# USP 634: Data Analysis I Spring 2023

- Times: **Tuesdays 9:00 am - 11:30 am**
- Location: In person URBN 225 and via Zoom (https://pdx.zoom.us/j/5037255130)
- Computer Lab: In person via Zoom, Tuesdays 11:40 am - 12:30 pm
- Instructor: Professor Liming Wang ([lmwang@pdx.edu](mailto:lmwang@pdx.edu))
- Office Hours: via Zoom ([https://pdx.zoom.us/j/5037255130](https://pdx.zoom.us/j/5037255130))
                          Mondays 2:00-3:30 pm or by appointment
- Course Website: [https://canvas.pdx.edu](https://canvas.pdx.edu)

## Synopsis

This course is designed to give students the necessary skills to analyze data for their research projects. Together with the second course, USP 654 Data Analysis II in the fall, the sequence will provide a thorough and reasonably comprehensive introduction to understanding, critically evaluating, conducting, and writing about analyses for most studies in social science-related disciplines. This course covers descriptive statistics, probability and distributions, hypothesis testing, association, and simple regression analysis. USP 654 Data Analysis II will cover regression analysis in more depth.

## Prerequisite:

College-level algebra, introductory probability and statistics and an understanding of research design. The most critical elements of these materials will be reviewed in class. I can provide optional materials that will help provide additional review if you need it.

## Format

Classes will be a combination of lectures, discussion and computer labs. Students are expected to read assigned readings before class and to participate in class discussions. Homework assignments will be given and analysis of these assignments will be the basis for some class discussion during the class following their due date, so it is essential to complete assignments on time.

## Software

The assignments for this course will use the R statistical software (or your choice of other stats software if you're already a proficient user, such as SAS, SPSS, or Stata).

The lab instructions will use R and RStudio. The software is installed on lab computers across the campus and you are encouraged to download and install on your own computer/laptop. R is

free and available for download at http://cran.r-project.org. We will use a RStudio (https://www.rstudio.com/) as our main interface to R. An alternative is jamovi (https://www.jamovi.org/), a beginner friendly interface to R.

Labs will be offered weekly to assist in using R to complete the assignments and R examples will be used during regular sessions. I can provide additional assistance with the software. If you plan to use other software, I will provide as much assistance as I can.

# Textbook and Reading

The recommended textbooks for the course are

- [Wickham] Grolemund, G., H. Wickham, 2017. R for Data Science, O'Reilly.
- [Fogarty] Brian J. Fogarty, 2019. Quantitative Social Science Data with R: An Introduction. SAGE Publications Ltd.  OR
- [Llaudet] Llaudet, E. and Imai, K. 2022. Data Analysis for Social Science: A Friendly and Practical Introduction. Princeton Press.

Optional textbooks are:

1. Myers, J.L., Well, A.D., and Lorch, R.F.,Jr., 2010. Research Design and Statistical Analysis: 3rd Edition. Mahwah, NJ: Erlbaum.
2. Wheelan, C., Davis, J., 2014. Naked Statistics, Unabridged edition. ed. Brilliance Audio.
3. Urdan, T.C., 2010. Statistics in Plain English, 3rd ed. Routledge, New York.
4. Gonick, L. and Smith, W., 1993. The Cartoon Guide to Statistics 1st Edition. Collins Reference.

There will also be supplemental readings that will be available for download from the course website.

# Grading

The course requirements include assigned readings, four assignments, attendance and class participation, and a research paper.

- **Datacamp "courses" (25%)** We will use the interactive courses on Datacamp to learn to use R for various aspects of data analysis. We will dedicate some of the lab hours for you to work on these courses, but you're expected to finish them outside labs. There are no deadlines for these courses as long as they are completed by 6/11. Use this link to register your account to get access to all Datacamp courses free of charge:
  - Introduction to R
  - Introduction to the Tidyverse
  - Data Visualization with ggplot2
  - Exploratory Data Analysis
  - Hypothesis Testing

- **Assignments (45%)**: Three homework assignments that involve completing analytical exercises designed to build skills in statistical analysis and to reinforce concepts discussed in the class lectures. Clarity in writing and presentation will be taken into account in grading. Students may collaborate on the approach to and analysis of homework assignments but are expected to prepare and present results and interpretations independently.
- **Research Project (25%)**: In this class project, you will explore a topic of your choosing in depth, applying statistical methods covered in this class (and beyond) for your data analysis. Ideally, you would build upon work you have already done for other courses, such as USP 630 Research Design, USP 532 Data Collection. The paper must include original data analysis. This most likely will be analysis of secondary data. If you want to collect your own quantitative data, it may be possible, but will take more planning and work. Check with me first.
  - 10-15 pages, double-spaced, not including figures, tables and references;
  - Initial proposal due on 04/18. It should include a brief description of the topic, a list of key references (including academic literature), specific research question(s) and hypotheses, and the data source(s) you plan to use;
  - Project updates due on 05/23. This should include a draft of the introduction/background sections, a brief literature review, the methodology and description and descriptive statistics of your data, as well as initial analysis results;
  - Project presentation on 06/13 (tentative);
  - Final paper due on 06/16.

