environmental & urban change | YORK

COURSE SYLLABUS

ENVS 6182 "Environmental Analytics: Data, Models, and Methods" 3.0
Winter 2024

Official course description

"The application of analytics including optimization, simulation, regression, and time series analysis, to problems in environmental studies such as food systems, political change, emergency response systems, and homeless shelter policy. Solutions will be implemented in spreadsheets and statistical software (Excel and R)."

This term we will be using Python, rather than R.

Pre-requisites

This course is for students who wish to use quantitative methods in their Major Research. No background is assumed, but a strong interest in quantitative methods is expected.

Course director

Dr Lina Brand Correa
Email: brand@yorku.ca
Office: HNE(UC) 273

Office hours: Tuesdays 10:00-11:00, HNE 273 or by appointment (arranged by email)

Times and locations

The course will be delivered in person, as long as public health recommendations continue to allow for it. Please review the guidelines at York University's site Better Together to keep everyone safe during our in-person interactions.

Lectures and practicals: Thursdays 14:30-17:30, HNE 253.

If remotely, you will be able to find Zoom links on eClass.

Course webpage

eClass: eclass.yorku.ca/

Please visit the Student Guide to eClass to familiarise yourself with the system.

Software requirements

For this course we will be using 2 main software packages:

- Microsoft office (Excel): this is available to any York University student and you probably already have it on your personal computer. Please visit the Computing for Students website for more information.
- Python: this is an open-source software. The computer lab where we will be working will have the Anaconda ecosystem installed in all its PCs. Anaconda, and within it Jupyter Lab, allows to

easily code in Python. If you wish to, you can also download Anaconda to your personal computer (https://www.anaconda.com/products/individual).

If you have any questions or require support, please get in touch with the course director.

Expanded course description (objectives and learning outcomes)

The field of analytics is about getting insight from data. Environmental is about getting insight from environmental data. To choose the right technique and implement it, it is necessary to be exposed to a range of techniques and develop a familiarity in the field.

Specific course learning outcomes are:

- Name a range of methods – statistical methods and models
- Identify variables, research questions, and what constitutes data that could answer a question
- Formulate problems from descriptions
- Develop familiarity with spreadsheets and with Python software that are used in solving problems

Problem-based learning will be used.

Future Career Skills

- Familiarity with the Python software environment
- Ability to search for, download and process secondary quantitative data
- Data analysis and visualisation
- Scenario development
- Ability to work independently in formulating and testing hypotheses based on the main theories of environment-economy linkages
- Critical thinking
- Improved presentation and public speaking skills
- Improved written communication skills
- Writing succinct report of findings

Organization of the course

This course is centred around problem-based learning, where students will be able to identify a real-world environmental problem, learn some analytical tools to understand them better and propose solutions. Given that students will be constantly learning data analytics tools, attendance is required. Furthermore, active participation in seminar discussions, as well of use of class time for undertaking the practical exercises will be key to achieving the learning outcomes of the course.

Each class will have 3 basic components (open to flexibility). First, a discussion of the reading(s), homework, presentation or real-world data (as relevant), either led by the course director or by the students themselves. Second, a short presentation of the week's topic by the course director. Third, a practical session where students apply what they've learnt using real-world data. The table below presents a summary of the organization of the course, with more details provided on eClass each week.

Item	Week	Topic	Date
Class 1	1 (w/c 08 Jan)	Introduction to module & coding/programming	Thursday, 11 Jan

