



SOC6707 – Advanced Data Analysis (Demographic Methods), Winter 2026

Monica Alexander (monica.alexander@utoronto.ca)

Office 9135 (Level 9, 700 University)

Mondays, 10-1, Room 17146

Course website: <https://github.com/MJAlexander/soc6708>

Course Description

This course introduces methods for demographic and population analysis. The goal of the course is to introduce graduate students to demography and illustrate how demographic methods could be used in their own work. Topics include population growth, mortality, fertility, migration, and kinship. We will talk about traditional demographic methods and more modern techniques using Bayesian statistical methods. We will also discuss classic and recent papers which draw on methods discussed in class.

The course will be structured as a combination of readings and examples. Each week, we will discuss topics, and work through example computer code which highlights the issues and the main concepts covered in the readings. The main assessment will be through a short project applying the concepts to a dataset of the student's choosing. Students will give a short presentation on their projects in Week 12.

Texts

A lot of the readings will come from:

- Preston, Heuvline and Guillot (2001): 'Demography: Measuring and Modeling Population Processes' (PHG)
- Wachter (2014): 'Essential demographic methods' (EDM)

Later in the course, we will switch gears to look at Bayesian methods. For that piece, the following books may be useful:

- Gelman, A., & Hill, J. (2007). 'Data analysis using regression and multilevel/hierarchical models.'
- Gelman, Andrew; Hill, Jennifer, and Vehtari, Aki. 2020. 'Regression and Other Stories'

In addition, a helpful resource (particularly for R) is:

- Alexander, Rohan. 2023. 'Telling Stories with Data'.
(<https://www.tellingstorieswithdata.com/>).

Software

A goal of this course is to familiarize students with using the statistical language R to perform demographic and statistical analyses. All code for the course will be put in the class GitHub repo. If you have not used R before, you will need to download R and RStudio:

- Download R here: <https://www.r-project.org/>
- Download RStudio (free version) here: <https://posit.co/download/rstudio-desktop/>

Later in the course, we will be using Stan to fit Bayesian models (through `rstan`). This can be annoying to install; here's some helpful notes if you run into trouble: <https://github.com/stan-dev/rstan/wiki/Rstan-Getting-Started>.

Course Requirements

In class labs (20%)

Each week during class we will go through some of the methods discussed in class with practical examples in R. During these labs there will be several questions that will be required to be submitted via GitHub. There will be a total of 10 labs worth 2% each.

Research project

Students will develop a research question of their choice using a dataset of your choice, address it by using the descriptive and inferential techniques presented in the course to analyze data in R, and write a short report summarizing your findings. Your overall grade on the research paper will be the sum of your grades on the following assignments:

- **Research proposal (10%)**: state the research question, the dataset that you will be using to answer this question, and the key independent and dependent variables in the dataset that you will use to answer this question. Show one or two key graphs or tables that help to inform your analysis.
- **Presentation (20%)**: Present short summary of your findings in the final class of semester.
- **Final report (40%)**: The final write-up of the report, should be structured as a short scientific report, and include an Introduction, description of Data, Description of methods of analysis, Results, and Discussion.

Each written component should be completed in Quarto/RMarkdown, and the submission should include the knitted pdf and Quarto/RMarkdown file containing the necessary code to produce the results. All parts should be handed in electronically via Quercus.

Course Policies

Communication

The best way to ask questions about course material or assignments is in person during the lecture, lab, or office hours. The following are guidelines for email communication: please make sure that you have a legitimate need before you write and that you cannot resolve your question

during the lecture, lab or office hours; email messages should state the course number and the purpose of the email clearly in the subject line.

Late Homework Assignments

If you are unable to turn in an assessment for medical reasons, you will need to email me and also declare your absence on ACORN, within one week of the missed assessment. For other reasons, such as family or other personal reasons, please contact your college registrar and have them email me directly.

Academic Integrity Clause

Copying, plagiarizing, falsifying medical certificates, or other forms of academic misconduct will not be tolerated. Any student caught engaging in such activities will be referred to the Dean's office for adjudication. Any student abetting or otherwise assisting in such misconduct will also be subject to academic penalties. Students are expected to cite sources in all written work and presentations. See this link for tips for how to use sources well:

(<http://www.writing.utoronto.ca/advice/using-sources/how-not-to-plagiarize>).

