

36-726: Statistical Practice Spring 2021

MW 9:00 -- 10:20 am Pittsburgh Time

(Note Time Change)

Zoom Meeting ID¹: 954 6049 2939 / Passcode: 042583

Class Materials: <https://canvas.cmu.edu/>

“It’s not necessary to know everything about statistics. It is better to understand the problem and use what ‘tools’ you have, than to know every advanced procedure and not quite understand the problem.”

– Fred Ruland, *Guide for the New Statistical Consultant*

“Data are just summaries of thousands of stories—tell a few of those stories to help make the data meaningful.” – Chip and Dan Heath, authors of *Made to Stick* and *Switch*

Instructors

Brian Junker (lead)

stat.cmu.edu/people/faculty/brian

brian@stat.cmu.edu

Zach Branson

stat.cmu.edu/people/faculty/zach-branson

zach@stat.cmu.edu

Valerie Ventura

stat.cmu.edu/people/faculty/valerie-ventura

vventura@andrew.cmu.edu

Brian Macdonald

stat.cmu.edu/people/faculty/brian-macdonald-0

brianmac@andrew.cmu.edu

Office Hours

Junker: Zoom² @ 923 9360 1531

Passcode: 491610

10:00–11:00am Thu, Pittsburgh Time

(or by appointment).

Others: TBA.

Prerequisites

You must be an MSP student to take this class. Beyond that, there are no formal prerequisites for this class. You should be familiar with IMRAD and IDMRAD paper formats, masters’ level applied linear models and statistical computing, and basic professional skills, all as taught in the first semester of the MSP program. You should be proficient with R/RStudio, and able to use or learn other statistical software that is available from the department or open-source. You should also be able to use or learn L^AT_EX or Microsoft Office, Zoom, and other communication tools as needed.

¹Class web link: <https://cmu.zoom.us/j/95460492939?pwd=dFlYbnBsRDdFVXF5MUh4TzIibWlnZ09>

²BJ Office hours link: <https://cmu.zoom.us/j/92393601531?pwd=TE50aDZKRFA5MmFpRnBQZjNUditSUT09>

Course Overview

Most professional statisticians in industry, government, and academia spend a fair bit of their time providing advice and analyzing data for colleagues whose specialties are in some other field. Often, the contributions of the statistician in a scientific collaboration is to sharpen the focus of an investigation by identifying what may, and may not, be learned from a particular data set. It is part of the appeal of our discipline that it is application oriented, yet in most coursework there is inadequate time to take the applications seriously. **The main purpose of this course is to help students develop skills in interacting with a client and in digging more deeply into problems that involve statistical practice.** This course is intended to help you make the transition from being a student of statistics to being a professional statistician.

Communication skills are essential for success. Through practice and discussion, this class will help you to improve your ability to interact with a client. A practicing statistician must be able to ask good questions, understand the context of the client's problem, summarize results and advise the client in a clear and useful manner, both verbally and in writing. Work on projects in this course will help you to improve these skills. Often, a lasting benefit to the client will be the education he or she receives from the statistician. This is another aspect of the process that improves with experience, and which you should be conscious of in your work here.

Texts and Course Materials

This course is very much a practical apprenticeship for you. There are no perfect textbooks for statistical consulting; it is best learned by practice. Below we have listed a few resources for those who would like a textbook close by for reference. The remaining instructional materials for this course will be available on-line, on Canvas. Materials for your consulting project will be available from the client you are working with.

Helpful Texts

Ruland, F. (2014). *Guide for the New Statistical Consultant: Some Suggestions and Three Key Questions to Ask*. Seattle: Amazon.com. Available in Kindle (free), or paperback (\$5.99).

The author has a long career in statistical consulting in commercial and university settings. This little booklet (26 pages!) contains a great deal of wisdom about consulting, which you will hear echoed in our classes together.

Glasman-Deal, H. (2021). *Science Research Writing for Native and Non-Native Speakers of English, Second Edition*. London: World Scientific Publishing. Available at amazon.com (\$25.00).

A detailed guide to writing empirical papers in science, focusing on IMRAD and related formats. Contains analysis of writing samples, suggested vocabulary for scientific papers, etc.

Cabrera, J. & McDougall, A. (2002). *Statistical Consulting*. NY: Springer. Available at amazon.com in hardcover (\$70.75) or paperback (\$99.99), or from link.springer.com (free, if you access through the CMU campus or VPN networks).

