

BUAD 312 Statistics and Data Science for Business COURSE SYLLABUS

Basic information

Instructors: Jacob Bien & Patrick Vossler, Data Sciences and Operations Lecture: TuTh 12:00pm-1:50pm PT (14924) & TuTh 4:00-5:50pm (14926) ¹ Units: 4 units Office Hours: Prof. Bien/Patrick Mon 1:30pm-2:30pm PT & Wed 5:30pm-6:30pm PT or by appointment. Zoom Room Links Email: buad312@marshall.usc.edu Website: BUAD 312's site on blackboard.usc.edu

"When every business has free and ubiquitous data, the ability to understand it and extract value from it... leads to intelligence, and the intelligent business is the successful business, regardless of its size. Data is the sword of the 21st century, those who wield it well, the Samurai."

—Jonathan Rosenberg, advisor to Google founder Larry Page.

Course description

Harnessing the power of data has become an essential skill to be successful in business. This course is designed to give you the tools for "wielding" data effectively. Data in its raw form is usually just gigabytes of numbers and letters. How do we go from piles of numbers to **useful information** that can inform our **business decision making**? How do we communicate these findings and critically evaluate the claims made by our business associates? This is a faster-moving, more advanced version of BUAD 310. The prerequisites are the same (nothing more than high school algebra), but the pace and expectations will be higher. Doing well in this class will **strategically position you in a job market that is desperate to hire business minds that can think and speak in data**.

¹Required to attend synchronously unless you are in a time zone where this falls outside of the 7am– 10pm range in local time.

Prerequisites

Just like BUAD 310, no mathematical background will be assumed other than high school algebra. No prior experience with programming is needed. It is expected that students enrolling in this class understand that this will be a greater challenge. They are excited to take this on because they understand that a greater challenge comes with a greater payoff.

How is this different from BUAD 310?

This is a more advanced alternative to BUAD 310 for students excited to learn the state-of-the-art tools for data science. Both classes teach students the fundamentals of statistics. BUAD 312 emphasizes the powerful combination of computation and statistics. Instead of Excel, students will learn basic programming in the R language. The R programming language has become one of the most common tools for data scientists working in business. As much as possible, BUAD 312 will put contemporary data sets in the spotlight. These will stimulate curiosity and motivate us to learn new techniques for making sense of it all.

Course Learning Objectives

In this class, you will develop the ability to...

- 1. Produce effective visualizations and data summaries, transforming raw data files into human understandable pieces of information.
- 2. Use statistical models to generate predictions about unseen data.
- Use statistical models to make inferences about the underlying processes generating the data.
- 4. Identify the proper statistical tool given a business question and a data set.
- Provide simple yet precise explanations of the purpose of the most common statistical tools.
- 6. Interpret the results of a statistical analysis as a basis for decision making.
- 7. Collaborate effectively to use statistical analysis to address business challenges.
- 8. Communicate your statistical analyses logically and persuasively
- 9. Use R along the entire statistical pipeline—from data import and wrangling, to data exploration and visualization, to analysis and the preparation of reproducible reports.
- 10. Understand the ethical guidelines for statistical practice.

See Appendix A for how these course-level objectives support the six Marshall undergraduate program learning goals.

This course is designed to fulfill the **GE-F (Quantitative Reasoning)** core literacy requirement, with an emphasis on empirical analysis. In particular,² we will (1) "understand the use and meaning of summary statistics of data sets," (2) "learn how to apply basic axioms of probability to analysis in environments characterized by uncertainty," (3) "understand the distinction between causation and correlation," and (4) "learn how to use methods of statistical inference to answer questions about natural, social, or artificial processes, systems, or phenomena, as well as to evaluate the success of a formal model in characterizing a given process, system, or phenomenon." Appendix B provides a mapping of the ten course learning objectives to the learning objectives of the GE-F (Quantitative Reasoning) core literacy category.

Required Materials

All course materials are available online:

- 1. *Textbook:* Statistical Inference via Data Science: A moderndive into R & the tidyverse by Chester Ismay and Albert Y. Kim.
- 2. *Software:* R and RStudio
- 3. Online videos: DataCamp see Problem Set 1 for how to get this for free.
- 4. *Online Platform:* Assignments, announcements, and other course materials will be distributed through Blackboard.