According to Section B.I.1.(e) of the Code of Behaviour on Academic Matters it is an offence "*to submit, without the knowledge and approval of the instructor to whom it is submitted, any academic work for which credit has previously been obtained or is being sought in another course or program of study in the University or elsewhere.*"

By enrolling in this course, you agree to abide by the university's rules regarding academic conduct, as outlined in the Calendar. You are expected to be familiar with the *Code of Behaviour on Academic Matters* (<https://governingcouncil.utoronto.ca/secretariat/policies/code-behaviour-academic-matters-july-1-2019>) and *Code of Student Conduct* (<http://www.viceprovoststudents.utoronto.ca/publicationsandpolicies/codeofstudentconduct.htm>) which spell out your rights, your duties and provide all the details on grading regulations and academic offences at the University of Toronto.

Normally, students will be required to submit their course essays to www.ouriginal.com for a review of textual similarity and detection of possible plagiarism. In doing so, students will allow their essays to be included as source documents in the www.ouriginal.com reference database, where they will be used solely for the purpose of detecting plagiarism. The terms that apply to the University's use of the www.ouriginal.com service are described on the www.ouriginal.com web site.

Accessibility Services

It is the University of Toronto's goal to create a community that is inclusive of all persons and treats all members of the community in an equitable manner. In creating such a community, the University aims to foster a climate of understanding and mutual respect for the dignity and worth

of all persons. Please see the University of Toronto Governing Council “Statement of Commitment Regarding Persons with Disabilities” at <http://www.governingcouncil.utoronto.ca/Assets/Governing+Council+Digital+Assets/Policies/PPDF/ppnov012004.pdf>.

In working toward this goal, the University will strive to provide support for, and facilitate the accommodation of individuals with disabilities so that all may share the same level of access to opportunities, participate in the full range of activities that the University offers, and achieve their full potential as members of the University community. We take seriously our obligation to make this course as welcoming and accessible as feasible for students with diverse needs. We also understand that disabilities can change over time and will do our best to accommodate you.

Students seeking support must have an intake interview with a disability advisor to discuss their individual needs. In many instances it is easier to arrange certain accommodations with more advance notice, so we strongly encourage you to act as quickly as possible. To schedule a registration appointment with a disability advisor, please visit Accessibility Services at <http://www.studentlife.utoronto.ca/as>, call at 416-978-8060, or email at: accessibility.services@utoronto.ca. The office is located at 455 Spadina Avenue, 4th Floor, Suite 400.

Additional student resources for distressed or emergency situations can be located at distressedstudent.utoronto.ca; Health & Wellness Centre, 416-978-8030, <http://www.studentlife.utoronto.ca/hwc>, or Student Crisis Response, 416-946-7111.

Equity and Diversity Statement

Equity and Diversity

The University of Toronto is committed to equity and respect for diversity. All members of the learning environment in this course should strive to create an atmosphere of mutual respect. As a course instructor, I will neither condone nor tolerate behaviour that undermines the dignity or self-esteem of any individual in this course and wish to be alerted to any attempt to create an intimidating or hostile environment. It is our collective responsibility to create a space that is inclusive and welcomes discussion. Discrimination, harassment and hate speech will not be tolerated.

Additional information and reports on Equity and Diversity at the University of Toronto is available at <http://equity.hrandequity.utoronto.ca>.

Use of Generative AI in Assignments

In general, students are discouraged from using generative AI tools, such as ChatGPT4 in this course. Specific course policies around the use of generative AI tools are as follows:

- **R code:** Students may wish to use generative AI tools to aid in initial development and writing of R code to answer assignment questions and carry out the research project analysis. If this is the case, the use of such tools should be explicitly acknowledged in the submitted work, and the relevant prompts should be included as an appendix to the assignment. In addition, all code must be thoroughly commented or described in the submitted work. Failure to do so may result in penalties.
- **Written work:** Using generative AI tools to generate written answers to assignment questions of text contained in the final research project is prohibited in this course. Representing as one's own an idea, or expression of an idea, that was AI-generated may be considered an academic offense in this course.