A detailed guide to statistical consulting practice, communication and methodology, with several detailed case studies. Worth a skim, but you will probably learn more by doing your own consulting in this class.

Course Organization and Objectives

Prior to the start of the semester, you've ranked projects and Jamie McGovern has matched you as teams to projects. For everyone's benefit, in the first several class periods, all of our clients will give short presentations of their projects, with time for Q&A for you after each presentation.

Each project is assigned one of the course instructors as a mentor. You will work for most of the semester with the client and the faculty mentor to develop solutions to the client's problem in the framework of guided practice. Typically, a project involves:

1. Learning to work with the client, and fully understanding the client's needs
2. Understanding the data, where it came from and cleaning it
3. Optionally gathering new data, for example, by web scraping or other means
4. Testing various solutions to the client's problem
5. Fully developing one or more solutions
6. Providing interim reports to the client
7. Providing a final written report, a final oral report with slides, and well-documented software during the last two weeks of the semester

During the semester, some of the classes will be lectures about presentation skills, computational approaches, and new statistical theory and practice related to specific projects. The class will be organized around milestones which include practice talks, presenting incremental progress to date, as well as written progress reports. Each project group will meet weekly with their mentor. Written reports and group presentations will be evaluated for feedback at least twice during the semester. Furthermore, a major learning component of the course comes from listening, evaluating and interacting with other groups about their work, so that you are not only learning from your project but also from the context and the issues that arise from all the course projects.

The specific objectives of this course are that by the end of the semester you will be able to:

1. Demonstrate effective interpersonal skills related to working with clients and with your group
2. Demonstrate effective skills for data cleaning, analysis, and model verification
3. Demonstrate effective written and oral presentation skills.
4. Work effectively in a group setting
5. Demonstrate production of software that is sufficiently well-documented to be useful to a client
6. Grasp the entire arc of the consulting project, and explain how to approach a new project
7. Demonstrate mastery of issues related to the ethics of human use.

Zoom & Participation

Most, likely all, of this course—classtime, office hours, etc.—will be conducted via Zoom. If public health conditions permit, we will consider some formal or informal face to face meetings. Classes on Zoom will be recorded and made available on a very limited basis for review.

In-class participation is especially important in this course, which is unfortunately a hardship for students in Asia. I will explore the possibility of holding class at a time that works for students in both North America and Asia, but you should plan on being awake and attending most or all classes, even if it is in the middle of the night.

Piazza

If you have questions that cannot be easily asked or answered in class or in office hours, Piazza is a good place to post your questions. You can answer or comment on each others' posts if you wish, or wait for one of the instructors to answer. Please be kind in your questions, answers and comments: On Piazza, there are no dumb questions, and no dumb mistakes.

While there may be some exceptions, Piazza is generally not a good place to ask specific questions about your project, your team, or your client; instead take this up directly with your project's faculty mentor or another course instructor.

Grading

Most of the class periods will be taken up by presentations by clients or student teams. You are expected to participate actively in class and interact both in class and outside with your team and with the client. The appropriate frequency of client meetings should be determined in discussion with your faculty mentor.

The final course grade will be calculated as follows:

Activity	Percentage of Final Grade
Project and related group activities	70%
Non-project homework	10%
Participation and engagement	20%

"Project and related group activities" includes meetings and interactions with your team, with your faculty mentor and with your client, as well as data analysis, report writing, oral presentations, etc. This category, and the category *"Participation and engagement"*, are graded somewhat subjectively, by necessity. Where possible we will provide rubrics and other guides for your work, but part of professional practice is simply showing up, engaging intellectually, and exhibiting pride of craft.

Other Course Policies

Attendance & Participation: It is essential to the class that you participate actively. This means you must do all the assignments including the readings; you must take part in discussions; and you must be present and prepared for class. On-time attendance is mandatory.

Team Participation: Every member of the project team is expected to contribute actively, productively, and professionally. We do not expect anything less than professional conduct and responsibility at all times.

Academic Integrity & Collaboration: Collaboration and discussion of readings and homework problem is encouraged. However, each student is required to write up any non-group HW solutions independently based solely on your own understanding. *See also the section on page 5 on academic integrity.*

Late-work/Make-up work policy: A part of professional practice is completing work on time. While I will usually allow a short "grace period" of an hour or so for classwork, I cannot guarantee that I will accept late submissions, unless we have discussed an extension together *in advance*. Work that you do for a client *must* be completed on time, and your ability to do so will be a factor in your course grade.