Total cost of materials: \$0.

Course expectations: This is a longer section than in a usual semester!

We expect you to be logged into Zoom on time for every class period, ready to engage actively in the course material. We also expect you to have a paper notebook (remember those?!) and pens/pencils for taking notes. If you prefer writing with a tablet, that is acceptable, but what's important is that you aren't just typing your notes in a text editor. We say this because we will be drawing pictures/diagrams and using mathematical notation. You'll see that we practice what we preach: when we are lecturing, we mostly write things out by hand. In person we would have done this on the whiteboard; online, we will do it with an iPad on screenshare. We do this because we believe it to be more effective for learning and clear thinking than if we were flipping through a slide deck during class. The act of writing notes is, we believe, an important part of the learning process. This is best said by the memorable tagline of a notebook company: "I'm not writing it down to remember it later, I'm writing it down to remember it now."

²Quoting from "Requirements of the USC General Education Program" GE-F.

The one exception to this rule is when we work in R. In that case, your "notes" will be a .R file, which we'll explain how to create when the time comes.

Zoom Camera Policy: Here is the Provost's statement of the in-class camera policy:

"Acknowledging that class dynamics are substantially compromised without the ability to see the people in class, faculty can set an expectation that students have their cameras on during synchronous online sessions. However, some students may be facing challenging situations, such as internet connectivity, illness, or home environments that make this difficult or impossible. To alleviate these concerns, faculty can encourage students to use virtual backgrounds, which will eliminate most privacy concerns, and earphones or headsets to improve audio quality. While faculty are at liberty to create a 'camera-on' policy in their class, they should communicate that accommodations are available to students who contact them directly with reasonable requests."

We are aware that we are all doing our best to adapt during these challenging times and that the challenges you may be facing associated with having your video on and connecting remotely to class may be different from those of other students. We will be happy to accommodate reasonable requests, but we would just ask that you please email me in advance about this so we are on the same page.

If you haven't reached out to us in advance, we will expect your video to be on for the full duration of class time. We are going to be working hard to make our class time engaging and worthwhile and to make the learning environment as close to the inperson experience as possible. Having videos on is one key component to accomplishing this goal.

If our class time does not fall within 7:00am to 10:00pm in your local time: Again, quoting the Provost:

"Remote learning, paired with the fact that USC students are spread across time zones, presents a number of challenges related to attendance and participation in synchronous class sessions. In general, students should plan to attend every synchronous session for the classes in which they are enrolled, irrespective of when it occurs in their time zone. However, faculty should only maintain normal attendance, participation, and assessment expectations for students when the class time falls within reasonable learning hours in the student's time zone, defined as 7:00am to 10:00pm in the student's time zone."

If our class time does not fall within 7:00am to 10:00pm in your local time, we do not expect you to attend the synchronous class sessions. Instead, we will require you to do the following:

- For each synchronous class session:
 - Watch the recording of the class.
 - Submit your handwritten notes (and .R file, if applicable) that you created while watching the synchronous class session. This must be submitted within 24 hours of the recording being made available.

Communicating with us

For each week of the semester, the instructor holding office hours will depend on who has taught the topics most recently covered in lecture. We have created a website that states which instructor is holding office hours during the current week and the expected office hours schedule for the entire semester. Furthermore, this website contains links to both of our Zoom rooms to make it easy to join our office hours.

Grading

Problem Sets	7%
DataCamp Assignments	2%
Engage-Integrate	11%
Quizzes	30%
Midterm Exam	25%
Final Exam	25%

Components of the grade

Problem Sets: There will be 7 problem sets assigned through the semester. Notice that each is worth only 1% of the grade. Thus, these should be thought of as a low-stakes opportunity to actively engage with the material on your own. Regular practice is especially important when it comes to learning to code. You will submit each assignment via blackboard in the format of an R markdown (.Rmd) file (click here for a one minute video introducing it). This is a best-practice format used by data scientists for sharing written reports that include reproducible data analyses. Despite only being worth a small fraction of the overall course grade, problem sets and reading are probably the **most valuable** thing for you to invest time and energy into doing well (along with active engagement in synchronous/asynchronous class time, of course). If you take the right approach to class time, problem sets, and reading"), you will find yourself ready for quizzes (worth 30%) and the exams (worth a total of 50%).

