QUANTITATIVE POLITICAL AND POLICY ANALYSIS

POLSCI 784
Term 2, Winter 2021

NOTE: The official course outline is posted on the course website. This version is for reference only, and links may not be enabled.

Instructor: Dr. Michelle L. Dion
Email: dionm@mcmaster.ca
Office: KTH-533
Office Hours: Mondays, 10:30am-11:00am, no appointment needed on Zoom

Seminar: Wednesdays, 08:30-11:20
Room: virtual

Contents

Course Description

This is an *introductory* graduate course in empirical research and statistical methods. For MA students, the intention is to provide you with basic statistical skills and familiarity for use on the job market. For PhD students, the goal is to provide a basic foundation for more advanced coursework or applications in your research. For some of you, the material presented in this course will be the beginning of a radically new way to approach research. To be successful in the course, you will NOT need to be a mathematician, statistician, or computer programmer, but you will need a desire to learn, to solve problems, and be open to new ways of thinking. Reading *all* assignments and instructions and asking questions if you need additional explanation or clarification is also important. You will also need some basic algebra skills. A copy of this outline and related web-based resources for the course can be found at the course websites:

- Course website for notes, discussion board, resources, downloading and uploading assignments, etc.: https://github.com/polsci784-2021.
    - You will receive an email invitation to join the course webpage.
    - Once you accept the invitation, you will create a github.com account, if you don't already have one.
- Curated collection of data, R, and related websites: https://bkmrk.michelledion.com/

Course Objectives

The course will provide an introduction to basic statistical methods in the social sciences through linear (and to a lesser extent logistic) regression. The emphasis will be on successful application of statistical methods and understanding the uses of such methods for public policy and social science. To gain experience in applying statistical analysis, students complete a series of practice assignments and an independent research project employing linear or logistic regression. Students will also gain experience using tools of data science, such as RStudio & GitHub. They will also learn how to plan, manage, and execute an applied quantitative research project.

Required Materials and Texts

All textbooks and resources for this course are available without cost online. Some books and resources are available through McMaster University Library subscriptions, and you'll [either need to be on campus or] login to the library website for access from off campus. Some are open source books or websites. Many are also available for purchase in paper formats, including as used textbooks (e.g., from an independent bookseller, brick & mortar chain, or online).

Link on item *title* goes to McMaster Library login for off-campus access.

- Lewis-Beck, Michael. 1995. *Data Analysis*. Thousand Oaks, CA: SAGE. https://dx.doi.org/10.4135/9781412983846
- Schroeder, Larry D., David L. Sjoquist, and Paula E. Stephan. 1986. *Understanding Regression Analysis: An Introductory Guide*. Thousand Oaks: SAGE. https://dx.doi.org/10.4135/9781412986410
- Lewis-Beck, Colin, and Michael Lewis-Beck. 2016. *Applied Regression: An Introduction*. Thousand Oaks: SAGE. https://doi.org/10.4135/9781483396774
- Fox, John. 1991. *Regression Diagnostics*. Thousand Oaks, CA: SAGE. https://dx.doi.org/10.4135/9781412985604
- Hardy, Melissa A. 1993. *Regression with Dummy Variables*. Thousand Oaks: SAGE. https://dx.doi.org/10.4135/9781412985628
- Kahane, Leo H. 2008. *Regression Basics*. 2nd ed. Thousand Oaks: SAGE. https://dx.doi.org/10.4135/9781483385662
- Menard, Scott. 2002. *Applied Logistic Regression Analysis*. 2nd ed. Thousand Oaks: SAGE. https://dx.doi.org/10.4135/9781412983433

Computing and Statistical Software

In this course, we will be using the R Statistical Computing language to analyze data. R is increasingly used in the private and public sectors for data analysis due to its flexibility and power (see Long & Turner 2020). Much of the power of R comes from user-contributed or developed packages that add new functions or types of analysis to "base" R. The tidyverse is a collection of R packages that implement coordinated and consistent data analysis approaches and tools. The "tidy" approach to data analysis in R is quickly displacing "old school" approaches that use base R commands.