Basics of Data Analytics			
Class 2	2 (w/c 15 Jan)	Quantitative methods – what, why, when and how Library session on data searching	Thursday, 18 Jan
Class 3	3 (w/c 22 Jan)	Data – how to organize and clean it	Thursday, 25 Jan
Homework 1: Data for final report			*Bring to class 3*
Class 4	4 (w/c 29 Jan)	Basic Python functions	Thursday, 1 Feb
Homework 2: Python Workbook 1			*Bring to class 4*
Class 5	5 (w/c 05 Feb)	Descriptive statistics and data visualization	Thursday, 1 Feb
Discourses of climate delay (presentations)			*Thursday, 1 Feb*
Class 6	6 (w/c 12 Feb)	Linear Regression (single and multi variate)	Thursday, 15 Feb
Discourses of climate delay (presentations)			*Thursday, 15 Feb*
Reading week (19-23 February)			
Class 7	7 (w/c 26 Feb)	Logistic Regression (single and multi variate)	Thursday, 29 Feb
Discourses of climate delay (presentations)			*Thursday, 29 Feb*
Social Metabolism			
Class 8	8 (w/c 04 Mar)	Social metabolism and cross-country analysis	Thursday, 07 Mar
Interim report on data (Friday March 7)			
Class 9	9 (w/c 11 Mar)	IPAT, KAYA and STIRPAT	Thursday, 14 Mar
Input-Output Analysis			
Class 10	10 (w/c 18 Mar)	Multi-regional input-output analysis with environmental extensions	Thursday, 21 Mar
Knowledge Consolidation and Final Report Consultation			
Class 11	11 (w/c 25 Mar)	Final report consultation	Thursday, 28 Mar
Class 12	12 (w/c 01 Apr)	Course conclusion: where to go next with environmental analytics	Thursday, 04 Apr
Final report (Thursday April 18)			

Course readings

This course draws mainly on academic journal articles as reading materials. These are specified in the detailed schedule of readings below, and are available through York University Library and linked on eClass. The course director might add to the list below as the course progresses.

This course also draws several sources for the basics of coding in Python:

- Whitington. 2019. *Python from the Very Beginning. An Introduction*: https://www.pythonfromtheverybeginning.com/
- Guide of Python courses: https://wiki.python.org/moin/BeginnersGuide/Programmers

Evaluation

The grade for the course will be based on the following items weighted as indicated:

Participation	15%	During all classes
Homework	10%	Due weeks 3 and 4
Discourses of climate delay (presentation)	20%	Due during class in weeks 5-7
Interim report on data	20%	Due on March 7

Final report	35%	Due on April 18

Brief description of each assessment activity

Below is a brief description of the assessment activities. Further details and expectations will be provided in class and over eClass.

Homework

The homework will get you started thinking about data and about Python coding during weeks 2 and 3. Students should complete the homework and be ready to discuss it during class.

Due date

During class in weeks 3 and 4.

Value (% of final grade)

10% (5% each homework)

Participation

This component is assessed based on class attendance, engagement with classmates' presentation and with in-class discussions.

Due date

During class in weeks 1-12.

Value (% of final grade)

15%

Discourses of climate delay

Each student will search a *popular press article* that makes an argument based on a "discourse of climate delay" (the typology we explored in Week 1). You can search for these types of articles in all major newspaper sites, as well in weekly or monthly magazines. Students will prepare a 15 minute presentation covering:

- A summary of the article
- An explanation of what elements of that article constitute discourses of climate delay (one or more) and an assessment of it's effectiveness (is the argument convincing?)
- An analysis of how data contributes to making the climate delay argument
- A presentation of additional data or research that debunks the climate delay argument.

Due date

During class in weeks 5-7.

Value (% of final grade)

20%

Interim report on data

Each student will search for and download one (or several) dataset(s) they are planning on analysing for their final report. In 2-pages (excluding tables and graphs), students will describe the main characteristics of the dataset(s), describe the origins of the dataset(s), critically analyze the strengths and limitations of the data, propose a research question to address in your final report and assess the suitability of the dataset(s) to answer it.

Due date

By 5pm, March 7.

Value (% of final grade)
20%

Final report

Each student will choose a topic or country they are interested in, find relevant data, and perform any of the data analysis techniques learnt during the course. Then students will produce a 2000-word report containing the following sections: introduction, data and methods, results, discussion and conclusion.

