PL/CS 362-1: Applied Computational Methods for Social Sciences

Details

Instructor: Bogdan G. Popescu
Hours: MW 10:00-11:15AM
Total Hours of Contact: 2:30 per week
Room: G.K.1.4-Guarini Campus, Kushlan Wing, First Floor, Room 4

Credits: 3
Prerequisites: None
Office Hours: TBA

Course Description

This is an introductory course to spatial data in R, a free programming language and environment developed for statistical computing and graphics. The first part of the course will introduce you to the R programming language. The second part will focus on the R's ability to manipulate spatial data: including how to process vector layers, and rasters. Through hands-on exercises, this course introduces students to the concepts and applications of geographic information systems (GIS). It provides a basic foundation of spatial analysis and GIS with laboratory applications in particular techniques and methodology utilizing R. Students will learn to perform spatial analyses and communicate their results through cartography, along with introduction to concepts such as spatial data collection, remote sensing, and spatial statistics.

Note: You should have a computer either Windows or Mac with at least 8GB of RAM

Summary of Course Content

This introductory course on spatial data in R offers a comprehensive exploration of the R programming language's capabilities in statistical computing and graphics, with a specific focus on spatial data manipulation. The initial segment of the course acquaints students with the fundamentals of the R programming language, emphasizing its status as a versatile and free tool. Subsequently, the course delves into R's ability to handle spatial data, covering the processing of both vector layers and raster data. Participants learn to perform spatial analyses using R, with a particular emphasis on laboratory applications that apply techniques and methodologies to real-world scenarios. A critical component of the course is the integration of cartography, enabling students to visually communicate spatial analyses results through map creation. The practical applications extend to various domains, allowing students to develop a solid foundation in spatial analysis skills. Overall, this course equips

students with a robust skill set for handling spatial data, making it valuable for those entering fields such as geography, environmental science, and urban planning.

Learning Outcomes

Upon successful completion of this course the students will be able to:

- execute basic programming tasks in R (e.g. loops, conditional statements, while statements, etc.)
- understand basic GIS terms and concepts
- utilize GIS for mapping and conducting spatial analyses.
- create a professional website containing a portfolio

Assessment methods:

You will be graded on four problem sets during the semester (each 12.5% of your grade) and a final report and presentation (each 25% of your grade).

- 4 problem sets: 50% of the final grade (12.5% each)
- final presentation: 25% of the final grade
- final project report: 25% of the final grade

Problem Sets

1.**Initial Individual Submission**: This component contributes 50% of the overall grade for the problem set. When you first complete the problem set independently and submit it to the instructor, your grade for this component will be calculated based on the quality of your independent work. This grade will be weighted at 50% of the total assignment grade.

2.**Final Submission After Group Consultation**: This component also contributes 50% of the overall grade for the problem sets. After discussing the problem set with your group members and documenting the correct answers, you will submit this revised version individually. Note: no group submission is permitted. Each one of you has to submit the second attempt of the assignment individually. Your grade for this component will be based on the quality of your final submission after group consultation. This grade will also be weighted at 50% of the total assignment grade.

Final Project

Students will undertake a GIS project emphasizing the practical application of spatial analysis techniques using the R programming language. This project offers an opportunity to showcase the acquired skills in manipulating spatial data and conducting meaningful analyses. Participants are encouraged to choose a topic of interest or relevance, utilizing datasets provided in the course or exploring new ones. The project entails a few steps:

- choosing a topic that involves spatial data
- acquire data either from the course materials or from external sources
- employ at least five GIS procedures in R such as: `st_join`, `st_centroid`, `st_area`, `st_distance`, `st_buffer`, `st_voronoi`, `st_union`, `st_combine`, `st_cast`, `st_intersection`, `st_difference`, dplyr for vector layer aggregation, `st_crop`, `st_rasterize`, `raster::aggregate`, etc.

OR

- produce some descriptive visualizations (maps, barplot, scatterplot, boxplots) that tell the same story AND make a Github website (for data, you can explore https://ourworldindata.org)
- craft a well-structured two-page report containing an intro to the problem, objectives, data sources, methodology, results, and conclusion
- appendix with the R code for the GIS procedures.

Here are examples of projects that former students produced:

- Swiss Glaciers Melting Rate
- Rising Temperatures and Droughts

Examples of Project Topics:

1. Understanding economic development in Eastern Europe since 1992 using satellite luminosity

2. Exploring temperature changes: what are the countries in Europe that have experienced the most dramatic changes in temperature since 1901

3. Exploring temperature changes: what are the regions within a country (e.g. Italy) that have experienced the most dramatic changes in temperature since 1901

4. Exploring business opportunities (e.g. restaurants, retail) in your favorite city using OSM

5. Exploring areas worst affected by an earthquake.

Students have to submit a two-page report in Quarto and give a 20-minute, in-class presentation in Quarto.

