



EVANS SCHOOL
OF PUBLIC POLICY & GOVERNANCE

UNIVERSITY of WASHINGTON

PUBPOL 542-A: COMPUTATIONAL THINKING
FOR GOVERNANCE ANALYTICS

Prof. José Manuel Magallanes, PhD.

Winter 2021

E-mail: magajm@uw.edu
Office Hours: By appointment

Website: <https://canvas.uw.edu>
Class Hours: 8:30-11:20 (Tuesdays)

Course Description

This course is an innovative combination of Data Science and Computational Social Science particularly structured for policy studies. It introduces algorithmic thinking to policy students to solve problems related to data pre-processing, optimization via linear programming, and the modeling of governance complexity via agent-based simulation.

The participants will also develop good practices in reproducibility for professional work in policy research. Every session of the course follows a hands-on approach (laptops needed). Tools to be taught and used include:

- Programming languages (Python/R/Netlogo).
- Version control apps (GitHub).
- Document preparation tools (markdown)
- Advanced spreadsheet functionality.
- Computational representation of governance complexity.

There is no pre-requisite for the course.

Course Objectives

1. Carry-out analytic work following a reproducibility/replicability approach.
2. Master basic data management and analysis principles.
3. Become an effective user of programming languages (without becoming a programmer!).
4. Be a reflective user of data science in complex governance settings.
5. Use computational tools to aid complex decision-making.

Course Book

Magallanes Reyes, J. M. (2017). *Introduction to Data Science for Social and Policy Research: Collecting and Organizing Data with R and Python*. Cambridge University Press, 1 edition edition.

Recommended Readings

- Python Language Documentation: <https://www.python.org/doc/>
- R Manuals: <https://cran.r-project.org/manuals.html>

Software installations required

Students have to install the following software in their computers:

- ANACONDA by Continuum Analytics (choose according to your Operating System):
<https://www.anaconda.com/products/individual>
- R (choose according to your Operating System):
<https://cran.r-project.org/>
- RStudio Desktop Personal License (choose according to your Operating System)
<https://www.rstudio.com/products/rstudio/download/>
- NetLogo (choose according to your Operating System)
<https://ccl.northwestern.edu/netlogo/download.shtml>
- GITHUB:
Get an account at <https://github.com/>,
and then download the desktop app from <https://github.com/> .
- OpenSolver: If you have a gmail account, we will use it from there (no need to install it now).

Evans School Community Conversation Norms

This course has adopted the Evans School Community Conversation Norms. Please be aware of these norms in interactions with the instructor and other students. At the Evans School, we value the richness of our differences and how they can greatly enhance our conversations and learning. As a professional school, we also have a responsibility to communicate with each other—inside and outside of the classroom—in a manner consistent with conduct in today’s increasingly diverse places of work. We hold ourselves individually and collectively responsible for our communication by:

- Listening carefully and respectfully
- Sharing and teaching each other generously
- Clarifying the intent and impact of our comments
- Giving and receiving feedback in a “relationship –building” manner
- Working together to expand our knowledge by using high standards for evidence and analysis

Changes to the Syllabus

The professor reserves the right to make changes to the syllabus during the quarter. The professor will notify students immediately by email and in class if any changes are made.

Grading Policy

Grades are based on two items:

- **Small Project:** A simple project on data analytics will be presented after the second unit (after Python and R are covered). The basic elements the small project should have are:
 1. Using Python to prepare a data set. This includes collecting, cleaning and integrating. As this work will be done in groups, each member of the group should bring a data set with one variable for the same unit of analysis (countries, counties, etc.). Integrating data already cleaned will give you **0.5** points on the final grade, if data goes through cleaning will give you **1.0** points.
 2. Using at least one of the techniques covered in R. Using one technique will give you **0.5** points on the final grade. Using for both dimension reduction and modeling will give you **1.5** points.

Two reports should be written in a tutorial-like fashion (one in Python and one in R), explaining a motivation and purpose in each case, as well a rationale for each step taken.

- **Homeworks:** There will be two homeworks based on the remaining topics. Each will give you **0.5** points on the final grade.

The remaining 0.5 points for the final grade will be based on within group evaluation. The members will decide who will receive those points.

Working in groups

This course requires working in groups. You should have decided who will be in your group by the end of the **first** class. The professor will decide the amount of groups and their sizes.

Course Schedule

Week 01, 01/04 - 01/08: Data Pre Processing using Python:

Data gathering

- Uploading files.
- Scraping tables from the web.
- Using APIs.
- Data from social apps.

Week 02, 01/11 - 01/15: Data Pre Processing using Python:

Data cleaning and Formatting

- Removing or cleaning cell values using regular expressions.
- Setting values in the right measurement scale.

Week 03, 01/18 - 01/22: Data Pre Processing using Python:

Integration

- Reshaping.
- Merging and appending different data sets.
- Saving files.

Week 04, 01/25 - 01/29: Data Analytics in R: Dimensionality reduction

- Cluster analysis.
- Factor analysis.

Week 05, 02/01 - 02/05: Data Modeling in R I

- Linear regression.

Week 06, 02/08 - 02/12: Data Modeling in R - II

- Logistic regression.

Week 07, 02/15 - 02/19: Optimization for Decision Making - I

- Introducing operations research.
- Modeling problems in linear programming.

Week 08, 02/22 - 02/26: Optimization for Decision Making II

- Using spreadsheet to solve linear programming problems.
- homework 1.

Week 09, 03/01 - 03/05: Modeling Governance Complexity I

- Network Analysis.
- homework 2 (alternative 1).

Week 10, 03/08 - 03/12: Modeling Governance Complexity II

- Complexity and policy making.
- Agent based-modeling.
- homework 2 (alternative 2).

Week 11, 03/15 - 03/19: Final week: submitting final version of small projects and homeworks.

Remote teaching protocol

- Attendance is recommended, but every session will be recorded in case you can not attend.
If a student attends, his/her camera should be activated.
- Group meetings need to be recorded.