

CS&SS/POLS 512

Lab1: Working with Time Series and Panel Data in R

Ramses Llobet

Welcome!

- ▶ Welcome to the first lab session of CS&SS / POLS 512!
- ▶ I am Ramses Llobet (rllobet@uw.edu), I am a Ph.D. candidate in Political Science.
- ▶ My research interest are in **political economy** and **applied statistics**.
- ▶ Please **DO NOT** hesitate to stop me if you don't hear or understand me properly.
- ▶ **DO NOT** hesitate to ask questions. No question is silly. :)

Acknowledgement

The section materials are adapted from previous versions made by former TAs to CSSS/POLS 512:

- ▶ Tao Lin
- ▶ Inhwan Ko
- ▶ Daniel Yoo

Department of Political Science, University of Washington.

Labs, Office Hours, and Homework

1. Lab Sessions: Fri, 1:30-3:20 pm via [Zoom](#).
 - ▶ Covers application of material from lecture using examples; clarification and extension of lecture material; Q&A for homework and lectures.
 - ▶ Materials will be available on the [course website](#), and will be offered in a compact .zip file.
2. Office Hours: by appointment via [Zoom](#).
 - ▶ In addition, [Slack channel](#) is available for trouble shooting and specific questions about homework and lecture materials, etc.
 - ▶ See the syllabus for more info.
3. Homework: tree problem sets.
 - ▶ Using Overleaf or RStudio with R Markdown with write up in LaTeX

Expectation and Goals

4. When this course is over, you should be able to do the following (and more):
 - ▶ Identify and understand time series **dynamics**: seasonality, deterministic trends, moving average and autoregressive processes.
 - ▶ Understand nonstationary time series, perform unit root tests, fit ARMA and ARIMA models, use cross validation for model assessment.
 - ▶ Analyze multiple **continuous** time series using vector autoregression, perform cointegration tests, and estimate error correction models for cointegrated time series.
 - ▶ Distinguish between **random** or **fixed** effects and decide when each of these are appropriate.
 - ▶ Understand **Nickell bias** and use an instrumental variable approach with GMM to address the issue.
 - ▶ Perform **multiple imputation** and in-sample simulations for panel data.

Expectation and Goals

5. The course **moves fast**: you should be comfortable doing the following for the homework assignments and project.
 - ▶ tidying and transforming data, especially time series and panel data.
 - ▶ importing and exporting datasets.
 - ▶ generating plots of your data and results.
 - ▶ writing basic **functions** and **loops** for repeated procedures.
6. Fortunately, for those of you new to R, there are many resources to get you up to speed
 - ▶ Books: e.g. [Grolemund \(2014\)](#), [Wickham et al. \(2023\)](#), etc.
 - ▶ Online Courses: e.g. [CSSS 508 Introduction to R for Social Scientists](#)

Three levels of Methodological Sophistication

▶ **Blind Consumer**

- ▶ *"I include two-way fixed-effects in my model because I read some famous econometrician say in top academic journals that everyone should."*
- ▶ *"I want to apply causal forest to my project because it seems so fancy."*

▶ **Critical Consumer**

- ▶ *Choose appropriate methods based on your research question and data from wide range of available tools.*
- ▶ *Understand how canned packages actually work internally.*

▶ **Methodologist/Developer**

- ▶ *Identify common methodological problems in your field and provide a solution.*
- ▶ *Write software packages for public use.*

Computation Software

7. Please make sure that you have R or RStudio installed on your computer
8. If you would like to learn how to use LaTeX, this is a great opportunity to do so
 - ▶ An easy way to get introduced to this is to use R Markdown within RStudio
 - ▶ Make sure you have TeX installed, which you can find [here](#)
 - ▶ Make sure you have R Markdown installed using `install.packages("rmarkdown")`
 - ▶ Now in RStudio, choose File → New File → R Markdown

R Markdown

8. Using R Markdown

- ▶ Choose to compile your document as a **PDF** and give it a title.
- ▶ Embed your code within the **code chunk** area, and write up your text outside.
- ▶ Then press `Knit` to render the document.
- ▶ Feel free to browse my [CS&SS 321 slides](#) if you need a refresh in R/RMarkdown basics.

More on R Markdown

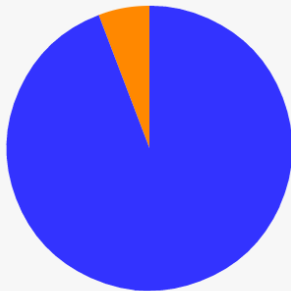
- ▶ You can use R Markdown to write an academic paper
 - ▶ Control chunk options such as `include` (hide code blocks) or `fig.align` (adjust the alignment of figures)
 - ▶ `install.packages("tinytex")` and install [pandoc](#).
 - ▶ You can load LaTeX packages or `.tex/.sty` files in the YAML header
- ▶ R Markdown can not only run R, but also run python!
 - ▶ `install.packages("reticulate")`
 - ▶ Use `py$...` to call objects from previous python chunk to R chunk.¹
- ▶ In 2022, RStudio release [Quarto](#).
 - ▶ An ambitious “next-gen” tool that aims to replace R Markdown and Jupyter Notebook.
 - ▶ The slides for today’s section is powered by [this!](#)

¹See https://rstudio.github.io/reticulate/articles/calling_python.html.

How to Look for Help?

How I feel when I debug code as a new hire

■ Time spent finding error
■ Time spent on coding



imgflip.com

How to Look for Help?

- ▶ Just google the error message or find them on Stack Overflow!
- ▶ We use [Slack](#) for Q&A and troubleshooting.
 - ▶ People encountering similar problems can see how to solve them (avoid reinventing the wheel).
 - ▶ It has the added benefit of facilitating **knowledge spillover** through peer discussion and **mutual assistance**.
 - ▶ If you are not comfortable with public post, you can send me a private message.
 - ▶ If you have more questions that cannot be covered by one single post, I encourage you to **set an appointment** for office hours.
- ▶ Besides coding issues, you are also welcomed to ask questions about the choice of methods, research design, etc.

How to Look for Help?

Developers searching for information



More on Troubleshooting

- ▶ Minimal Reproducible Example (MRE)
 - ▶ “minimal”: “look in a *smaller* stack to find a needle”
 - ▶ inputs are small and simple
 - ▶ fewer packages loaded
 - ▶ fewer function calls
 - ▶ reproducible: provide code that someone else could run

Also see: <https://www.rstudio.com/resources/webinars/help-me-help-you-creating-reproducible-examples/>

Potential Data Source for Your Final Project

- ▶ Awesome lists of public datasets in GitHub
 - ▶ [Awesome Public Datasets](#)
 - ▶ [Awesome collections on DataHub](#)
 - ▶ Just search “dataset” in GitHub!
- ▶ Comparative politics/world politics
 - ▶ Government: [World Economic and Politics \(WEP\)](#), [Quality of Government \(QoG\)](#), [Variety of Democracy \(V-Dem\)](#), etc.
 - ▶ Survey: [World Value Survey \(WVS\)](#);
[\(Asian/Afro/Latino/...\)Barometor Survey](#), etc.

Questions?