Data and Network Science for Public Policy

Timur Naushirvanov
naushirvanov_timur@phd.ceu.edu

Fall/Winter, 2024-25

Course Description

In the era where Big Data has taken center stage, influencing our daily lives, its application in Public Policy is still emerging. This course mirrors the current trend by exploring the realm of applying Data and Network Science for Public Policy objectives, actively contributing to the surge in popularity and adoption of Data Science methodologies.

Welcome to an engaging and introductory journey! This multidisciplinary course is designed for CEU undergraduate students keen on discovering the dynamic applications of Data and Network Science within the realm of Public Policy. Over 12 weeks, the course unfolds in three different parts. First, students learn or review the fundamentals of Public Policy, Data and Network Science, and Artificial Intelligence. Next, the spotlight turns to designing and implementing a Data Science Project tailored to address real-world public policy challenges. Finally, students explore the variety of applications of Data and Network Science usage cases in different Public Policy domains, with topics spanning from Health and Education to Energy and Environment issues.

Upon completing the course, students will design their own Data and Network Science project, tackling a public policy issue of their choice.

My Background and Teaching Philosophy

My academic background spans across various disciplines, encompassing Political Science, Public Policy, and Data Science. This diverse foundation motivates my desire to teach a course that integrates these interdisciplinary areas, shaping the core of my teaching philosophy.

In my teaching approach, I embrace two educational philosophies: student-centered teaching and liberalism. Student-centered teaching focuses on putting students at the center of the learning experience by creating a safe and supportive environment where every student can thrive individually. Liberalism underscores the importance of cultivating well-rounded individuals with diverse and multidisciplinary knowledge and skills.

During the class, I introduce all the key learning concepts, and I design various group and individual activities to facilitate students' comprehension of these concepts. The interdisciplinary nature of this lesson aligns with the second facet of my teaching philosophy, as it assists students

from various backgrounds in gaining additional knowledge and skills from related disciplines such as data science, network science, political science, and public policy.

Prerequisites - Relation to other Bachelor courses

The course is tailored for undergraduate students across various CEU Undergraduate programs. Within the Data Science and Society (DSS) program, it integrates skills from three key modules: Programming and Data Analysis, Data Science, and Social Sciences. In the context of the Culture, Politics and Society (CPS) program, it is particularly relevant for majors in Political, Legal, and Governmental Studies, as well as Social and Environmental Studies. For students in the Philosophy, Politics, and Economics (PPE) program, the course delves into topics related to Economics and Public Policy.

Given the course's focus on applying Data and Network Science to address public policy issues, prospective participants are recommended to have completed the Fundamentals of Data Analysis course first and have good knowledge of Python programming language. While it is advisable to have prior exposure to Introduction to Machine Learning and Data Mining, and/or Introduction to Public Policy, such prerequisites are not mandatory for enrollment.

Course Objectives

The purpose of this course is to:

1. Build Comprehensive Understanding: By the end of the initial weeks, students will identify and analyse foundational concepts in Public Policy, Data Science, and Artificial Intelligence, establishing a solid groundwork for informed decision-making in public policy through data-driven insights.

2. Apply Data Science Techniques: Students will apply their knowledge to design and implement a Data Science Project addressing real-world public policy challenges, evaluating project outcomes critically and completing the implementation within a specified timeframe.

3. Explore Diverse Applications: Throughout the course, students will identify and analyse patterns in various Data and Network Science applications across public policy domains, evaluating and judging the effectiveness of these applications through case studies.

Learning Outcomes

By the end of this course, students should be able to:

1. demonstrate in-depth understanding by identifying and analysing intricate patterns in Data Science concepts, achieving high scores in assessments and establishing a solid foundation for further exploration and application in practical scenarios;

2. design and implement a Data Science Project, critically evaluating outcomes and cultivating skills for translating theoretical concepts into actual policies;

3. identify and analyse interdisciplinary connections between data science, network science, political science, and public policy, demonstrating how these connections contribute to critical analysis and problem-solving in diverse public policy scenarios;

4. refine their skills in evaluating and judging the effectiveness of various Data and Network Science applications in public policy domains, showcasing a deep understanding of patterns and outcomes through rigorous case studies and assessments.