Due date
By 5pm, April 18.
Value (% of final grade)
35%

Important course information for students

All students are expected to familiarize themselves with the following information, available on the Senate Committee on Academic Standards, Curriculum & Pedagogy webpage

- Senate Policy on Academic Honesty and the Academic Integrity Website
- Ethics Review Process for research involving human participants
- Course requirement accommodation for students with disabilities, including physical, medical, systemic, learning and psychiatric disabilities
- Student Conduct Standards
- Religious Observance Accommodation

Intellectual Property Notice

Detailed schedule of readings and activities

Week 1: Introduction to module & coding/programming

This week we will introduce ourselves to each other, the course format and the assessment activities. We will also discuss data analytics and coding/programming as a language and way of thinking.

Readings:

- Course syllabus
- Fedorenko, E., Ivanova, A., Dhamala, R., & Bers, M. U. (2019). The Language of Programming: A Cognitive Perspective. *Cell Press Reviews*, *23*(7), 525–528.
- Wing, J. (2006). Computational Thinking. *Communications of the ACM*, *49*(3), 33–35.

Week 2: Quantitative methods – what, why, when and how

This week we will delve into the world of research philosophies, discuss 2 readings in the context of different research philosophies, and go through an overview of the wide array of possible quantitative methods.

During a library-run session, we will also discuss different sources of data (primary and secondary), citation rules, copyright and permissions.

Readings:

- Oreskes, N. and E. M. Conway (2013). "The collapse of western civilization: A view from the future." *Daedalus* 142(1): 40-58.
- IPCC 6th Assessment Report – Summary for Policymakers: https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_SPM_final.pdf

Week 3: Data – where to find it, how to organise it

We will then explore the types of formats in which data can be found and the basics of data processing. We will discuss the importance of consistency in units and different ways in which data can be transformed (e.g. intensive vs extensive data, log transformation).

Readings:

- Listen to this podcast: https://www.bbc.co.uk/programmes/p0b2hfhh
- Listen to this other podcast: https://www.theguardian.com/science/audio/2022/jun/09/why-would-boris-johnson-want-to-bring-back-imperial-units
- Data errors: https://theconversation.com/the-reinhart-rogoff-error-or-how-not-to-excel-at-economics-13646
- Lamb, W. F., Mattioli, G., Levi, S., Timmons Roberts, J., Capstick, S., Creutzig, F., … Steinberger, J. K. (2020). Discourses of climate delay. *Global Sustainability*, *3*(17), 1–5. https://doi.org/10.1017/sus.2020.13

Optional complementary readings:

- Paper supporting The Conversation article (optional, if you want to go into detail):

Herndon, T, M Ash, and R Pollin. "Does High Public Debt Consistently Stifle Economic Growth? A Critique of Reinhart and Rogoff." *Cambridge Journal of Economics* 38.2 (2014): 257–279. https://ocul-yor.primo.exlibrisgroup.com/permalink/01OCUL_YOR/sqt9v/cdi_crossref_primary_10_1093_cje_bet075

- Paper referred to in the first podcast (optional, if you want to go into detail):

Schweinsberg, M., … Luis Uhlmann, E. (2021). Same data, different conclusions: Radical dispersion in empirical results when independent analysts operationalize and test the same hypothesis. *Organizational Behavior and Human Decision Processes*, *165*, 228–249. https://ocul-yor.primo.exlibrisgroup.com/permalink/01OCUL_YOR/sqt9v/cdi_swepub_primary_oai_prod_swepub_kib_ki_se_147185627

Homework 1:

Before this class, students must come up with a shortlist of topics they are interested in exploring with quantitative data for their final report. This can be your own major research topic or anything else that sparks curiosity. Think about the following:

- What question would you like to be answered by data?
- What would you need to measure (quantified) in order to be able to answer your question?
- At what level of disaggregation would you need your data to be?
    - Geographically: world, country, province, municipality, neighbourhood, street.
    - Temporally: year, month, week, day, hour, minute, second.
    - Population: groups of people (by socio-demographic characteristics), households, individuals.
    - Other relevant disaggregation (e.g. by species, by types of contaminant, etc.)
- Where might you find data related to your topics?