To interact with R, we will be using Desktop RStudio, which is a free Integrated Development Environment that provides an interface for R that is similar to interfaces for other statistical software, such as SPSS or STATA. In RStudio, you can run command scripts, interactively analyze your data, and view your data or results, including tables and plots. RStudio also has a range of other features that make it easy to prepare your results to share with collaborators or add to papers.

RStudio also integrates with GitHub.com :octocat: to track and archive your data and files. Git is an open source program that tracks versions of code and documents, and combined with GitHub.com, you can transfer your files to your account directly from RStudio. We will be using GitHub as our course website, where we will have a private space to share class notes, access and submit practice assignments, and ask questions on a class discussion board. *You will receive an invitation to join the course on*

github.com. If you don't already have an account on github.com, you will be prompted to create one.

Primary R textbooks & websites

- Grolemund, Garrett, and Hadley Wickham. n.d. *R for Data Science*. O'Reilley. *Also available in hard copy.*
    - Solutions manual for *R for Data Science*: Arnold, Jeffrey B. n.d. *R for Data Science: Exercise Solutions*.
- Phillips, Nathaniel D. n.d. *YaRrr! The Pirate's Guide to R*. (solutions included in Chapter 18)

Reference or alternative resources

Statistics textbooks (some available from the library, order from basic to advance

- Pollock, Phillip. 2015. *The Essentials of Political Analysis*. 5th ed. edition. Los Angeles: CQ Press.
- Neil J. Salkind. *Statistics for People who (Think They) Hate Statistics*. 4th edition. Sage. ISBN 1412979595 (covers most of material in class)
- Alan Agresti and Barbara Finlay. *Statistical Methods for the Social Sciences*, 4th edition. Upper Saddle River, NJ: Prentice Hall. ISBN 0130272957 (covers class material)
- Earl Babbie. *The Practice of Social Research*, 13th ed. Wadsworth. ISBN: 1133049796. (covers class material and qualitative methods and data collection methods, like survey design)
- Janet Buttolph Johnson and H.T. Reynolds. *Political Science Research Methods*. CQPress. ISBN: 1608716899. (covers class material, qualitative methods, and general research strategies)
- Damodar N. Gujarati and Dawn Porter. *Essentials of Econometrics* 4th Ed. ISBN: 0073375845 (solid basic introduction to regression and some advanced topics, does not use matrix algebra)
- Damodar N. Gujarati and Dawn Porter. *Basic Econometrics.* 5th ed. ISBN: 0073375772 (solid basic introduction to regression and some advanced topics, used in many political science grad programs, uses matrix algebra)
- William H. Greene. *Econometric Analysis*. 7th ed. Prentice Hall. ISBN: 0131395386. (advanced econometric text that covers most advanced methods in more detail than Gujarati, used in many political science grad programs)

Other statistics reference materials online

Before consulting Wikipedia or Google for basic statistical concepts or terms, search one of these methods encyclopedias available through the McMaster Library website:

- Lewis-Beck, Michael, Alan Bryman, and Tim Liao. 2004. *The SAGE Encyclopedia of Social Science Research Methods*. 3 vols. Thousand Oaks, California. https://doi.org/10.4135/9781412950589

- SAGE Little Green Books.

NB: SAGE also has a Little Blue Book series for qualitative methods.

Additional R reference material and learning resources

- RStudio cheat sheets for various R and tidyverse packages
- Michelle Dion's collection of R resources
- UCLA's IDRE. R Statistical Computing.
- Healy, Kieran. 2018. *Data Visualization: A Practical Introduction*. 1st ed. Princeton, NJ: Princeton University Press.
- Healy, Kieren. 2018. "Appendix". In *Data Visualization*. Princeton: Princeton UP. *Also available in hard copy.*
- McConville, Chester Ismay and Albert Y. Kim Foreword by Kelly S. 2019. *Statistical Inference via Data Science*. New York: CRC Press. *Also available in hard copy.*
- Mailund, Thomas. 2019. *R Data Science Quick Reference: A Pocket Guide to APIs, Libraries, and Packages*. Berkeley, CA: press. DOI. Separate chapters for different tidyverse packages, similar to *R for Data Science*.