Course Structure

Class Structure

Depending on the weekly topic, most classes consist of short lectures that introduce the topic and learning goals (2 lectures 20 minutes each per class), small quizzes and interactive tasks (including case discussions), and group discussions. During practical coding sessions, the lecturer additionally presents coding notebooks. During the final lesson, students present their group projects related to applying Data and Network Science to Public Policy issues.

Assessments

Attendance and Class Participation (20 points): as there are many in-class quizzes and discussions, they all assist in understanding the students' level of interest and motivation, as well as identifying their weak points.

In-Class Topic/Paper Presentation (20 points): each student has an opportunity to present the case of interest during Part III: Domain Areas of Data Science in Public Policy.

Final Project Presentation (40 points): during the final week, students present the group projects that they have been developing during the semester.

Final Project Policy Paper (20 points): after the final presentation, students write individual policy papers (1400 words) summarising the policy problem they presented and how their project helps resolve it.

For the final course grade, the CEU Grading system is used:

Grade	Points (0-4 scale)	Points (0-100 scale)
A	3.68 - 4.00	100-91
A-	3.34 - 3.67	90-81
B+	3.01 - 3.33	80-71
B	2.68 - 3.00	70-63
B-	2.34 - 2.67	62-58
C+	2.33 - (minimum pass)	57-53

Final Project Presentation

During the concluding week of the semester, students showcase the group projects they have developed over the term. This assignment is a platform for fostering teamwork and enhancing practical skill sets. Students apply the knowledge acquired from lectures and practical tutorials to develop their final projects. Project topics, aligned with the application of Data and Network Science in Public Policy, will be collectively discussed in class. The projects are expected to encompass the full spectrum of data science stages: from data collection and processing to analysis, modelling, and visualisation. Each group member should participate in presenting the results. The students form groups on their own.

Grading Schema

Criterion	Fully Completed (5 points)	Partially Completed (3 points)	Needs Improvement (1 point)
Project Relevance to Public Policy	Demonstrates a clear and strong connection between the project topic and its application in public policy.	Shows some relevance to public policy, but lacks depth or clarity in connection.	The connection between the project topic and its relevance to public policy is unclear or absent.
Data Source and Data Collection	Thoroughly identifies relevant data sources and effectively collects data, demonstrating a comprehensive understanding of data collection methods.	Identifies and collects data adequately, but may lack completeness or clarity in source selection or data gathering methods.	Data source identification or collection methods are incomplete or lack clarity, requiring further development.
Data Processing and Preparation	Conducts comprehensive data processing and preparation, demonstrating proficiency in data cleaning, integration, and transformation.	Processes and prepares data adequately, but may lack completeness or effectiveness in cleaning, integration, or transformation.	Data processing and preparation are incomplete or lack effectiveness, requiring additional work for proper analysis.
Data Analysis and Modelling	Conducts thorough data analysis and modelling, providing insightful interpretations and effective utilisation of appropriate techniques and algorithms.	Performs data analysis and modelling to some extent, but lacks depth or effectiveness in interpretation or technique utilisation.	Data analysis and modelling are rudimentary or incomplete, requiring further development for meaningful insights.
Ethical Considerations	Carefully evaluates the ethical implications of the project, demonstrating a thoughtful approach to potential ethical issues and their resolution.	Considers ethical implications to some extent but may lack thoroughness or depth in evaluation.	Ethical considerations are overlooked or insufficiently addressed, requiring a more comprehensive examination.
Presentation Quality	Presents final project in a clear, organised, and engaging manner, effectively communicating key findings and insights.	Presents final project adequately, but may lack coherence or engagement in delivery.	Presentation is unclear, disorganised, or lacks engagement, hindering effective communication of findings.

Feedback for the Final Project Presentation

Feedback consists of two parts:

1. Rubric Assessment is shared with the students.
2. After seeing the result, each student needs to complete a short individual self-evaluation report. In this self-evaluation, each member should assess their contributions to the project, reflecting on their strengths, areas for improvement, and the extent of their involvement in each criterion.

