PSYC 81.06: Storytelling with Data (Spring 2020)

Meeting times: MWF 2:10 - 3:15 PM **X-hour:** Th 1:20 - 2:10 PM

Classroom: Remote learning via Zoom (link may be found on Canvas)

Instructor: Dr. Jeremy R. Manning
Email: jeremy@dartmouth.edu

Office Hours: Held remotely on Zoom during X-hours

Course Description

In a world plagued by "alternative facts" but flush with "big data," how can we find truth? For example, can truth be objectively defined, or are there many equally valid truths? And does truth depend on the question we're asking, or is it a fixed property that we could somehow uncover with the right analysis? These sorts of question align with other deep questions about how we can really "know" something. For example, can we really ever hope to prove that the universe works in a particular way? If so, how? Or if not, what's the point of observing the world around us at all, or of becoming a scientist?

In this course we will define truth from a (somewhat cynical, but embarrassingly practical) psychological perspective: truth is the story about data that others find most convincing. To that end, we will examine tools and strategies for finding patterns in complex datasets, crafting convincing stories about those patterns, and communicating them to others. One major focus will be on creating impactful data visualization and infographics. The second major focus will entail you "getting your hands dirty" by applying the tools we learn about to various datasets and telling stories about data to your classmates. Along the way, we will leverage findings from the psychological and neuroscientific literatures to create more convincing stories by exploiting the visual system's ability to extract patterns from the world around us, and by examining what distinguishes effective stories from ineffective ones.

Course Goals

The primary goals of this course are to help you develop intuitions and tools for (a) finding meaningful patterns in complex datasets and (b) communicating your findings in an intuitive way. To that end, you will be learning to write basic computer programs (in Python), analyze and create plots, and present your work effectively. I see data interpretation and visualization as a deep psychological issue: creating compelling figures and stories is just as much about theory of mind of the receiver as it is about accurately reporting the deeper truths underlying the dataset. We will therefore motivate the tools and skills we develop from a psychological and neuroscientific perspective. My hope is that you will draw heavily on what you learn in my course in "real life," in whatever fields, professions, and/or hobbies, you choose to pursue.

Pre-Requisites

You should have taken a course on statistics or probability (e.g. PSYC 10, AP Stats, or similar). An online statistics course is fine as a stand-in, but I will expect you to know about hypothesis testing, probability distributions, and have some intuitions about how to design and interpret statistical analyses before you start this class. Prior course work or experience in Psychology, Neuroscience, Computer Science, Math, Engineering, Creative Writing, Theater, and/or Graphic Arts will also be useful.

In defining these pre-requisites somewhat vaguely, I hope we will end up with a group of students with a diverse array of backgrounds, interests, and skills. Some students may be more "quant" minded (they might have an easier time learning and applying the programming and analysis components of the course) and others will be more "humanities" minded (they might have an easier time crafting and presenting compelling stories). However, although you don't need to know how to program before you start this course, every student will eventually need to both program analyses and present results to the class. Ultimately we'll all build on our existing skills and experiences to help each other grow intellectually and learn new practical skills.

Course Materials

We will reference miscellaneous figures and chapters from three classics in data visualization by Edward Tufte (purchasing these books is **optional**, but they are excellent for drawing inspiration from a true master of data visualization):

The Visual Display of Quantitative Information (ISBN: 0961392142)

Envisioning Information (ISBN: 0961392118)

Visual Explanations: Images and Quantities, Evidence and Narrative (ISBN: 0961392126)

I will also provide various (free) PDFs for download via throughout the term.

You will also need an internet-enabled computer or tablet capable of displaying and outputting graphics and running a standard web browser (e.g., Firefox or Chrome). You'll use your computer to participate in Zoom meetings, engage with self-paced learning materials, and complete your assignments.

If obtaining these materials presents a financial or logistical hardship for you, please contact me as early in the term as possible.

Format and Overview

The overall format of the course will adapt to student interests and needs. My plan is for the first 4(ish) weeks of class to be primarily instructor-directed, supplemented by in-class discussions and assignments. We'll cover four "modules," with each module lasting for roughly a week: (1) What makes a good story?, (2) Visualizing data, (3) Python and Jupyter notebooks as a medium for data storytelling, and (4) Data science tools. These modules will introduce you to the fundamentals of storytelling with data.

During the remainder of the course, you (either in groups or individually) will download datasets, analyze them, produce figures, and informally present your results as "data stories" to the class each week. We will then constructively critique and make suggestions about each others' analyses and presentations, helping each other to become better data storytellers.

Slack

We will use <u>Slack</u> (a tool for organizing notes, files, and conversations) to provide a forum for asking and answering questions, posting demos, etc. You will need to join the class workspace at the beginning of the term by following the link. We'll set up a series of channels (one for each topic we decide to explore) and we will use the tool during and outside of class to keep track of our thoughts and ideas. Your first two assignments will be submitted using Slack.

GitHub

We will use <u>GitHub</u> to manage and share data and code. GitHub provides an easy way of managing multiple versions of data and code that may be easily shared and tracked. To use GitHub, you will need to create a (free) GitHub account at the beginning of the term. After the second assignment, all other coursework will be submitted via GitHub.

Google Colaboratory

We will use <u>Google Colaboratory</u> to develop and program analyses. This tool provides an easy means of organizing notes, code, and graphics in a single cohesive format ("notebook"). We will also use Google Colaboratory notebooks to tell our data stories; they combine analyses, reports, and presentations into a single document. The programming environment is accessed via a web-browser. (You will learn how to set them up and run the notebooks in Module 3 during the first part of the course.)

Grading

All Dartmouth courses in the Spring, 2020 term will be graded on a Pass/Fail basis. Achieving a "Pass" in this course will require engaging with the material regularly and submitting assignments. You will receive feedback on each assignment from me and your classmates as part of our learning process. However, I will not assign formal grades to your assignments this term.

The Academic Honor Principle

I expect you to abide by Dartmouth's <u>Academic Honor Principle</u> at all times. I encourage (and expect) you to discuss your assignments with your classmates. The class will be heavily collaborative, and I encourage group presentations and collaboration on assignments. However, it's also important that you contribute something substantive to each project that you participate in, and that you actively engage with the material. In addition, you cannot "re-use" projects from other courses without modifying them, although some projects will allow you to (optionally) build on prior work. Put simply, you should hand in your own (new) work, even if you collaborated or discussed your assignment with a classmate. If you have any questions about the Academic Honor Principle and how it applies generally to this course, or specifically to a particular assignment, please ask me.

Student Needs

I strive to maintain a welcoming and accessible classroom environment. I want you to be an active participant and contributor to ongoing discussions and activities, and that means that every student should feel comfortable in my (virtual) classroom. If you would like me to be aware of any issues that arise during the term, or any personal needs that may require adjusting how I run my class or how you participate, I encourage you reach out to me privately. Dartmouth's <u>Student Accessibility Services Office</u> can also help assist with setting up disability-related accommodations.