Grading Criteria for the project:

- Relevance
- Methodology: demonstration of at least five distinct GIS procedures in R.
- Analysis
- Presentation
- Code Quality

You can find below an example of a project: **Wildfires in 2023**

- Project Memo | Project Memo QMD file
- Presentation | Presentation QMD file

Attendance Requirements

Students are required to attend classes following the University's policies. Students with more than four unexcused absences (two weeks) will get an F on the course. Thus, students must attend classes.

Academic Honesty

As stated in the university catalog, any student who commits an act of academic dishonesty will receive a failing grade on the work in which the dishonesty occurred. In addition, acts of academic dishonesty, irrespective of the weight of the assignment, may result in the student receiving a failing grade in the course. Instances of academic dishonesty will be reported to the Dean of Academic Affairs. A student who is reported twice for academic dishonesty is subject to summary dismissal

from the University. In such a case, the Academic Council will then make a recommendation to the President, who will make the final decision.

Students with Learning or Other Disabilities

John Cabot University does not discriminate on the basis of disability or handicap. Students with approved accommodations must inform their professors at the beginning of the term. Please see the website for the complete policy

Recommended Books

General Books on R

- Wickham, Hadley, and Garrett Grolemund. 2023. *R for Data Science.* O'Reilly Media. https://r4ds.hadley.nz.
- Wickham, Hadley. 2023. *ggplot2: Elegant Graphics for Data Analysis.* Springer. https://ggplot2-book.org/index.html.
- Nguyen, Mike. 2023. *A Guide on Data Analysis.* https://bookdown.org/mike/data_analysis/.

Spatial Analysis

- Pebesma, Edzer, and Roger Bivand. 2023. *Spatial Data Science: With Applications in R.* Chapman & Hall/CRC. https://r-spatial.org/book/.
- Gimond, Manuel. 2023. *Intro to GIS and Spatial Analysis.* https://mgimond.github.io/Spatial/index.html.
- Lovelace, Robin; Jakub Nowosad; and Jannes Muenchow. 2023. *Geocomputation with R.* https://r.geocompx.org.

Week 1

Class 1: Intro to R and GIS
01/15/2024 - Mon - Lecture

- Introduction and course overview
- What is R and GIS
- Programming in R

Class 2: Intro to R, Quarto, and R evironments
01/17/2024 - Wed - Lecture

- Installing R
- Quarto files: notebooks vs. presentations
- The R environment

Reading
Good Enough Practices in Scientific Computing, *Good Enough Practices in Scientific Computing.* PLoS Comput Biol 13(6): e1005510.
Wickham, Hadley and Grolemund, Garett, 2023. *R for Data Science. O'Riley.* https://r4ds.hadley.nz. C1-4

Week 2

Class 1: Classes, Vectors, Operations, Matrices, Dataframes
01/22/2024 - Mon - Lecture

- Numeric and string objects
- Vectors, Matrices, and Dataframes
- Conditionals

Class 2: Importing Data
01/24/2024 - Wed - Lecture

- Functions
- Loops
- the `%in%` operator

Reading
Wickham, Hadley and Grolemund, Garett, 2023. *R for Data Science. O'Riley.* C. Program https://r4ds.hadley.nz/program
Thulin, Måns. 2021. *Modern Statistics with R.* C6 R programming https://www.modernstatisticswithr.com/progchapter.html

Assignment 1

Week 3

Class 1: Reading Data and Working Directories
01/29/2024 - Mon - Lecture

- Importing Data into R
- Relative vs. Absolute Path
- Tidy Data
- Missing Values

Class 2: Relational Data, Dates and Joins
01/29/2024 - Wed - Lecture

- Joins
- Date objects

Reading
Wickham, Hadley and Grolemund, Garett, 2023. *R for Data Science. O'Riley.* C. Transform https://r4ds.hadley.nz/transform
Thulin, Måns. 2021. *Modern Statistics with R.* C. Exploratory data analysis and unsupervised learning. https://www.modernstatisticswithr.com/eda.html#outliers-and-missing-data

Week 4

Class 1: Problems
02/05/2024 - Mon - Exercises

- Functions

- Loops
- Joins
- Dates

Class 2: Data Visualization 1
02/07/2024 - Wed - Lecture

- `ggplot` Basics
- Simple Geoms vs. Complex Geoms
- Layers

Class 3: Data Visualization 2
02/09/2024 - Fri - Lecture 1 | Lecture 2

- Colors, Shapes, Transparency
- Facets
- Coordinates
- Annotations

Reading
Wickham, Hadley and Grolemund, Garett, 2023. *R for Data Science. O'Riley.* https://r4ds.hadley.nz. C5-8
Wickham, Hadley. *2023. ggplot2. Elegant Graphics for Data Analysis.* Springer. https://ggplot2-book.org/index.html