30/40 points are based on Rubric (meaning all group members receive the same amount of points) and 10/40 points are based on Individual Self-Assessment.

Schedule and weekly learning goals

Part I: Introductions

Week 01 Introduction to the Course and Public Policy

Learning Outcomes	Assessment/ Assignment	Teaching/ Learning	Prepare/ Engage/ Consolidate
Form an understanding of the course and its relevance for the students. Discuss the foundations of the policy-making process and its evolution. Identify and analyse current trends in using Big Data for Public Policy.	In-class Group Work with a short Presentation on a lecture's topic (pass/fail)	Individual introductions Lecture combined with guided discussions of the topic (including polls) Finishing with a group work and brief presentations on current trends	**Prepare:** Students should think about their motivation and interest to join the course. **Engage:** In-class discussion of these ideas and motivations. **Consolidate:** After the session, students start developing their projects and ideas.

Mandatory Resources:

- Moody, R., Bekkers, V. (2023). Introduction. In: Big Data and Public Policy. Palgrave Macmillan, Cham.
- Moody, R., Bekkers, V. (2023). Public Policymaking. In: Big Data and Public Policy. Palgrave Macmillan, Cham.

Optional Resources:

- Högberg, C. (2023). Big Data: Consulting the Public in Public Policy. In M. van Gerven, C. Rothmayr, & K. Schubert (Eds.), Encyclopedia of Public Policy Springer.
- Högberg, C. (2023). Big data: Automated Decision-Making in Public Policy. In M. van Gerven, C. Rothmayr, & K. Schubert (Eds.), Encyclopedia of Public Policy Springer.

Week 02 Introduction to Data and Network Science

Learning Outcomes	Assessment/ Assignment	Teaching/ Learning	Prepare/ Engage/ Consolidate
Understand the foundations of Data Science and Networks as its relevant subfield. Assess the relevance of Data and Network Science for the policy-making objectives. Start identifying and evaluating potential data sources relevant to Public Policy research.	In-class Group Work with a short Presentation on a lecture's topic (pass/fail)	Lecture combined with guided discussions of the topic (including polls) Finishing with a group work and brief presentations on datasets identification	**Prepare:** Students should familiarise with the literature. **Engage:** In-class discussions. **Consolidate:** Students start thinking about development of their own projects based on identified data sources.

Mandatory Resources:

- Barabási, A.-L., Pósfai, M. (2016). Network science. Cambridge: Cambridge University Press. ISBN: 9781107076266 1107076269 – Chapter 1 http://networksciencebook.com/chapter/1
- What is Big Data? https://www.guru99.com/what-is-big-data.html
- Data Science In 5 Minutes | Data Science For Beginners | What Is Data Science? | Simplilearn: https://www.youtube.com/watch?v=X3paOmcrTjQ

Optional Resources:

- The most important skills of data scientists | Jose Miguel Cansado | TEDxIEMadrid: https://www.youtube.com/watch?v=qrhRfPY4F4w
- Moody, R., Bekkers, V. (2023). Big Data. In: Big Data and Public Policy. Palgrave Macmillan, Cham.
- Moody, R., Bekkers, V. (2023). Big Data and Public Policymaking. In: Big Data and Public Policy. Palgrave Macmillan, Cham.

Week 03 Introduction to Artificial Intelligence and Remote Sensing

Learning Outcomes	Assessment/ Assignment	Teaching/ Learning	Prepare/ Engage/ Consolidate
Understand the foundations of Artificial Intelligence for Public Policy purposes. Formulate policy questions and design Data Science and AI interventions that can help answer these questions. Continue identifying and comparing different data sources relevant to Public Policy Research.	In-class Group Work with a short Presentation on a lecture's topic (pass/fail) In-class test checking the understanding of materials of the first three lectures (Pass / Please Revise)	Lecture combined with guided discussions of the topic (including polls) Group work on formulating relevant policy questions and designing Data Science interventions Finishing with 15-minute test summarising materials of introductory lectures.	**Prepare:** Students should familiarise with the literature. **Engage:** In-class discussions. **Consolidate:** Students continue thinking about development of their own projects based on identified data sources.