  The best papers (high A grade) will be worthy of submission to an academic journal. Use empirical journal articles as a model to follow in terms of format, style, etc.

  You must cite all of your sources in your work. Please review this web site about citations: http://library.pdx.edu/citing_sources.html

- **Research Project Presentation (5%)**

**Late assignments will be marked down** – one-third of a grade per day late. "One-third of a grade" is, for example, from A to A-, B+ to B, etc. (or 3.3 points using the 100 point scale). We generally do not allow students to turn things in late without assessing this penalty, except in unusual circumstances, e.g. medical emergencies. A request for such an extension must be made before the due date. Having too much work in other classes or at work/internship does not count. All students have those challenges.

# Schedule

The table below shows by date lecture topics, computer labs, and readings, and dates that assignments will be handed out and due. Supplement readings will be posted on course website.

| Week Date | Topic | Readings | Assignment / Deadline | Datacamp Courses |
|---|---|---|---|---|
| 1 04/04 | Course Overview; Stats and Research Design Computer lab: R/Rstudio installation; Computer setup; R basics | Why Statistics? Freese, Secondary Analysis of Large Social Survey | | Introduction to R |
| 2 04/11 | Data analysis workflow; tidy data RStudio Workflow; tidyverse | Wickham Ch 1-2 | HW1 assigned | Introduction to the Tidyverse |
| 3 04/18 | Data Visualization Computer lab: data visualization | Wickham Ch 3 | Initial research proposal due | Data Visualization with ggplot2 |
| 4 04/25 | Descriptive Statistics; Exploratory Data Analysis Computer lab: descriptive stats; exploratory data analysis | Wickham, Ch 7 Fogarty, Ch 7 | | Exploratory Data Analysis |
| 5 05/02 | Probability and Distributions; Foundation for Inferential Statistics Computer lab | Llaudet, Ch 6-7 | HW1 due; HW2 assigned | |
| 6 05/09 | Confidence Interval and Hypothesis Testing (NHST) Computer lab | Fogarty, Ch 6 & 9 Stefano 2004 | | |
| 7 05/16 | Hypothesis Testing for Continuous Variable: t-test and ANOVA Computer lab: t-test and ANOVA | Fogarty Ch 9 | Project update due | Hypothesis Testing |
| 8 05/23 | Hypothesis Testing for Categorical Variable Computer lab | Fogarty Ch 10 Spence 2007 | HW2 due; HW3 assigned | |
| 9 05/30 | Association and Correlation Computer lab | Fogarty Ch 10 | | |
| 10 06/06 | Regression Analysis Computer lab | Fogarty Ch 11-12 | | |
| 11 06/13 | Class project presentation | | HW3 due | |
| 11 06/16 | | | Class project due | |

# ACCESS AND INCLUSION FOR STUDENTS WITH DISABILITIES

PSU values diversity and inclusion; My goal is to create a learning environment that is accessible, equitable, inclusive, and welcoming. I am committed to fostering mutual respect and full participation for all students. If any aspects of instruction or course design result in barriers to your inclusion or learning, please notify me. Additionally, the Disability Resource Center (DRC) provides reasonable accommodations for students who encounter barriers in the learning environment. The DRC works with students who have physical, learning, cognitive, mental health, sensory, and other disabilities.

If you have, or think you may have, a disability that may affect your work in this class and feel you need accommodations, contact the Disability Resource Center to schedule an appointment and initiate a conversation about reasonable accommodations.

If you already have accommodations, please contact me to make sure that I have received your faculty notification letter from the DRC so we can discuss your accommodations.

The DRC is located in 116 Smith Memorial Student Union, Suite 116. You can also contact the DRC at 503-725-4150 or, drc@pdx.edu. Visit the DRC online at https://www.pdx.edu/disability-resource-center.

# Title IX reporting obligations

As an instructor, one of my responsibilities is to help create a safe learning environment for my students and for the campus as a whole. Please be aware that as a faculty member, I have the responsibility to report any instances of sexual harassment, sexual violence and/or other forms of prohibited discrimination. If you would rather share information about sexual harassment, sexual violence or discrimination to a confidential employee who does not have this reporting responsibility, you can find a list (http://www.pdx.edu/sexual-assault/get-help) of those individuals. For more information about Title IX please complete the required student module Creating a Safe Campus in your Canvas.