This course policy is designed to promote your learning and intellectual development and to help you reach course learning outcomes.

Course Schedule

Date	Lecture	Lab	Readings
5 January Week 1	Introduction <ul style="list-style-type: none"> - What is demography and why does it matter? - Demographic identity - Demographic rates - Models without age - Age, periods, cohorts 	<ul style="list-style-type: none"> - R/RStudio/Quarto - Introduction to GitHub - Intro demog plotting 	<ul style="list-style-type: none"> - PHG Chapters 1-2 - EDM Chapters 1-2 - Keyfitz, N. 1975. 'How Do We Know the Facts of Demography?' Population and Development Review. 1(2):267-288. - Malthus, T. 1798. 'An Essay on the Principle of Population'. Chapters 4-5 - Ryder, NB. 1964. 'Notes on the Concept of a Population.' American Journal of Sociology 69: 447-62. - *Wallace, M., Gillispie-Bell, V., Cruz, K., Davis, K., & Vilda, D. (2021). Homicide during pregnancy and the postpartum period in the United States, 2018–2019. Obstetrics & Gynecology, 138(5), 762-769.
12 January Week 2	Mortality <ul style="list-style-type: none"> - Decomposition - Life tables - Multi-state life tables - Cause-deleted life expectancy - Entropy, lifespan disparity 	<ul style="list-style-type: none"> - Decomposition - Life tables 	<ul style="list-style-type: none"> - PHG Chapters 3,4 - EDM Chapters 3,7 - Graunt, J. 1662. 'Natural and Political Observations Mentioned in a Following Index, and Made Upon the Bills of Mortality' in Smith, D and Keyfitz, N. 'Mathematical Demography'. Chapter 2. - Van Raalte, A. A., Sasson, I., & Martikainen, P. (2018). The case for monitoring life-span inequality. Science, 362(6418), 1002-1004. - Nigri, A., Barbi, E., & Levantesi, S. (2022). The relationship between longevity and lifespan variation. Statistical Methods & Applications, 31(3), 481-493. - *Lundberg, D. J., Wrigley-Field, E., Cho, A., Raquib, R., Nsoesie, E. O., Paglino, E., ... & Stokes, A. C. (2023). COVID-19 mortality by race and ethnicity in US metropolitan and nonmetropolitan areas, March 2020 to February 2022. JAMA network open, 6(5), e2311098-e2311098.

19 January Week 3	Mortality II <ul style="list-style-type: none">- Mortality models: parametric, relational- Hazard models, heterogeneity- Oldest-old mortality	<ul style="list-style-type: none">- Gompertz- SVD, Lee-Carter	<ul style="list-style-type: none">- PHG Chapters 3,4, 9.1, 11.1- EDM Chapters 3,7,8- *Barbi, E, Lagona, F, Marsili, M, Vaupel, J and Wachter, K. 2018. 'The plateau of human mortality: Demography of longevity pioneers', <i>Science</i>, 360: 1459-1461.<ul style="list-style-type: none">+ Critique: Newman, SJ. 2018. 'Plane inclinations: A critique of hypothesis and model choice in Barbi et al'. <i>PLoS Biol</i> 16(12): e3000048.+ Ken's response: Wachter, K. 2018. 'Hypothetical errors and plateaus: A response to Newman'. <i>PLoS Biol</i> 16(12): e3000076.- Dang, L. H. K., Camarda, C. G., Ouellette, N., Meslé, F., Robine, J. M., & Vallin, J. (2023). The question of the human mortality plateau. <i>Demographic Research</i>, 48, 321-338.- Feehan, D. 2018. 'Separating the Signal From the Noise: Evidence for Deceleration in Old-Age Death Rates'. <i>Demography</i> 55(6):2025--2044.- Lee, RD, and Carter, LR. 1992. 'Modeling and Forecasting US Mortality.' <i>Journal of the American Statistical Association</i> 87 (419). Taylor & Francis: 659-71.- Oeppen, J, and Vaupel, JW. 2002. 'Broken limits to life expectancy.' <i>Science</i> 296.5570: 1029-1031. (Updated graph: https://ourworldindata.org/the-rise-of-maximum-life-expectancy)- Van Raalte, A. A. (2021). What have we learned about mortality patterns over the past 25 years?. <i>Population Studies</i>, 75(sup1), 105-132.
26 January Week 4	Fertility <ul style="list-style-type: none">- Basic indicators- Parity- Childlessness- Mean age at childbearing- Fertility models- Tempo and quantum	<ul style="list-style-type: none">-Switching between indicators- Tempo and quantum	<ul style="list-style-type: none">- PHG Chapters 5, 9.3- EDM Chapters 4, 6- Bongaarts, J and Feeney, G. 1998. 'On the Quantum and Tempo of Fertility'. <i>Population and Development Review</i>, 24(2):271-291.- Lima, E., Zeman, K., Sobitka, T., Nathan, M., and Castro, R. (2018). The Emergence of Bimodal Fertility Profiles in Latin America. <i>Population and Development Review</i>, 44(4). 723-743.