All problem sets must be submitted. Late problem sets will receive a 0. However, we will drop the lowest two problem set scores.

DataCamp Assignments: These are interactive online segments that we will be assigning throughout the semester. A video will introduce you to a skill in R and then you will immediately write code to practice that skill in an interactive environment. Your score will be based on the accuracy of your answers and on how many hints you required to get to the right answer. Note: You will get free access to DataCamp.

Engage-Integrate: The "Engage-Integrate" component will be earned primarily through completing "micro-assignments." The micro-assignments will be more varied, sporadic, or open-ended than homework assignments. Examples of micro-assignments we might assign: read a statistics-related news article or watch an online video and then write a two-sentence response; apply something we've learned to a real-world context; find a data set you find interesting and analyze it. The hope is that micro-assignments will be a

fun way to interact with the course material in a less structured way that will encourage you to connect what we are learning to your life and other interests. Micro-assignments are different in nature from homework. For example, if you don't do a micro-assignment, it will not have negative consequences on your ability to do well on quizzes/exams. The number of micro-assignments varies each semester, but you can expect roughly 15–20.

Quizzes: There will be five quizzes (see schedule), each worth 6% of the grade. No collaboration/communication is allowed. These will be taken on your computer through blackboard using Proctorio. If you take the right approach to class time, problem sets and reading, (see section called "The right approach to class time, problem sets, and reading") then you should find these quizzes straightforward. And if you do not do well, this is a valuable low-stakes, early-indicator that you should consider changing how you are approaching the course work (and we would encourage you to seek out my help in office hours).

Midterm and Final Exams: Same rules as for the quizzes. we will hold review sessions before both exams.

The right approach to class time, problem sets, and reading

In class: Please see course expectations section for the handwritten notes policy.

After class: At some point later in the day, take some time to review your notes from class. Make a note of anything that wasn't clear to you and consider coming to office hours or sending an email to the course email with your question.

Problem sets, reading, and office hours: The schedule has a column called "Readings" that shows the section of the book you are responsible to have read by a certain date. For example, by class 2, you should have read Chapter 1.3 of the textbook.

Before you start a problem set, you should first have done all of the reading. Just as with your after-class review of class notes, make a note of anything you are confused about and bring this question to office hours or to the online forum. You should start by working on each problem on your own, without anyone else. It is quite likely that while working on a problem you will need to look back at your class notes or re-read a part of the textbook. This is a valuable part of the learning process.

Then (and only after this first step), when you've identified problems that are difficult or require discussion, you may talk about the *general approach* with others in the class (or a teaching staff member in office hours). Then, after this discussion you can independently go back to writing your homework. This is very different from doing your homework side-by-side with a classmate. This policy is not just to help us evaluate people separately. **Doing homework carefully on your own, which includes spending time being stumped on a problem, looking back at the book and your class notes to clarify misunderstandings, is the best way to make sure you are actually internalizing the course material in a deep and lasting way. If you don't let yourself be stuck on a problem (before getting help), it's hard to notice areas of material that you actually should be studying more closely. If you find that you have no idea how to do any of the problems on** your own, that is a good indication that you should be approaching this class differently and that you should come to office hours for guidance in how to go about learning the material in this class more effectively. (Better to realize this while doing a problem set than during an exam!)

Further Advice: Office hours, when used well by students, can be an excellent resource. There's a common misconception that office hours are mainly for getting homework help. Given that the material covered in class is challenging, it is important to make sure you understand the material covered in class and in the book very well. If you identify material that you are confused about, coming to office hours is an excellent way to hone your understanding.

<u>Course Schedule</u> This schedule an approximation and may change.

