	10
13 <b>Lecture 8</b> Dummy variables	14	15 <b>Lecture 9</b> Interactions	16	17
20 <b>Lecture 10</b>	21	22 <b>Lecture 11</b> Diagnostics and transformations	23	24
27 Midterm 1	28	Mar 1 <b>Lecture 12</b>	2	3
6 <b>Lecture 13</b> Model selection	7	8 <b>Lecture 14</b> Logistic regression	9	10
13 No class Spring break	14	15 No class Spring break	16	17

MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY
20 <b>Lecture 15</b>	21	22 <b>Lecture 16</b> Autoregression	23	24
27 <b>Lecture 17</b> Moving averages and smoothing	28	29 <b>Lecture 18</b> Decision trees	30	31
Apr 3 Midterm 2	4	5 <b>Lecture 19</b>	6	7
10 <b>Lecture 20</b>	11	12 <b>Lecture 21</b>	13	14
17 <b>Lecture 22</b>	18	19 <b>Lecture 23</b> Simulation	20	21
24 <b>Lecture 24</b>	25	26 <b>Lecture 25</b>	27	28
May 1 <b>Lecture 26</b>	2	3 <b>Lecture 27</b>	4	5
8	9	10	11	12 Final exam (7-10 PM)