

SOCI 502 – 101 Research Design and Techniques (Quantitative)

Location: ANSO 134
Time: Thursday, 2-5 pm, Sept 8th – December 1st
Instructor: Nathan Lauster
E-mail: nathan.lauster@ubc.ca
Office: ANSO 1325

Course Description:

This course focuses on first principles for quantitative research design and analysis techniques in hands-on, how-to fashion. We will work with the R (and RStudio) data analysis software and Tidyverse package (all freely downloadable). Topics covered include basic programming & measurement; data structuring & management; data visualization; probability; prediction & causal inference; regression & uncertainty. The course is designed to highlight the value of basic “bread and butter” statistical proficiency and the construction of stylized facts. At the same time, the course will provide cursory exposure to research design and preparation for learning more advanced statistical techniques and engaging in discussions of causal inference.

Structure: I anticipate most classes will be divided into three parts, including a targeted discussion (with occasional lecture), a hands-on, small group workshop, and a Show & Tell “Figure of the Day!” segment.

Targeted Discussion – I’ll provide some basic concepts and grounds for discussion each day, often concerning but also extending the readings. Students should come prepared having read the day’s readings, but also with their own questions and thoughts.

Small Group Workshop - Statistics – and particularly statistical software - comprises a distinct language rewarding regular practice and, yes, homework. Students will be expected to complete exercises between classes. A portion of each class will be devoted to workshops dealing with exercises, where students will talk through their work together in small groups and assist one another with the previous assignment as well as set up work on the next assignment. Assignments will generally be turned in to me at the end of each class. Students are encouraged to assist one another, but all should turn in their own, distinct assignments.

Show & Tell – Applied discussion and critique are also an important part of learning quantitative methods. A set of students will be assigned to bring an interesting figure or other construction involving quantitative data (or another aspect of research, as assigned by topic) to class each day, ideally one evoking a strong reaction (love it or hate it). We’ll talk through the figures of the day or other show & tell items as a way of engaging with basic data analysis in the wild.

Software Requirements: R and RStudio (an interface for R that comes strongly recommended) can be freely downloaded at the CRAN <https://cran.rstudio.com/> & RStudio websites: <https://www.rstudio.com/products/rstudio/download/>

Readings: A selection of course readings will be referenced in the week-by-week section below, with full citation at the end of the syllabus. Overall we'll be drawing heavily upon Rafael Irizarry's (2022) open book, *Introduction to Data Science: Data Analysis and Prediction Algorithms with R*, an on-line and free copy of which can be found here: <https://rafalab.github.io/dsbook/>

Basic Schedule

M	Day	TOPICS	DUE	READINGS
Sept	8	Authority		<i>In Class</i> [Fry (2021) ; Borwein & Bailey (2013)]
	15	Intro to R	ex 1-2	Irizarry 1-2; Gailey (2021)
	22	Keeping Tidy	ex 3-4	Irizarry 3-4; Gryzmala-Busse (2021)
	29	Importing Data	ex 5-6	Irizarry 5-6; Pettinicchio & Maroto (2021)
Oct	6	Data Visual I	ex 7-8	Irizarry 7-8; Hirschman (2020)
	13	Data Visual II	ex 9-10	Irizarry 9-10; open
	20	Data Visual III	ex 11-12	Irizarry 11-12; open
Nov	27	Probability	ex 13-14	Irizarry 13-14; open
	3	Statistic infer	ex 15-16	Irizarry 15-16; Bradley, et al (2021)
	10	HOLIDAY		
	27	Regression	ex 17-18	Irizarry 17-18; open
	24	Linear Models	ex 19-20	Irizarry 19-20; Lundberg, et al (2021)
Dec	1	Presentations	Stylized Facts	

Note: "open" readings may still be assigned, but are otherwise left open for students to explore articles suiting their interest

Assignments:

Exercises – Irizarry provides exercise sets accompanying each chapter. Students should attempt to complete each exercise set due that day prior to class, as assigned. Bring questions about exercises to class to work on and improve during group work. Submit completed answers via Canvas. Note, some questions may be headscratchers, with more than one possible answer and programming solution. Feel free to explore and explain attempts that don't work out. Students' answers should be their own, but students are encouraged to help & learn from one another. Additional help and hints can be found on-line (e.g. <https://github.com/huynguyen250896/Introduction-to-Data-Science-BOOK-in-R->)

Stylized Fact – For their final project, students will be expected to construct a Stylized Fact of their own, illustrated and supported by an interesting figure. See examples from Data Visualizations within the journal Socius. Follow Socius guidelines for submission (as per here: <https://journals.sagepub.com/author-instructions/SRD>) but note that unlike Socius, I will not be charging a submission fee! Cap of approximately 500 words, plus figure, with room to add supporting figures and verbiage in appendices. Students will present their stylized facts on the final day of class (Dec 1st), with final document due date one week later. DUE Dec 8th.

Marking

Participation	20%	(come to class informed & engage)
Show & Tell	20%	(bring us data visualizations & critiques to discuss)
Assignments	30%	(complete & engage with assignments)
Stylized Fact	30%	(apply visualization lessons in publishable form)

Additional Readings:

Borwein, Jonathan & David Bailey. 2013. The Reinhart-Rogoff error – or how not to Excel at economics. *The Conversation*. April 22, 2013. <https://theconversation.com/the-reinhart-rogooff-error-or-how-not-to-excel-at-economics-13646>

Bradley, V.C., Kuriwaki, S., Isakov, M. et al. Unrepresentative big surveys significantly overestimated US vaccine uptake. *Nature* (2021). <https://doi.org/10.1038/s41586-021-04198-4>
<https://www.nature.com/articles/s41586-021-04198-4>

Fry, Hannah. 2019. What Statistics Can and Can't Tell Us About Ourselves. *The New Yorker* Sept 2, 2019. <https://www.newyorker.com/magazine/2019/09/09/what-statistics-can-and-cant-tell-us-about-ourselves/amp>

Fry, Hannah. 2021. When Graphs are a Matter of Life and Death. *The New Yorker* June 14, 2021. <https://www.newyorker.com/magazine/2021/06/21/when-graphs-are-a-matter-of-life-and-death>

Gailey, Sarah. 2021. Do Hippos Count as Dragons: An Examination of Identity and Taxonomy. *Tor.com*, July 26, 2021. <https://www.tor.com/2021/07/26/do-hippos-count-as-dragons-an-examination-of-identity-and-taxonomy/>

Gryzmala-Busse, Anna. 2021. Tilly Goes to Church: The Medieval and Religious Origins of the European State. *Broadstreet Blog*. <https://broadstreet.blog/2021/07/23/tilly-goes-to-church-the-medieval-and-religious-origins-of-the-european-state/>

Hirschman, Daniel. 2016. Stylized Facts in the Social Sciences. *Sociological Science* 3: 604-626.

Lundberg, Ian, Rebecca Johnson, Brandon Stewart. 2021. What Is Your Estimand? Defining the Target Quantity Connects Statistical Evidence to Theory. *American Sociological Review* 86(3): 532-565. <https://doi.org/10.1177/00031224211004187>

Pettinicchio, David & Michelle Maroto. 2021. Who Counts? Measuring Disability Cross-Nationally in Census Data. *Journal of Survey Statistics and Methodology* 9(2): 257-284. <https://doi.org/10.1093/jssam/smaa046>

COVID adjustments – We will be following university & public health guidelines throughout the course. See <https://covid19.ubc.ca/> for more. The class will always welcome & support mask use indoors. If you are feeling ill, please do NOT attend class in person! Supports will be provided to catch you up, as possible. Please contact me with any questions or concerns or important updates through the term.