Weekday	Date	Class	Торіс	Readings	Problem Sets	DataCamp	Quizzes
T Th T Th T	2021-01-19 2021-01-21 2021-01-26 2021-01-28 2021-02-02	1 2 3 4 5	Intro lecture; syllabus R RStudio, packages, RMarkdown effective visualization of data and the grammar of graphics	1-1.3 1.4-1.5, 2-2.3 2.4-2.9 3-3.3	PS1 due	Tidyverse 1-2 Tidyverse 3-4	
Th T Th T Th	2021-02-04 2021-02-09 2021-02-11 2021-02-16 2021-02-18	6 7 8 9 10	Data wrangling , summary statistics, and life within the tidyverse Correlation & regression Regression with 1 categorical feature	3.4-3.6, 3.8.1, 3.9 5-5.1.1, 5.3.1 5.1.2, 5.3.2, 5.1.3, 5.3.3 5.2	PS2 due PS3 due	Modeling 1 Modeling 2	Quiz 1
T Th T Th T	2021-02-23 2021-02-25 2021-03-02 2021-03-04 2021-03-09	11 12 13 14 15	Multiple regression Interactions vs parallel slopes Sampling framework; midterm review Sampling : the good, bad, and ugly Midterm	6.2-6.2.3 6-6.1 7-7.2.3 7.2.4-7.3.2	PS4 due		Quiz 2
Th T Th T Th	2021-03-11 2021-03-16 2021-03-18 2021-03-23 2021-03-25	16 17 18 19 20	Sampling, probability, and simulations Resampling and the bootstrap Confidence intervals: their goal and Wellness Day: no class how to construct and interpret	7.3.3-7.5 8-8.2 8.3 8.4-8.7	PS5 due		Quiz 3
T Th T Th T	2021-03-30 2021-04-01 2021-04-06 2021-04-08 2021-04-13	21 22 23 24 25	Simulation Hypothesis testing: the intuition, the definitions, and interpretation Inference for regression	9-9.2 9.3 9.4-9.5 10-10.2	PS6 due		Quiz 4
Th T Th T Th	2021-04-15 2021-04-20 2021-04-22 2021-04-27 2021-04-29	26 27 28 29 30	regression tables + conditions for inference Effective data storytelling Wellness Day: no class Data ethics : Privacy, fairness, + p-hacking Final review	10.3-10.4	PS7 due		Quiz 5

Please note: The final exam is **Wednesday, May 5, 8am–10am**. There is no make-up exam time, so please make sure this time works for you before enrolling in this class. This time is as listed here under Exceptions: https://classes.usc.edu/term-20211/finals/

Collaboration Policy

Discussion of problem sets is permitted. However, please note that "discussion" means talking with a classmate or teaching staff about the *general approach* to a problem. When you actually write up your assignment (including any code), you must do so on

your own. At the top of each problem set you must write down the names of any people (including the teaching staff) you had discussions with for that assignment.

Collaboration of any sort on quizzes and exams is prohibited and will result in a 0 on that exam/quiz in addition to further penalties—we take violations of this rule very seriously and potential cheating issues will be brought to the administration.

Add/Drop Process

The last day to register and add classes is Februrary 5. For more information, such as the deadline to drop a class without a mark of "W" on your transcript, visit https://classes.usc.edu/term-20211/calendar/.

Academic Integrity and Conduct

Plagiarism—presenting someone else's ideas as your own, either verbatim or recast in your own words—is a serious academic offense with serious consequences. Please familiarize yourself with the discussion of plagiarism in SCampus in Part B, Section 11, "Behavior Violating University Standards" https://policy.usc.edu/scampus-part-b/. Other forms of academic dishonesty are equally unacceptable. See additional information in SCampus and university policies on scientific misconduct, http://policy.usc. edu/scientific-misconduct.

Discrimination, sexual assault, and harassment are not tolerated by the university. You are encouraged to report any incidents to the Office of Equity and Diversity http: //equity.usc.edu or to the Department of Public Safety http://adminopsnet.usc.edu/ department/department-public-safety. This is important for the safety of the whole USC community. Another member of the university community—such as a friend, classmate, advisor, or faculty member – can help initiate the report, or can initiate the report on behalf of another person. The Center for Women and Men http://www.usc. edu/student-affairs/cwm/ provides 24/7 confidential support, and the sexual assault resource center webpage http://sarc.usc.edu describes reporting options and other resources.

Students with Disabilities

USC is committed to making reasonable accommodations to assist individuals with disabilities in reaching their academic potential. If you have a disability which may impact your performance, attendance, or grades in this course and requires accommodations, you must first register with the Office of Disability Services and Programs (www.usc.edu/disability). DSP provides certification for students with disabilities and helps arrange the relevant accommodations. Any student requesting academic accommodations based on a disability is required to register with Disability Services and Programs (DSP) each semester. A letter of verification for approved accommodations can be obtained from DSP. Please be sure the letter is delivered to me as early in the semester as possible. DSP is located in GFS (Grace Ford Salvatori Hall) 120 and is open 8:30 a.m.–5:00 p.m., Monday through Friday. The phone number for DSP is (213) 740-0776. Email: ability@usc.edu.

