

POLI 888 – Transatlantic Data Analysis

Spring 2026

Lecture: Friday 9:05am–12:05pm
Murphey Hall – Room 0115

Instructor Information

Instructor: Jaime Lindh
Office: Hamilton 451
Email: jlindh@unc.edu
Office Hours: Thursday 1:00pm–4:00pm ([schedule here](#))

Course Description

This course introduces applied data analysis for social science research, with an emphasis on fostering critical analytical thinking in the creation and communication of information within professional and academic settings. Throughout the course, we will examine applied examples from scholarly articles that focus on transatlantic content. After completing this course, students will gain the skills to analyze data and present multivariate data analysis. Additionally, the course aims to provide students with the quantitative toolkit to develop the master's thesis.

To achieve these goals, the course not only reviews basic concepts in statistics but also explores practical programming skills using R for effective data analysis and communication. Students will learn to transform raw data into understandable information by leveraging R to import, tidy, transform, and communicate their findings. Starting with the basics of the R working environment, the course covers arithmetic and logical operators, as well as essential functions for data manipulation. Additionally, students will explore common data structures, variables, and data types in R, gaining hands-on experience in writing R scripts and constructing publishable graphs, tables, and reports. In terms of statistics, the course covers measures of central tendency, dispersion, and association, probability, hypothesis testing, and linear regression.

Learning Objectives

General

- Develop critical thinking about the development and presentation of information in professional and academic settings.
- Expand skills in the application of multivariate data analysis.

Specific

- Understand the fundamental syntax of R.

- Import data from a variety of formats into R.
- Tidy data for preparation and analysis.
- Transform data by applying programming language in R.
- Analyze data and present findings using appropriate R packages.
- Visualize data attributes using R packages.
- Develop a report which communicates information in a clear and professional manner.
- Understand linear regression analysis.

Class Structure

Lectures will be divided into two parts. The first portion will consist of an exposition of the fundamental content related to the weekly topic, accompanied by examples primarily drawn from transatlantic studies to illustrate the application of analysis in R. This will enable students to familiarize themselves with datasets in transatlantic studies pertaining to public opinion, party systems, and governmental administration.

In the second part, a lab session will be conducted, giving students the opportunity to implement some of the techniques covered earlier. This portion will function as a workshop in which students work together with a peer to solve exercises. Throughout the semester, this session will also serve as an opportunity for students to apply newly acquired concepts to their research projects.

Readings

We will primarily rely on four resources:

- *An Introduction to Political and Social Data Analysis Using R* by Thomas M. Holbrook (IPS).
- *Data Analysis for Social Science* by Elena Llaudet and Kosuke Imai (DASS).
- *R for Data Science* by Hadley Wickham, Mine Çetinkaya-Rundel, and Garrett Grolemund (RDS).
- *ggplot2: Elegant Graphics for Data Analysis* by Hadley Wickham, Danielle Navarro, and Thomas Lin Pedersen (EGD).

Requirements and Evaluation

Your grade in this course will be determined by the components outlined below. Failure to fulfill the course requirements will lead to a failing grade.

Lab Exercises (10%)

After each lab session, students must submit their work. Grading follows an incomplete/complete structure.

Problem Sets (35%)

During the semester, you can expect a series of four assignments. These assignments will become available two weeks before the deadline. It is imperative that all problem sets are submitted by Friday of the specified week (please consult the calendar below). Collaborative efforts on these assignments are welcomed and encouraged, but it is essential for each student to submit individual answers. Although some similarities in submissions are expected, it is also vital that your code and responses differ from those of your classmates. Lastly, make sure to take advantage of office hours for discussions related to these assignments.

Research Project (55%)

The class features a research project in transatlantic studies in which students will conduct and present original data analysis. Students have the option to work on a project based on a previous course submission or choose a completely new one (if the former is the case, please inform me). I encourage you to discuss your research project with me at the beginning of the semester. The project will involve the following four stages:

- **First delivery (5%):** In this submission, students must present four components: (i) a research question to answer; (ii) a brief motivation, explaining why answering this question is important; (iii) a theoretical discussion explaining the relationship between the main variables (i.e., why we should expect a relationship between the main independent variable and dependent variable); (iv) a list of hypotheses to be tested.
- **Second delivery (10%):** Students are required to submit the following additional elements: a research design, encompassing (i) justification for their case selection, (ii) how the main variables of interest will be conceptualized and measured, and (iii) the empirical analysis planned to test the hypotheses. Additionally, they must present (iv) the data to be used and (v) descriptive statistics.
- **Presentation (10%):** Students are required to present an advanced stage of their final research project in class, covering the research question, motivation, theoretical discussion, hypotheses, research design, data analysis (descriptive and regression analysis), and conclusion. Feedback from both the instructor and classmates will be provided before the final submission.
- **Final delivery (30%):** The final delivery must include feedback from previous submissions and incorporate a section with the data analysis, including tables and figures.

Grading Scale

Score	Grade
≥ 95	H
≥ 75	P
≥ 60	L
< 60	F

Class Policies

Attendance

Students are expected to attend all sessions, with attendance understood to be in-person. While your grade will not be directly affected by attendance, participating in class is an integral part of the learning process. If, for any reason, you cannot attend a class, please submit the necessary paperwork when appropriate.

Late Work

Late submissions with valid justification will be accommodated. In the event of anticipated delays due to medical or other emergencies, please contact me at the earliest opportunity to explore possible solutions. For all other situations, a penalty will apply.

Technology and Generative A.I.

Given the course's nature, you will typically be utilizing your computers during each class. However, keep in mind that any external use of your computer will adversely affect not only your learning process but also that of your classmates. Unless absolutely necessary, please stow away your phone before the class starts.