Also, using the knowledge gained during the Library Session in Week 2, students will bring a shortlist of potential datasets for their final report.

Week 4: Basic Python functions

During this class, we will move from excel to Python. We will use the data that students have obtained from week 2 to perform some basic Python functions and produce descriptive statistics of the data. They include: mean, median, std. dev., etc. We will also learn how to undertake correlation tests. Finally, we will learn some common data visualisation tools.

Readings:

- TBC
- TBC

Homework 2:

Before this class, students should go through Workbook 1. Coding requires learning a new language, and trial and error is key. So please do not be discouraged if things don't work on your first try (they rarely do!). Please come prepared to ask questions about what worked and what you found challenging. Workbook 1 will be released at the end of the week 2 class, and the answers for Workbook 1 will be released after week 3 class.

Week 5: Descriptive statistics and data visualization

Readings:

The following two entries from Bergstrom and West's "Calling Bullshit" course:

- Something on descriptive stats TBC
- Tools and tricks: Misleading axes
- Tools and tricks: Proportional Ink

Discourses of climate delay

A student (or several students) will present their analysis of a popular press article that makes a discourse of climate delay.

Week 6: Linear Regression (single and multi variate)

In this class we will go through the basics of performing single and multi variate linear regressions with continuous variables. This will include selection of variables, required tests to perform on variables, and analysis and reporting of regression outputs.

Readings

Read at least one of the following academic articles:

- Lamb, W. F., Steinberger, J. K., Bows-Larkin, A., Peters, G. P., Roberts, J. T., & Wood, F. R. (2014). Transitions in pathways of human development and carbon emissions. *Environmental Research Letters*, *9*, 014011. https://doi.org/10.1088/1748-9326/9/1/014011
- Steinberger, J. K., Roberts, J. T., Peters, G. P., & Baiocchi, G. (2012). Pathways of human development and carbon emissions embodied in trade. *Nature Climate Change*, *2*(2), 81–85. https://doi.org/10.1038/nclimate1371
- Knight, K. W., & Rosa, E. A. (2011). The environmental efficiency of well-being: A cross-national analysis. *Social Science Research*, *40*(3), 931–949. https://doi.org/10.1016/j.ssresearch.2010.11.002

Discourses of climate delay

A student (or several students) will present their analysis of a popular press article that makes a discourse of climate delay.

Week 7: Logistic Regression (single and multi variate)

In this class we will go through the basics of performing single and multi variate logistic regressions with categorical variables (either dichotomous or grouped). This will include selection of variables, required tests to perform on variables, and analysis and reporting of regression outputs.

Readings

Read at least one of the following academic articles:

- Riva, M., Kingunza Makasi, S., Dufresne, P., O'Sullivan, K., & Toth, M. (2021). Energy poverty in Canada: Prevalence, social and spatial distribution, and implications for research and policy. *Energy Research & Social Science*, *81*, 102237. https://doi.org/10.1016/J.ERSS.2021.102237
- O'Neill, D. W., Fanning, A. L., Lamb, W. F., & Steinberger, J. K. (2018). A good life for all within planetary boundaries. *Nature Sustainability*, *1*(2), 88–95. https://doi.org/10.1038/s41893-018-0021-4
- Baltruszewicz, M., Steinberger, J. K., Ivanova, D., Brand-Correa, L. I., Paavola, J., & Owen, A. (2021). Household final energy footprints in Nepal, Vietnam and Zambia: composition, inequality and links to well-being. *Environmental Research Letters*, *16*(2), 025011. https://doi.org/10.1088/1748-9326/abd588

Discourses of climate delay

A student (or several students) will present their analysis of a popular press article that makes a discourse of climate delay.