Swirl interactive learning exercises

Swirl is an interactive R package that walks users through fundamental concepts and basic commands of the R Statistical Computing language. In addition to the original core Swirl modules that cover basic data types and related concepts in R, several new collections of modules are interactive learning tools for probability and statistical inference, data cleaning using the tidyverse, linear regression, and data visualization using ggplot. Swirl modules vary in length and are not a substitute for class or practice assignment exercises. Instead, they are useful reference tools for those who want to develop more confidence and intuition using R.

Instructions for how to install Swirl & explore R are below.

LinkedIn Learning (Lynda.com) via McMaster subscription

McMaster University maintains a subscription to various online courses hosted at LinkedIn Learning (formerly Lynda.com). These courses are not a substitute for class lectures or working through practice assignment exercises, but I include links to short (all < 10 min. and most < 5 min.) videos that cover some of each week's content as an additional resource. These are optional and supplemental. The full list of available courses is available at: https://uts.mcmaster.ca/services/teaching-and-learning/linkedin-learning/.

To use the links listed in the Weekly Schedule below: 1. Login at: https://uts.mcmaster.ca/services/teaching-and-learning/linkedin-learning/ 2. While your

browser is still open, in another tab, click on the links to the videos on this page you want to view.

Git & GitHub background & resources

Some of these may discuss some advanced features that we do not use.

- JennyBC's instructions on installing git
- RStudio and Git - an Overview (Part 1)
- A simple overview
- RStudio and Git - an Example (Part 2)
- RStudio & Github Integration
- How to use Git and GitHub with R
- Happy Git and GiHub for the useR - online book for students explaining git/github
- Background on git, markdown, etc. by Jeff Racine (Econ, McMaster)
- Another nice overview
- Hadley Wickham's (tidyverse creator) page on git, GitHub & RStudio, lots of detail.

Other class-related website

On this website, I save links to other resources on the web, with a focus on: - websites with data, tagged "data" + other keywords that describe the dataset; - websites related to statistical methods and tools, tagged "stats" + other keywords that describe the method; and - websites related to statistical computing and related tools, tagged "stats" + name of the tool (e.g., R, R Studio, tidyverse, etc.)
The data links are particularly useful for finding data for projects.

Getting Help

If you've been working on the same line of code for more than 20 minutes or are seeing an error message that you haven't seen before, your classmates and instructor are the best source of help for specific assignments in this course. To post questions about assignments, we have a discussion board. - Use this space to ask questions about assignments, class content, or to organize yourselves into study groups. - I prefer to answer most questions on the discussion board because then everyone has the same information about an assignment or class content. I find that if one student has a question, it is likely that at least one other (and often more!) student has the same or a similar question. You can always tag me @michelledion to make sure that I see your question. - Answering each others' questions here also helps everyone practice asking for help online and providing it to others. Most people will find that using an online forum like this is useful not just for this class but for other technical issues (even outside of class or work) that you may encounter in the future. - I will often leave a query unanswered for a day to give others a chance to answer, but then will either confirm advice provided by others or suggest additional strategies. - I will sometimes suggest that you remind me to answer a question in class or to come to office hours. In those cases, the question may be better answered when I can "show" you the solution or add nuance to the answer. - This StackOverflow page has some useful guidance about how

to post a good question, and I would recommend trying to follow this guidance when you post to the class discussion board.

Class notes

Class materials (including all the code and data used to generate the notes) are stored in this repository in the notes folder.

I will demonstrate how to use the materials in this folder to take notes during class.

Class Format

The class runs as a weekly 3 hour seminar. Assignment due dates are included in the Weekly Schedule below and on the course Google Calendar.

Course Evaluation – Overview

1. Practice assignments, 30%, due various weeks
2. Final research project, 70%, due April 21, 2021

Course Evaluation – Details

Practice assignments (30%), due various weeks

Practice assignments are designed to build skills and develop confidence and competence with collaborative assignments that are marked based on completeness and effort. We will sometimes use class time to complete some parts of practice assignments (e.g., some piece of code), and students will need to answer remaining questions (e.g., those that require interpretation) and assemble and submit the complete assignment for credit.