			<ul style="list-style-type: none"> - Beaujouan, E. (2020). Latest-late fertility? Decline and Resurgence of Late Parenthood across the Low Fertility Countries. <i>Population and Development Review</i>, 46(2), 219-247. - Lazzari, E., Mogi, R., & Canudas-Romo, V. (2021). Educational composition and parity contribution to completed cohort fertility change in low-fertility settings. <i>Population Studies</i>, 75(2), 153-167 - Nobles, J. Gemmill, A., Hwang, S., and Torche, F. (2024). Fertility in a Pandemic: Evidence from California. <i>Population and Development Review</i>, 50(s1), 101-128. - *Chen, S., & Gietel-Basten, S. (2025). Only Children and Low Family Size Ideals: Did the One-Child Policy Create a “Low-Fertility Trap” in China?. <i>European Journal of Population</i>, 41(1), 32.
9 February Week 5	Population projection and stable populations <ul style="list-style-type: none"> - Leslie matrices - Lotka's r - Reproductive value - Population momentum - Generational waves - Probabilistic projections 	<ul style="list-style-type: none"> - Building Leslie matrices - Population projections 	<ul style="list-style-type: none"> - PHG Chapters 6,7 - EDM Chapters 5,10 - Keyfitz, N. 1985. <i>Applied mathematical demography</i>. Second edition. Chapter 6. - Bernadelli, H. 1941. 'Population Waves' in Smith, D and Keyfitz, N. 'Mathematical Demography'. Chapter 23. - Espenshade, Thomas J., Analia S. Olgiati, and Simon A. Levin. "On nonstable and stable population momentum." <i>Demography</i> 48, no. 4 (2011): 1581-1599. - Goodkind, D. 2017. 'The Astonishing Population Averted by China's Birth Restrictions: Estimates, Nightmares and Reprogrammed Ambitions'. <i>Demography</i>. 54:1375-1400. + Lots of responses, e.g.: Zhao, Z and Zhang, G. 2018. 'Socioeconomic Factors Have Been the Major Driving Force of China's Fertility Changes Since the Mid-1990s'. <i>Demography</i>. 55: 733-742. - Peak global population and other key findings from the 2024 UN World Population Prospects. Our World in Data. https://ourworldindata.org/un-population-2024-revision - * Polizzi, A., & Tilstra, A. M. (2024). The impact of early death on birth counts in the United States, 1950 to 2019. <i>PNAS nexus</i>, 3(6), pgae058.
16 February	READING WEEK		