Week 8: Social metabolism and cross-country analysis

We will discuss the basics of social metabolism through cross-country analysis. The history and evolution of countries and region in relation to material and energy use, as well as economic activity and the wellbeing of their population.

Readings

Read at least one of the following academic articles:

- Haberl, H., M. Fischer-Kowalski, F. Krausmann, J. Martinez-Alier and V. Winiwarter (2011). "A socio-metabolic transition towards sustainability? Challenges for another Great Transformation." *Sustainable Development* 19(1): 1–14.
- Krausmann, F., Gingrich, S., Eisenmenger, N., Erb, K.-H., Haberl, H., Fischer-Kowalski, M., 2009. Growth in global materials use, GDP and population during the 20th century. *Ecological Economics* 68 (10), 2696-2705.
- Schaffartzik, A., Mayer, A., Gingrich, S., Eisenmenger, N., Loy, C., & Krausmann, F. (2014). The global metabolic transition: Regional patterns and trends of global material flows, 1950–2010. *Global Environmental Change*, *26*(1), 87–97. https://doi.org/10.1016/J.GLOENVCHA.2014.03.013

Week 9: IPAT, KAYA and STIRPAT

We will explore the widely-used IPAT identity

Readings:

- Chertow, M.R., 2000. The IPAT equation and its variants. *Journal of Industrial Ecology* 4 (4), 13-29.
- York, R., Rosa, E. A., & Dietz, T. (2003). Footprints on the earth: The environmental consequences of modernity. *American Sociological Review*, *68*(2), 279–300. https://doi.org/10.2307/1519769

Week 10: Multi-regional input-output analysis with environmental extensions

Readings:

Read at least one of the following academic articles:

- Peters, G. P., J. C. Minx, C. L. Weber and O. Edenhofer (2011). "Growth in emission transfers via international trade from 1990 to 2008." Proceedings of the National Academy of Sciences **108**(21): 8903-8908.
- Afionis, S., Sakai, M., Scott, K., Barrett, J., & Gouldson, A. (2017). Consumption-based carbon accounting: does it have a future? *Wiley Interdisciplinary Reviews: Climate Change*, *8*(1), e438. https://doi.org/10.1002/wcc.438
- Wood, R., Moran, D., Stadler, K., Ivanova, D., Steen-Olsen, K., Tisserant, A., & Hertwich, E. G. (2018). Prioritizing Consumption-Based Carbon Policy Based on the Evaluation of Mitigation Potential Using Input-Output Methods. *Journal of Industrial Ecology*, *22*(3), 540–552. https://doi.org/10.1111/JIEC.12702
- KITZES, J. 2013. An Introduction to Environmentally-Extended Input-Output Analysis. *Resources,*2**,** 489. https://doi.org/10.3390/resources2040489
- Hickel, J. (2020). Quantifying national responsibility for climate breakdown: an equality-based attribution approach for carbon dioxide emissions in excess of the planetary boundary. *The Lancet Planetary Health*, *4*(9), e399–e404. https://doi.org/10.1016/S2542-5196(20)30196-0
- Oswald, Y., Owen, A., & Steinberger, J. K. (2020). Large inequality in international and intranational energy footprints between income groups and across consumption categories. *Nature Energy*, *5*(3), 231–239. https://doi.org/10.1038/s41560-020-0579-8

Week 11: Final report consultation

Readings:

- Wall-Kimmerer, R. (2013). “Mishkos Kenomagwen: The Teachings of Grass” in *Braiding Sweetgrass*. Minneapolis, Minnesota: Milkweed Editions. https://ocul-yor.primo.exlibrisgroup.com/permalink/01OCUL_YOR/j50f41/cdi_askewsholts_vlebooks_9781571318718

Final report

Students can use the practical time to work on their final report

Week 12: Course conclusion: where to go next with environmental analytics

Readings:

- Listen to this audiobook chapter: https://timharford.com/2021/02/a-free-chapter-of-the-data-detective-audiobook/

Final report

Students can use the practical time to work on their final report