Class Notes Policy

All classes will be automatically recorded using Zoom. These recordings are solely for use of students enrolled in BUAD 312 this semester and may not be shared in any form. In particular, they are considered "educational records" subject to federal privacy laws (Family Educational Rights and Privacy Act - FERPA).

Furthermore, notes made by students based on a university class or lecture may only be made for purposes of individual or group study, or for other non-commercial purposes that reasonably arise from the student's membership in the class or attendance at the university. This restriction also applies to any information distributed, disseminated or in any way displayed for use in relationship to the class, whether obtained in class, via email or otherwise on the Internet, or via any other medium. Actions in violation of this policy constitute a violation of the Student Conduct Code, and may subject an individual or entity to university discipline and/or legal proceedings.

Support Systems

Counseling and Mental Health (213) 740-9355 – 24/7 on call

studenthealth.usc.edu/counseling

Free and confidential mental health treatment for students, including short-term psychotherapy, group counseling, stress fitness workshops, and crisis intervention.

National Suicide Prevention Lifeline 1 (800) 273-8255 – 24/7 on call

 ${\tt suicide prevention lifeline.org}$

Free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week.

Relationship and Sexual Violence Prevention and Services (RSVP) (213) 740-9355(WELL), press "0" after hours – 24/7 on call

studenthealth.usc.edu/sexual-assault

Free and confidential therapy services, workshops, and training for situations related to genderbased harm.

Campus Support & Intervention (CSI) (213) 740-0411

https://campussupport.usc.edu/

A team of professionals here to assist students, faculty, and staff in navigating complex issues. Whether you are here seeking support for yourself or someone else, we are available to help you problem solve, understand options, and connect with resources. Please note that we are not an emergency resource and are not available 24/7.

Office of Equity and Diversity (OED) (213) 740-5086 — Title IX – (213) 821-8298

equity.usc.edu,titleix.usc.edu

Information about how to get help or help someone affected by harassment or discrimination, rights of protected classes, reporting options, and additional resources for students, faculty, staff, visitors, and applicants. The university prohibits discrimination or harassment based on the following protected characteristics: race, color, national origin, ancestry, religion, sex, gender, gender identity, gender expression, sexual orientation, age, physical disability, medical condition, mental disability, marital status, pregnancy, veteran status, genetic information, and any other characteristic which may be specified in applicable laws and governmental regulations. The university also prohibits sexual assault, non-consensual sexual contact, sexual misconduct, intimate partner violence, stalking, malicious dissuasion, retaliation, and violation of interim measures.

Reporting Incidents of Bias or Harassment (213) 740-5086 or (213) 821-8298

usc-advocate.symplicity.com/care_report

Avenue to report incidents of bias, hate crimes, and microaggressions to the Office of Equity and Diversity —Title IX for appropriate investigation, supportive measures, and response.

The Office of Disability Services and Programs (213) 740-0776

dsp.usc.edu

Support and accommodations for students with disabilities. Services include assistance in providing readers/notetakers/interpreters, special accommodations for test taking needs, assistance with architectural barriers, assistive technology, and support for individual needs.

USC Support and Advocacy (213) 821-4710

uscsa.usc.edu

Assists students and families in resolving complex personal, financial, and academic issues adversely affecting their success as a student.

Diversity at USC (213) 740-2101

diversity.usc.edu

Information on events, programs and training, the Provost's Diversity and Inclusion Council, Diversity Liaisons for each academic school, chronology, participation, and various resources for students.

USC Emergency UPC: (213) 740-4321, HSC: (323) 442-1000 - 24/7 on call

dps.usc.edu,emergency.usc.edu

Emergency assistance and avenue to report a crime. Latest updates regarding safety, including ways in which instruction will be continued if an officially declared emergency makes travel to campus infeasible.

USC Department of Public Safety UPC: (213) 740-6000, HSC: (323) 442-120 - 24/7 on call

dps.usc.edu

Non-emergency assistance or information.

