DACSS 690C: Computational Social Science Methods (3 credits)

Prof. José Manuel Magallanes, PhD.

Spring 2025

E-mail: jmagallanesr@umass.edu
Modality: 100% Asynchronous

Course Website: Please visit *UMass Canvas*.
Personal assistance: Upon request

Course Description

This course reviews different computational social science methods that are applied under different academic and professional situations. This includes different but complimentary methods to format and explore data as tables, maps, and graphs. The course also includes basic methods for decision-making support such as optimization and social simulation. The course makes use of version control applications, as well as promoting the practice of reproducibility and transparency. The course makes use of Python and R, including NetLogo for social simulation, Gephi for network exploration, and Git for version control. Those languages are not a pre-requisite for the course, as they will be used as templates.

Prerequisite: *DACSS 601 Data Science Fundamentals* or equivalent (with instructor permission)

Course Objectives

1. Using different methods to explore different data structures.
2. Become familiar with the exploration of spatial data, graph/networks, text, multivariate data in tables.

3. Become familiar with optimization techniques (linear programming) and social simulation (agent-based modeling).
4. Learn how Python and R, as well as other tools, are used in this applications.
5. Follow a reproducible/transparent approach as a data scientist.

About Class format

The format for this class includes:

Course Objectives

1. This class is 100% asynchronous. No concurrent meetings.
2. Material for each class. Material will be organized as tutorials, so that students have concepts and examples to help them with homework.
3. The class is homework-based. Each session will present a homework.
4. Tutoring. Personal tutoring is offered to the students before the submission of each homework. If needed, synchronized will be offered individually.
5. Pregrading: Students may submit homewok for pregrading, that way they will know the grade that they would receive with the homework draft sent. The lecturer will give advice on how to improve towards 100%. Students not sending for pre-grade meeting, implicitly agree to receive the mark given without previous feedback.
6. Publishing: Generally, every homework will be submitted as a GitHub page.
7. Recordings: Each session will offer a video guide which is recorded for further use.
8. Late submission. Every homework will have a deadline to get 100 points. After that late, submissions will be accepted after 72 hours, with a penalty of 5 points for every 24 hours.

Course Book

There is no course book required. The lecturer will prepare ad-hoc reading material for each session. The lecturer will also recommend some readings for further exploration or reference.

Recommended Readings

Anselin, L. (1995). Local Indicators of Spatial Association—LISA. *Geographical Analysis*, 27(2):93–115.

BasuMallick, C. (2022). What Is Linear Programming? Meaning, Methods, and Examples.

Bates, C. (2020). Getting Started with R Markdown — Guide and Cheatsheet. https://www.dataquest.io/blog/r-markdown-guide-cheatsheet/.

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences*, 99(suppl_3):7280–7287.

Herries, J. (2018). Better Breaks Define Your Thematic Map's Purpose. https://www.esri.com/arcgis-blog/products/arcgis-online/mapping/better-breaks-define-your-thematic-maps-purpose/.

Jagoda, J. A., Schuldt, S. J., and Hoisington, A. J. (2020). What to Do? Let's Think It Through! Using the Analytic Hierarchy Process to Make Decisions. https://kids.frontiersin.org/articles/10.3389/frym.2020.00078.

Jayawickrama, T. D. (2021). Community Detection Algorithms. https://towardsdatascience.com/community-detection-algorithms-9bd8951e7dae.

Katija, S. (2021). Basic Operations in Git. https://medium.com/codex/basic-operations-in-git-72714ded4180.

Pandey, K. (2022). Graph Data Structure - Explained With Examples. https://www.masaischool.com/blog/graph-data-structure-explained-with-examples/.

Satopay, H. (2020). The Ultimate Markdown Guide (for Jupyter Notebook). https://medium.com/analytics-vidhya/the-ultimate-markdown-guide-for-jupyter-notebook-d5e5abf728fd.

Smith, H. (2020). Geographic vs Projected Coordinate Systems. https://www.esri.com/arcgis-blog/products/arcgis-pro/mapping/gcs_vs_pcs/.

Team, E. (2020). Agent-Based Modeling in NetLogo – Towards AI. https://towardsai.net/p/artificial-intelligence/agent-based-modeling-in-netlogo.

You may find helpful for data preprocessing, and visualization:

Magallanes Reyes, J. M. (2022). *Data Visualization for Social and Policy Research: A Step-by-Step Approach Using R and Python*. Cambridge University Press, 1 edition.

Magallanes Reyes, J. M. (2017). *Introduction to Data Science for Social and Policy Research: Collecting to Organizing Data with R and Python*. Cambridge University Press, Cambridge, United Kingdom New York, USA Port Melbourne, Australia Delhi, India Singapore.

Software installations required

Students have to install the following software in their computers:

- 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/
  A good alternative can be RStudio/Posit Cloud: **https://posit.cloud/**
- Python:
  https://www.anaconda.com/download
  Google Colab will be used too, which requires no installation.
- NetLogo:
  https://ccl.northwestern.edu/netlogo/6.4.0/
- GITHUB:
  Get an account at **https://github.com/**,
  and then download the desktop app from **https://github.com/**

Assessment and Grading

The grades are computed from homework. There are **EIGHT** assignments, two for each topic. You only need to complete **ONE** in each topic. The grade for each homework is computed considering:

1. The data is in a repo in GitHub (20 %).
2. The code is in a repo in GitHub (20 %).
3. The code is:
   - Explained in a tutorial-like fashion (20%).
   - Commented to explain a particular piece of code (20%).
4. The homework is published as a GitHub page (20%).