Academic Integrity

As members of a top-ranked academic institution, your academic integrity is assumed and expected. Please see <http://www.cmu.edu/academic-integrity/>; however, I expect each of you to behave well above these lower bounds. For all oral reports, intermediate and final papers, and other communications about your project, if you get ideas or words from a website, journal article, book, another person (in or out of this class), etc., cite the source in your communication, right where you use it. Then put a bibliography or list of sources cited at the end of the communication. *If you are not sure what is allowed, or required, please ask.*

Disability and other Special Needs

Carnegie Mellon makes great efforts to provide physical and programmatic campus access to everyone, as guaranteed by the Americans With Disabilities Act (ADA) and Section 504 of the Rehabilitation Act of 1973. If you have a documented disability, please let me know so that we can take whatever steps are needed to accommodate your needs. Contact CMU's Disability Resources office (<http://www.cmu.edu/hr/eos/disability/>) if you think you may have a disability and want to document it; or you have a documented disability that is not being adequately accommodated. For other issues and special needs, please contact a faculty mentor in this class, your academic advisor or another trusted mentor, and/or the Office of the Dean of Student Affairs (<https://www.cmu.edu/student-affairs/resources.html>).

A Note on Diversity

In this class, I will affirm and promote the inherent worth and dignity of every person, and I expect that every member of the class will do the same. The University is enhanced by the diversity of its members, in gender, sexuality, disability, age, socioeconomic status, ethnicity, race, nationality, religion and culture: each of us can contribute ideas and perspectives that no one else can. I will endeavor to present materials that are respectful of and accessible to all of our backgrounds and perspectives. Please let me know ways to improve the effectiveness of the course for you personally or for other students or student groups.

CMU Resources that may be useful to you include:

- The Center for Diversity and Inclusion: <https://www.cmu.edu/student-diversity/>
- The Intercultural Communication Center: <https://www.cmu.edu/icc/>
- The Office of Title IX Initiatives: <https://www.cmu.edu/title-ix/>

Wellness

At Carnegie Mellon University, we believe our individual and collective well-being is rooted in healthy connections, to each other and to campus resources. How we care for ourselves and others is important to our success. There are a wide variety of resources, opportunities, and people that want to help you thrive inside and outside of the classroom. Please see <https://www.cmu.edu/wellness/> for general wellness resources at Carnegie Mellon.

If you are experiencing challenges that interfere with your learning, your academic performance, or your ability to participate in daily activities, you can find out about confidential mental health services at <http://www.cmu.edu/counseling/>. Support is always available 24/7 from Counseling and Psychological Services: (+1) 412-268-2922.

Tentative Schedule

- The timing outlined below is tentative and subject to change, depending on how the course progresses.
- Depending on how projects go, we may decide to schedule some final presentation to clients in the last week of classes also.

Week	Dates	Agenda
Week 1	Feb 1, Feb 3	Project Proposal Presentations
Week 2	Feb 8, Feb 10	Project Proposal Presentations
Week 3	Feb 15, Feb 17	Topics in Statistical Practice
Week 4	Feb 22, Feb 24*	First Team Progress Report Presentations
Week 5	Mar 1, Mar 3	First Team Progress Report Presentations
Week 6	Mar 8, Mar 10	Topics in Statistical Practice
Week 7	Mar 15, Mar 17**	Second Team Progress Report Presentations
Week 8	Mar 22, Mar 24	Second Team Progress Report Presentations
Week 9	Mar 29, Mar 31	Topics in Statistical Practice
Week 10	Apr 5*** , Apr 7	Slack, or Additional Topics
Week 11	Apr 12, Apr 14	Third Team Progress Report Presentations
Week 12	Apr 19, Apr 21	Third Team Progress Report Presentations
Week 13	Apr 26, Apr 28	Finishing Project Work
Week 14	May 3, May 5	Finishing Project Work
Finals Week	May 10–17	Final Presentations to Clients, Schedule TBA

*Tuesday Feb 23 is a break day; no classes on this day.

**Friday Mar 19 is Mid-Semester Break; no classes on this day.

***Monday Apr 5 is a break day; no classes on this day.