23 February Week 6	Migration <ul style="list-style-type: none">- Fundamental issues of measurement- Incorporating migration into Leslie Matrix- Migration models- Social media data	<ul style="list-style-type: none">- Rogers Castro- Facebook	<ul style="list-style-type: none">- PHG Chapter 9.4, EDM Chapter 11- Wilson, T. (2010). Model migration schedules incorporating student migration peaks. <i>Demographic research</i>, 23, 191-222.- Yildiz, D., & Abel, G. (2021). Migration stocks and flows: data concepts, availability and comparability. In <i>Research handbook on international migration and digital technology</i> (pp. 29-41). Edward Elgar Publishing.- Abel, G. J., & Cohen, J. E. (2022). Bilateral international migration flow estimates updated and refined by sex. <i>Scientific Data</i>, 9(1), 173.- Chi, G., Abel, G. J., Johnston, D., Giraudy, E., & Bailey, M. (2025). Measuring global migration flows using online data. <i>Proceedings of the National Academy of Sciences</i>, 122(18), e2409418122.- *Alexander, M., Polimis, K., and Zagheni, E., ‘The impact of Hurricane Maria on out-migration from Puerto Rico: Evidence from Facebook data’, <i>Population and Development Review</i>, 2019, 45(3): 617-630.
2 March Week 7	Kinship <ul style="list-style-type: none">- Terminology- Keyfitz-Goodman equations- Matrix-based models	<ul style="list-style-type: none">- KG ‘by hand’- DemoKin- Comparison to survey data	<ul style="list-style-type: none">- Goodman, L, Keyfitz, N and Pullum, T. 1974 'Family Formation and the Frequency of Various Kin Relationships'. <i>Theoretical Population Biology</i>. 5: 1-27.- Caswell, H. (2019). The formal demography of kinship. <i>Demographic Research</i>, 41, 679-712. (there's subsequent parts, I think we're now up to part VI).- Alburez-Gutierrez, D., Mason, C., & Zagheni, E. (2021). The “sandwich generation” revisited: Global demographic drivers of care time demands. <i>Population and Development Review</i>, 47(4), 997-1023.- Smith-Greenaway, E., Alburez-Gutierrez, D., Trinitapoli, J., & Zagheni, E. (2021). Global burden of maternal bereavement: indicators of the cumulative prevalence of child loss. <i>BMJ global health</i>, 6(4).- *Schluter, B. S., Alburez-Gutierrez, D., Bibbins-Domingo, K., Alexander, M., and Kiang, M. V. ‘Youth experiencing parental death due to drug poisoning and firearm violence in the US, 1999-2020’. <i>JAMA</i>, 2024, doi:10.1001/jama.2024.8391.

9 March Week 8	Introduction to Bayes - Bayes' rule - Bayes' rule for parameters - Prior choice	- Bayes' rule - Visualization of prior	- Gelman and Hill - *Schmertmann, C and Hauer, M. 2018. 'Bayesian estimation of total fertility from a population's age–sex structure'. Statistical Modelling 19(3): 1-23.
16 March Week 9	Bayes II - Computation - Introduction to Stan - Bayesian regression	- Stan Stan Stan	- Gelman and Hill - *Alexander, M., and Root, L., 'Competing effects on the average age of infant death'. Demography, 2022. 59 (2): 587–605.
23 March Week 10	Bayesian demography - Bayesian Gompertz - Probabilistic projection, revisited - Indirect estimation - Measurement error	- Gompertz - 1000 Leslie Matrices	- Bijak, J., & Bryant, J. (2016). Bayesian demography 250 years after Bayes. Population studies, 70(1), 1-19. - Lynch, S. M., & Bartlett, B. (2019). Bayesian statistics in sociology: Past, present, and future. Annual Review of Sociology, 45(1), 47-68. - Raftery, A. E., Alkema, L., & Gerland, P. (2014). Bayesian population projections for the United Nations. Statistical science, 29(1), 58. - Alexander, M., & Raftery, A. E. (2024). Developing and implementing the UN's probabilistic population projections as a milestone for Bayesian demography. Demographic Research, 51, 1-16. - Byant, J., and Zhang, J. 'Bayesian Demographic Estimation and Forecasting' - Alkema, L, Raftery, A, Gerland, P and Clark, S 2011. 'Probabilistic Projections of the Total Fertility Rate for All Countries'. Demography. 48(3): 815-839. - Alexander, M., and Alkema, L., 'A Bayesian Cohort Component Projection Model to Estimate Women of Reproductive Age at the Subnational Level in Data-Sparse Settings', Demography, 2022. 59(5): 1713–1737.

30 March Week 11	The future - The surveys don't work they just make it worse - Estimating the probability of events without denominators - Generative models	TBD	
6 April Week 12	Research presentations and recap		
13 April	Research project due		