Office of the Ombuds (213) 821-9556 (UPC) / (323-442-0382 (HSC)

ombuds.usc.edu

A safe and confidential place to share your USC-related issues with a University Ombuds who will work with you to explore options or paths to manage your concern.

Emergency Preparedness/Course Continuity

In case of a declared emergency if travel to campus is not feasible, USC executive leadership will announce an electronic way for instructors to teach students in their residence halls or homes using a combination of Blackboard, teleconferencing, and other technologies.

A BUAD 312 and Marshall's Six Undergraduate Program Learning Goals

 Marshall program learning goal Our graduates will demonstrate critical thinking skills so as to become future-oriented decision makers, problem solvers and innovators. Specifically, students will: Students will understand the concepts of critical thinking, entrepreneurial thinking and creative thinking as drivers of innovative ideas (not explicit for this course). Critically analyze concepts, theories and processes by stating them in their own words, understanding key components, identifying assumptions, indicating how they are similar to and different from others and translating them to the real world. Be effective at gathering, storing, and using qualitative and quantitative data and at using analytical tools and frameworks to understand and solve business problems. Demonstrate the ability to anticipate, identify and solve business problems. They will be able to identify and assess central problems, identify and evaluate potential solutions, and translate a chosen solution to an implementation plan that considers future contingencies. 	Course objectives that support this goal 1–6,8
 Our graduates will develop people and leadership skills to promote their effectiveness as business managers and leaders in the 21st century's evolving work and organizational structures. Specifically, students will: Students will recognize, understand and analyze the roles, responsibilities and behaviors of effective managers and leaders in diverse business contexts e.g., functionally diverse, culturally diverse, geographically diverse, etc. Students will understand factors that contribute to effective teamwork including how to elicit, manage and leverage diverse perspectives and competencies. Students will recognize, understand, and analyze the motivations and behaviors of stakeholders inside and outside organizations (e.g., teams, departments, consumers, investors, auditors) 	N/A
 Our graduates will be effective communicators to facilitate information flow in organizational, social, and intercultural contexts. Specifically, students will: Identify and assess diverse personal and organizational communication goals and audience information needs. Understand individual and group communications patterns and dynamics in organizations and other professional contexts. Demonstrate an ability to gather and disseminate information and communicate it clearly, logically, and persuasively in professional contexts. 	1,5,6,7,8
 4. Our graduates will demonstrate ethical reasoning skills, understand social, civic, and professional responsibilities and aspire to add value to society. Specifically, students will: 4.1 Understand professional codes of conduct. 4.2 Recognize ethical challenges in business situations and assess appropriate courses of action. 	10
 Our graduates will develop a global business perspective. They will understand how local, regional, and international markets, and economic, social and cultural issues impact business decisions so as to anticipate new opportunities in any marketplace. Specifically, students will: Understand how local, regional and global markets interact and are impacted by economic, social and cultural factors. Understand that stakeholders, stakeholder interests, business environments (legal, regulatory, competitor) and business practices vary across regions of the world. 	N/A
 Our graduates will understand types of markets and key business areas and their interaction to effectively manage different types of enterprises. Specifically, students will: Demonstrate foundational knowledge of core business disciplines, including business analytics and business economics. Understand the interrelationships between functional areas of business so as to develop a general perspective on business management. Apply theories, models, and frameworks to analyze relevant markets (e.g. product, capital, commodity, factor and labor markets). Show the ability to utilize technologies (e.g., spreadsheets, databases, software) relevant to contemporary business practices. 	1–10

B BUAD 312 and GE-F category

This syllabus meets the learning objectives of critical thinking, logical integrity, and application as defined in the USC General Education Program requirement document of 2017. Below is a mapping showing how this course's learning objectives fulfill the learning objectives of the GE-F (Quantitative Reasoning) core literacy category.

Quantitative Reasoning Learning Objectives ³	Course objectives that support this goal
1. Critical Thinking: Students will learn how to use a set of formal tools (e.g., logical or statistical inference,	2,3,4,6
probability, or mathematical analysis) to pose and evaluate hypotheses, claims, questions, or problems within a given (formal) mode of thought.	
2. Logical Integrity: Students will be able to understand the logical structure of a given formal system, to distinguish between its assumptions and implications.	3,5,6
3. Application: Students will be able to identify useful and specific applications of the formal systems they study.	4,6,7