The grade will be an average of the grades obtained, and will be translated according to Table 1:

	A	B	C	F
+		87-89	77-79	0-72
	93-100	83-86	73-76	
−	90-92	80-82		

Table 1: Grade scale

Changes to the Syllabus

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

Working in groups

If you prefer, you can work in groups of 2, and present the deliverables as a group.

Course Support

Accommodations

The University of Massachusetts Amherst is committed to making reasonable, effective and appropriate accommodations to meet the needs of students with disabilities and help create a barrier free campus. If you are in need of accommodation for a documented disability, register with Disability Services to have an accommodation letter sent to your faculty. It is your responsibility to initiate these services and to communicate with faculty ahead of time to manage accommodations in a timely manner. For more information, consult the **Disability Services website**. .

Title IX

In accordance with Title IX of the Education Amendments of 1972 that prohibits gender-based discrimination in educational settings that receive federal funds, the University of Massachusetts Amherst is committed to providing a safe learning environment for all students, free from all forms of discrimination, including sexual assault, sexual harassment, domestic violence, dating violence, stalking, and retaliation. This includes interactions in person or online through digital platforms and social media. Title IX also protects against discrimination on the basis of pregnancy, childbirth, false pregnancy, miscarriage, abortion, or related conditions, including recovery. There are resources here on campus to support you. A summary of the available Title IX resources (confidential and non-confidential) can be found at the following link: **https://www.umass.edu/titleix/resources**. You do not need to make a formal report to access them. If you need immediate support, you are not alone. Free and confidential support is available 24 hours a day / 7 days a week / 365 days a year at the SASA Hotline 413-545-0800.

Academic Honesty

Since the integrity of the academic enterprise of any institution of higher education requires honesty in scholarship and research, academic honesty is required of all students at the University of Massachusetts Amherst. Academic dishonesty is prohibited in all programs of the University. Academic dishonesty includes but is not limited to: cheating, fabrication, plagiarism, and facilitating dishonesty. Appropriate sanctions may be imposed on any student who has committed an act of academic dishonesty. Instructors should take reasonable steps to address academic misconduct. Any person who has reason to believe that a student has committed academic dishonesty should bring such information to the attention of the appropriate course instructor as

soon as possible. Instances of academic dishonesty not related to a specific course should be brought to the attention of the appropriate department Head or Chair. The procedures outlined below are intended to provide an efficient and orderly process by which action may be taken if it appears that academic dishonesty has occurred and by which students may appeal such actions. Since students are expected to be familiar with this policy and the commonly accepted standards of academic integrity, ignorance of such standards is not normally sufficient evidence of lack of intent. For more information about what constitutes academic dishonesty, please see the Dean of Student 's website: **http://umass.edu/dean_students/codeofconduct/acadhonesty/**

Course Schedule

Week 00, January/27 - January/31: Guidelines for the course

HOMEWORK 1: Spatial Data

Week 01, February/03 - February/07: Intro to Spatial Data

- The spatial data files.
- Formatting the GeoDataframe.
  Readings:Smith (2020).
- **NO homework offered**.

Week 02, February/10 - February/14: Exploring Spatial Data

- Producing a Choropleth.
- Normalizing and Discretizing data.
  Readings:Herries (2018)
- **Homework 1 option #1**

Week 03, February/17 - February/21: Spatial Analytics

- Spatial operations on data frames.
- Global and Local Spatial autocorrelation.
  Readings:Anselin (1995)
- **Homework 1 option #2**

Week 04, February/24 - February/28: Pre-grading of homework 1

HOMEWORK 2: Data on graphs (networks)

Week 05, March/03 - March/07: Intro to Graph Data

- Formatting social data into networks.
  Readings:Pandey (2022)
- **NO homework offered**.

Week 06, March/10 - March/14: Exploring social networks

- Exploration of network structure.
- Network visualization.
  Readings:Disney (2020).
- **Homework 2 option #1**

Week 07, March/17 - March/21: Communities in networks

- Community detection.
  Readings:Jayawickrama (2021).
- **Homework 2 option #2**

Week 08, March/24 - March/28: Pre-grading of homework 2

HOMEWORK 3: Optimization

Week 09, March/31 - April/04: Linear programming

- Organizing information for Linear programming.
- Linear programming.
  Readings:BasuMallick (2022).
- **Homework 3 option #1**

Week 10, April/07 - April/11: Decision Making

- The Analytics Hierarchical Process (AHP).
  Readings:Jagoda et al. (2020)
- Organizing information for AHP.
- **Homework 3 option #2**

Week 11, April/14 - April/18: Pre-grading of homework 3

HOMEWORK 4: Social Simulation

Week 12, April/21 - April/25: Social Simulation I

- Principles of agent-based modeling.
  Readings:Bonabeau (2002)
- Basic use of NetLogo.
  Readings:Team (2020)
- **Homework 4 option #1**

Week 13, April/28 - May/02: Social Simulation II

- Scenarios in NetLogo.
- **Homework 4 option #2**

Week 14, May/05 - May/09: Pre-grading of homework 4

Week 15, May/12 - May/16: Final Grading.