Mandatory Resources:

- Introduction To Artificial Intelligence | What Is AI?| Artificial Intelligence Tutorial |Simplilearn: https://www.youtube.com/watch?v=SSE4M0gcmvE
- Janga, B., Asamani, G. P., Sun, Z., & Cristea, N. (2023). A Review of Practical AI for Remote Sensing in Earth Sciences. Remote Sensing, 15(16), 4112.

Optional Resources:

- Flasiński, M. (2016). Introduction to artificial intelligence. Springer.

Part II: Implementing a Data Science Project

Week 04 Accessing and Cleaning Data

Learning Outcomes	Assessment/ Assignment	Teaching/ Learning	Prepare/ Engage/ Consolidate
Identify the topic of interest for each student. Discuss the most relevant datasets for each project and how to access them. Practice Data Cleaning in Python.	Students have to bring their initial ideas about their own projects (ungraded) *Homework:* Coding Game on Processing Data	Individual discussions of a topic of interest 60-minute lecture on Accessing and Cleaning Data combined with a coding in-class tutorial	**Prepare:** Student should complete Python Review sheet before the session. **Engage:** During the session, everyone will be practicing data preprocessing. **Consolidate:** After the session, students will be able to practice new knowledge in a coding game.

Optional Resources:

- Géron, A. (2022). Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow. " O'Reilly Media, Inc.". – Chapter 2
- Müller, A. C., Guido, S. (2016). Introduction to machine learning with Python: a guide for data scientists. " O'Reilly Media, Inc.". – Chapter 4

Week 05 Exploratory Data Analysis (EDA)

Learning Outcomes	Assessment/ Assignment	Teaching/ Learning	Prepare/ Engage/ Consolidate
Understand key EDA goals and techniques. Discuss the role of good data visualisations in exploring data. Apply Exploratory Data Analysis in Python.	In-class Group Work: Doing coding tutorial together *Homework:* Coding Game on Exploratory Data Analysis	60-minute lecture on Exploratory Data Analysis combined with a coding in-class tutorial 40-minutes group work on implementing Exploratory Data Analysis	**Prepare:** Student should complete Python EDA sheet before the session. **Engage:** During the session, everyone will be practicing EDA. **Consolidate:** After the session, students will be able to practice new knowledge in a coding game.

Optional Resources:

- Müller, A. C., Guido, S. (2016). Introduction to machine learning with Python: a guide for data scientists. " O'Reilly Media, Inc.". – Chapter 3
- Data Presentation, Step-by-Step | Google Data Analytics Certificate: https://www.youtube.com/watch?v=CzrCADdsXwE

Week 06 Modelling and Model Assessment

Learning Outcomes	**Assessment/ Assignment**	**Teaching/ Learning**	**Prepare/ Engage/ Consolidate**
Learn the foundations of data modelling, prediction, and its assessment. Experiment with interpreting and communicating the results of a Data Science Project. Practice Modelling and Model Assessment in Python.	In-class Group Work: Doing coding tutorial together *Homework:* Coding Game on Modelling and Model Assessment	60-minute lecture on Modelling and Model Assessment combined with a coding in-class tutorial 40-minutes group work on implementing Modelling and Model Assessment	**Prepare:** Student should complete Python Modelling sheet before the session. **Engage:** During the session, everyone will be practicing Python Modelling and Model Assessment. **Consolidate:** After the session, students will be able to practice new knowledge in a coding game.