You might have had the opportunity to work with generative A.I. models like ChatGPT. These tools are highly potent and can serve as significant resources for data analysis. Nevertheless, it is crucial to recognize that these tools supplement rather than substitute the critical analysis integral to researchers' work. Importantly, this class fully adheres to the university's guidelines on generative A.I. use, which you can review at the following [link](#). Failure to adhere to these guidelines may be a reportable violation to the UNC Honor Court.

Academic Integrity

The Honor Code applies to all your assignments. Please review it [here](#). Plagiarism involves intentionally or carelessly presenting someone else's words, thoughts, or ideas as your own without proper attribution when submitting academic work, whether for grading or other purposes. This includes (but is not limited to) using someone else's code to answer your assignment. You have the option to take a tutorial [take a tutorial](#) at the library.

Title IX Resources

Any student who is impacted by discrimination, harassment, interpersonal (relationship) violence, sexual violence, sexual exploitation, or stalking is encouraged to seek support on campus or in the community. Reports can be made online to the [University Compliance Office](#). Please contact the [University's Title IX Coordinator, Report and Response Coordinators in the Equal Opportunity and Compliance Office](#) , or the [Gender Violence Services Coordinators](#) to discuss your specific needs.

Students with Disabilities

UNC supports the provision of reasonable accommodations, encompassing resources and services, to assist students with disabilities, chronic medical conditions, temporary disabilities, or pregnancy complications that may pose barriers to fully engaging in University courses, programs, and activities. If you are a student with specific needs or a disability, please reach out to me to discuss and make appropriate arrangements. You will also need to contact the [University Compliance Office](#).

Mental Health and Counseling

Graduate school can present challenges even under typical circumstances. The Counseling and Psychological Services is dedicated to supporting the mental health requirements of student population. They provide prompt access to consultations and connect students with clinically services, whether for short-term or long-term needs. Explore their [website](#) to gather more information.

Calendar

Session	Topic	Objective	Contents	Readings	Assignment due
1 (01/09)	Introduction to data analysis	• Review of the Syllabus and course objectives • Introduce data analysis	• Research process: an overview • Research question • Theory and hypotheses • Conceptualization and measurement	IPS 1	
2 (01/16)	Introduction to R programming	• Introduce R programming • Understand good programming practice	• Installing R and RStudio • RStudio Overview • Arithmetic Operators • Logical Operations • Getting Help in R • Quitting Rstudio • Writing R scripts – code style	IPS 2.1–2.3 RDS 2	
3 (01/23)	Data import and cleaning	• Understand data importation • Identify different types of data and objects in R • Introduce basics in cleaning data in R	• Setting up your working directory • Downloading and importing data • Installing and loading packages • Types of data: numeric, character, logical • Types of objects: vector, factor, list, dataframe • Extracting a subset of a data frame • Sorting numeric, character, and factor vectors • Renaming variable names, recoding, creating variables • Missing data	IPS 2.4–2.6 DASS 1	Assignment 1

Session	Topic	Objective	Contents	Readings	Assignment due
4 (01/30)	Data visualization	• Comprehend how to develop univariate and bivariate visual analyses using <code>ggplot</code>	• Introduce to <code>ggplot</code> • Histograms • Box plots • Scatter Plots • Combination • Customize charts and graphs (labels, annotations, themes)	RDS 1 EGD 2 RDS 9–11	Term project: First delivery
5 (02/06)	Descriptive statistics	• Explore data patterns to inform research design decisions • Understand how to do and interpret univariate and bivariate analyses in R	• Counting outcome • Central tendency measures • Variability measures • Functions: summary, describe, and others • Analysis by group • Correlations • Crosstabulation using <code>gmodels</code>	IPS 3–5, 6	
6 (02/13)	Data transformation	• Apply conditional statements for data transformation • Introduce <code>dplyr</code>	• If/else • Boolean logical operators • <code>dplyr</code> and <code>when_case</code> • Rows transformation • Column transformation • Pipe • By group	RDS 3	Assignment 2
7 (02/20)	Statistical inference I	• Understand basic of probability and sampling	• Probability • Normal distribution • Statistics and parameters • Sampling error • Sampling distribution • Confidence intervals	IPS 7–8 DASS 6	
Canceled (02/27)					
8 (03/06)	Statistical inference II	• Use statistical inferences through the process of hypothesis testing	• Significant test for means • Two-samples difference in means • Test for proportions	IPS 9–10	
9 (03/13)	Linear Regression Model – Part I	• Introduce and interpret simple linear regression models	• Simple linear regression • Ordinary Least Squares estimator • Linear regression in R • Coefficients interpretation • Model assumptions and goodness of fit	IPS 15 DASS 4	Assignment 3

Session	Topic	Objective	Contents	Readings	Assignment due
Spring break (03/20)					
10 (03/27)	Linear Regression Model – Part II	• Understand multiple linear regression and omitted variable bias • Report regression results	• Omitted variable bias • Control variables • Multiple linear regression • Categorical independent variables • Interpreting partial effects • Export and plot OLS results	IPS 16 DASS 5	
Spring holiday (04/03)					
11 (04/10)	Linear Regression Model – Part III	• Incorporate nonlinear and interaction terms in linear regression	• Interaction terms • Quadratic terms • Predicted outcomes • Marginal effects • Robustness checks	IPS 16 DASS 5	Term project: Second delivery
12 (04/17)	Research project workshop				
13 (04/24)	Final Presentation				Assignment 4
Exam week (05/06)					Term project: Final delivery