Week 5

Intro to Geographic Information Systems (GIS)
02/12/2024 - Mon | 02/14/2024 - Wed - Lecture

- The Cartesian System
- Coordinate Reference Systems
- Vectors: points, lines, polygons

Reading
Gimond, Manuel. 2023 *Intro to GIS and Spatial Analysis.* C9. Coordinate Systems. https://mgimond.github.io/Spatial/chp09_0.html

Assignment 2 | Assignment 2 Presentation Example | Assignment 2 Presentation Template

Week 6

Shapefiles, Coordinate Systems, and Basic Mapping
02/19/2024 - Mon - Lecture

- Vector file formats
- The `sf` package
- EPSG codes
- Creating point layers
- Reading Vector layers

Reading
Gimond, Manuel. 2023 *Intro to GIS and Spatial Analysis*. C3. GIS Data Management. https://mgimond.github.io/Spatial/gis-data-management.html

Spring Break 02/26/2024 - 03/03/2024

Week 7

Class 1: Fundamentals of Mapping
03/04/2024 - Mon - Lecture

- Mapping points, lines, polygons
- Choropleth Maps
- Colors Choice for Choropleth Maps
- Statistical Maps: quantile, interquantile, standard deviation, outlier maps

Class 2: Geometric Operations with Vectors 1
03/06/2024 - Wed -Lecture

- Join by Location
- Spatial Joins
- Geometric Calculations

Reading
Gimond, Manuel. 2023 *Intro to GIS and Spatial Analysis*. C7. Good Map Making Tips. https://mgimond.github.io/Spatial/good-map-making-tips.html#elements-of-a-map

Week 8

Geometric Operations with Vectors 2
03/11/2024 - Mon | 03/13/2024 - Wed - Lecture

- Distance
- Centroids
- Combining and Dissolving
- Buffers
- Vector Layer Aggregation

Reading
Lovelace Robin, Nowosad, Robin, and Muenchow, Jannes. 2023. *Geocomputation with R*. C. Geometry operations. https://r.geocompx.org/geometry-operations

Assignment 3

Week 9

Class 1: Spatial Operations - Applications
03/18/2024 - Mon - Lecture

- Intersections
- Unions
- Changing Coordinate Systems
- Open Street Maps (OSM)

Class 2: Rasters and Raster Operations 1
03/20/2024 - Wed - Lecture

- Reading Rasters
- Raster Attributes
- Satellite luminosity
- Raster Subsetting

Week 10

Rasters and Raster Operations 2
03/25/2024 - Mon | 03/27/2024 Wed - Lecture

- Cropping Raster to Shape
- Zonal statistics: mean and max
- Converting rasters (stars) to polygons
- Distance to Line

Reading
Gimond, Manuel. 2023 *Intro to GIS and Spatial Analysis*. C. Raster operations in R. https://mgimond.github.io/Spatial/raster-operations-in-r.html
Lovelace Robin, Nowosad, Robin, and Muenchow, Jannes. 2023. *Spatial data operations*. C. https://r.geocompx.org/spatial-operations

Assignment 4

Week 11

Class 1:
04/01/2024 - Mon - NO CLASS

Class 2: Rasters and Raster Operations 3
04/03/2024 - Wed - Lecture

- Temperature rasters (NC Data)
- Raster algebra: mean
- Writing rasters

Reading
Gimond, Manuel. 2023 *Intro to GIS and Spatial Analysis*. C. Raster operations in R. https://mgimond.github.io/Spatial/raster-operations-in-r.html
Lovelace Robin, Nowosad, Robin, and Muenchow, Jannes. 2023. *Spatial data operations*. C. https://r.geocompx.org/spatial-operations

Week 12

Class 1: Geometric Operations with Rasters 1
04/08/2024 - Mon - Lecture

- Grids and Fishnets
- Rasterizing grids

Class 2: Student Proposal Presentations
04/10/2024 - Wed

1. Question
2. Data sources
3. Proposed GIS functions
4. Expected policy implications

Week 13

Class 1: Combining Vectors and Rasters
04/15/2024 - Mon | 04/17/2024 Wed - Lecture

- Making grids
- Calculating zonal statistics
- Calculating number of points within a polygon
- Learning more about extracting Open Street Map data

Week 14

Class 1: Geometric Operations with Rasters 2
04/22/2024 - Mon - Lecture

- Mosaicing and Resampling
- Reprojecting
- Raster classification
- Raster cropping and masking

Class 2: Concluding Remarks
04/24/2024 - Wed - Lecture

- Overview of the Course
- Skills Acquired
- Jobs where these skills are valued

Supplementary Tutorials

Georeferencing Maps and Creating Shapefiles

- Digitizing and Georeferencing Maps
- Making a shapefile

Making a website on GitHub: `username.github.io`

- Making a website on Quarto
- Storing a website on GitHub
- Structuring a professional website for Data Analytics Jobs