Optional Resources:

- Géron, A. (2022). Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow. " O'Reilly Media, Inc.". – Chapters 3, 4, 9
- Müller, A. C., Guido, S. (2016). Introduction to machine learning with Python: a guide for data scientists. " O'Reilly Media, Inc.". – Chapters 2, 3

Part III: Domain Areas of Data Science in Public Policy

Week 07 Data and Network Science for Social Policy (Health and Education)

Learning Outcomes	Assessment/ Assignment	Teaching/ Learning	Prepare/ Engage/ Consolidate
Explore the application of Data and Network Science in addressing challenges related to social policy, specifically in the domains of Health and Education. Compare and analyse different cases of using Data and Network Science for Social Policy. Gain hands-on experience using Python to implement Data and Network Science techniques for Social Policy analysis.	In-class coding test (graded on a scale 0-100) Student presentations	25-minutes in-class coding test on summarising the ideas of lectures 4-6 10-minute mini-lecture 40-minute student presentations of 3 cases with their discussion (assigned students are supposed not only present their cases but to discuss and compare it between each other in advance) 20-minute guided discussions with some coding examples	**Prepare:** Critically read mandatory literature, some students are supposed to prepare case presentations. **Engage:** While some students giving presentations, others are encouraged to ask questions and discuss strong and weak sides of presented cases. **Consolidate:** After the session, students can integrate discussed ideas into their projects.

Mandatory Resources:

- Data Science in Education:
  https://www.discoverdatascience.org/articles/data-science-in-education/
- Data Science and Healthcare: The Impact on Medicine:
  https://datascientest.com/en/data-science-and-healthcare-the-impact-on-medicine

Optional Resources:

- Network science: translating big data into precision medicine:
  https://www.youtube.com/watch?v=IYW3N7Slqn0
- Using network science to forecast the spread of emerging diseases:
  https://www.youtube.com/watch?v=5uEBl2x78SI

Week 08 Data and Network Science for Economic Development

Learning Outcomes	Assessment/ Assignment	Teaching/ Learning	Prepare/ Engage/ Consolidate
Understand the role of Data and Network Science in fostering economic development. Compare and analyse different cases of using Data and Network Science for Economic Development. Evaluate economic datasets using modelling techniques to derive meaningful conclusions and recommendations.	Student presentations	2*20-minute mini-lecture 40-minute student presentations of 2-3 cases with their discussion (assigned students are supposed not only present their cases but to discuss and compare it between each other in advance) 20-minute guided discussions with some coding examples	**Prepare:** Critically read mandatory literature, some students are supposed to prepare case presentations. **Engage:** While some students giving presentations, others are encouraged to ask questions and discuss strong and weak sides of presented cases. **Consolidate:** After the session, students can integrate discussed ideas into their projects.

Mandatory Resources:

- Data Science in the New Economy: A new race for talent in the Fourth Industrial Revolution | WEF: https://www3.weforum.org/docs/WEF_Data_Science_In_the_New_Economy.pdf
- Data Science for Development: https://cega.berkeley.edu/theme/data-science-for-development/

Optional Resources:

- Emmert-Streib, F., Tripathi, S., Yli-Harja, O., & Dehmer, M. (2018). Understanding the world economy in terms of networks: a survey of data-based network science approaches on economic networks. Frontiers in Applied Mathematics and Statistics, 4, 37.

Week 09 Data and Network Science for Disaster Management

Learning Outcomes	Assessment/ Assignment	Teaching/ Learning	Prepare/ Engage/ Consolidate
Examine how Data and Network Science can be leveraged for effective disaster management and response. Compare and analyse different cases of using Data and Network Science for Disaster Management. Assess and interpret data patterns to enhance disaster preparedness and mitigation strategies.	In-class Group Work with a short Presentation on a lecture's topic (pass/fail)	2*20-minute mini-lecture combined with guided discussions of the topic (including polls) 40-minute student presentations of 2-3 cases with their discussion (assigned students are supposed not only present their cases but to discuss and compare it between each other in advance) 20-minute guided discussions with some coding examples	**Prepare:** Critically read mandatory literature, some students are supposed to prepare case presentations. **Engage:** While some students giving presentations, others are encouraged to ask questions and discuss strong and weak sides of presented cases. **Consolidate:** After the session, students can integrate discussed ideas into their projects.

Mandatory Resources:

- Cao, L. (2023). AI and data science for smart emergency, crisis and disaster resilience. International journal of data science and analytics, 15(3), 231-246.

Optional Resources:

- Ghaffarian, S., Taghikhah, F. R., & Maier, H. R. (2023). Explainable artificial intelligence in disaster risk management: Achievements and prospective futures. International Journal of Disaster Risk Reduction, 98, 104123.
- Webinar on Digital Technologies for Disaster Risk Management: https://www.youtube.com/watch?v=Zq2vkXgYJzM

Week 10 Data and Network Science for Energy and Environment

Learning Outcomes	Assessment/ Assignment	Teaching/ Learning	Prepare/ Engage/ Consolidate
Explore the intersection of Data and Network Science with issues in Energy and Environment policy. Compare and analyse different cases of using Data and Network Science for Energy and Environment. Evaluate environmental datasets using modelling tools to inform sustainable policies.	Student presentations	2*20-minute mini-lecture combined with guided discussions of the topic (including polls) 40-minute student presentations of 2-3 cases with their discussion (assigned students are supposed not only present their cases but to discuss and compare it between each other in advance) 20-minute guided discussions with some coding examples	**Prepare:** Critically read mandatory literature, some students are supposed to prepare case presentations. **Engage:** While some students giving presentations, others are encouraged to ask questions and discuss strong and weak sides of presented cases. **Consolidate:** After the session, students can integrate discussed ideas into their projects.

Mandatory Resources:

- Holme, P., Rocha, J. C. (2023). Networks of climate change: connecting causes and consequences. Applied Network Science, 8(1), 1-20.
- AI in Renewable Energy: How Is It a Game Changer? https://www.youtube.com/watch?v=pghjLyAmc5g

Optional Resources:

- McKenna, E., Higginson, S., Hargreaves, T., Chilvers, J., Thomson, M. (2020). When activities connect: Sequencing, network analysis, and energy demand modelling in the United Kingdom. Energy Research Social Science, 69, 101572.

Part IV: Course Summary

Week 11 Ethics in Data Science, AI, and Public Policy

Learning Outcomes	Assessment/ Assignment	Teaching/ Learning	Prepare/ Engage/ Consolidate
Identify and analyse the benefits and threats of using Data Science and AI for Public Policy. Compare ethical components of various cases of Data Science and AI interventions. Evaluate own projects through the ethical perspective.	In-class Group Work with a short Presentation on a lecture's topic (pass/fail)	40-minute lecture combined with guided discussions of the topic (including polls) 60-minute group work on ethical cases	**Prepare:** Critically read mandatory literature and reflect on the relevance of ethical considerations for your own projects. **Engage:** Participate in discussions sharing your opinions and relevant examples from your projects. **Consolidate:** After the session, students can integrate discussed ideas into their projects.

Mandatory Resources:

- The ethics of AI development in public policy | By Dr Emma Carmel: https://www.openaccessgovernment.org/ethics-of-ai-development/84472/
- Why develop a data science code of ethics? https://www.youtube.com/watch?v=s8qjmImu1LQ

Optional Resources:

- Artificial intelligence: From ethics to policy | European Parliament: https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641507/EPRS_STU(2020)641507_EN.pdf

Week 12 Conclusions and Student Presentations

Learning Outcomes	Assessment/ Assignment	Teaching/ Learning	Prepare/ Engage/ Consolidate
Present student policy projects based on the application of concepts and techniques learned. Evaluate student policy projects and their outcomes: group discussions, comparison of ideas and recommendations for further improvement. Reflect on course results and provide final feedback (both from the lecturer and students).	Group Project Presentations (graded 0-100) Students are supposed to submit individual policy papers based on their projects (graded 0-100)	80-minute – group project presentations 20-minute – mini-lecture (key takeaways from the course)	**Prepare:** Students have to prepare a group project **Engage:** Project presentation during the session **Consolidate:** Students write an individual policy paper based on a received feedback

Optional Resources:

- 8 Tips for Creating a Compelling Presentation for Data Science: https://dataknowsall.com/blog/storytelling.html
- A Blueprint for Data Science Presentations | Medium: https://towardsdatascience.com/a-blueprint-for-data-science-presentations-9c79e86bdf0c