Students are encouraged to discuss practice assignment assignments and work together to solve coding bugs and confirm understanding of course content.

However, each student must turn in their own practice file that includes all output and answers *in their own words*.

This approach is meant to provide timely feedback to students (when we review the "answers" in class), help me quickly identify material that requires collective review, and minimize everyone's worry about marks in order to facilitate student engagement and learning. However, this approach to marking also requires that students complete assignments according to the course schedule, and late practice assignments will be accepted only *under exceptional circumstances and prior approval from the instructor*.

Links to all practice assignments are below. The link will take you to the GitHub Education tool that makes a new, private *copy* of the assignment repository with your

GitHub username in the repository title. In class, we will walk through this process together.

1. Informational notecard 0/30 (We will complete this practice exercise together in class.)
2. Article comparison 4/30
3. Descriptive statistics 5/30
4. Bivariate regression 7/30
5. Multivariate regression 7/30
6. Regression assumptions 7/30 (Assignment does not require coding/data analysis)

Final research project (70%), due in phases

More than half of your final mark in this course will be based upon your completion of an original research project using quantitative data (e.g., a microdata file of survey responses or an aggregate dataset that you compile based on published sources) and linear (or logistic) regression.

The project will proceed in phases to give you guidance and feedback throughout the research process. You will present your findings on the last day of class, and submit a revised final report that assembles all the project assignments into one revised, final documenht.

Though students will complete the project in stages, there is only one project repository per student.

Detailed instructions for each project assignment are in the README.md file in the project repostiory.

Use the above link to “accept” the project assignment *after the first day of class* (link will not be *live* before class) and create your private repository.

Submission of research project assignments through the project repository will be discussed and demonstrated in class.

I may edit project assignment templates or add other resources during term, so please “pull” updates to the repository before you begin working on a particular project assignment.

Your final research project will proceed in phases:

1. Statement of research question with clear identification of dependent variable and preliminary bibliography (5%)
2. Description of research hypotheses and bibliography (10%)
3. Diagram of research design (5%)

4. Description of data and sources bibliography (10%)
5. Description of analysis and results (15%)
6. Final presentation of results (10%)
7. Final report (15%)

Weekly Course Schedule and Required Readings

Week 1 (Jan 13) Introduction

Due: Install R, RStudio, git & create GitHub.com [educational] account

Content readings:

- Achen, Christopher H. 2002. "Advice for Students Taking a First Political Science Graduate Course in Statistical Methods," *The Political Methodologist* 10 (2): 10–12.
- Lewis-Beck, Michael S. 1995. *Data Analysis: An Introduction*. Thousand Oaks: SAGE, pgs. 1-8.

Review mainly for structure & description of the data & analysis: - Davis, Darren W., and Brian D. Silver. 2004. "Civil Liberties vs. Security: Public Opinion in the Context of the Terrorist Attacks on America." *American Journal of Political Science* 48 (1): 28–46. - Dion, Michelle L., and Catherine Russler. 2008. "Eradication Efforts, the State, Displacement and Poverty: Explaining Coca Cultivation in Colombia during Plan Colombia." *Journal of Latin American Studies* 40 (3): 399–421.

R resources:

From course website:

- Install R & R Studio Desktop
- Install TeX & git

Other resources

- Grolemund, Garrett, and Hadley Wickham. n.d. *R for Data Science*. O'Reilley. Chapters 1-2
- Healy, Kieren. 2018. *Data Visualization*. Princeton: Princeton UP. Chapter 2
- Github Education Classroom video.
- Phillips, Nathaniel D. n.d. *YaRrr! The Pirate's Guide to R*. Chapters 1-2.

Swirl:

- Install Swirl & explore R
- Swirl 1-2

LinkedIn:

- Navigating the RStudio Environment: Learning R
- Why Is the Tidyverse Unique?: R Programming in Data Science: Setup and Start
- Swirl: R for Data Science: Lunchbreak Lessons
- Packages for R: Learning R
- The Tidyverse: Learning R

GitHub & RStudio videos:

- RStudio and Git - an Overview (Part 1)
- RStudio and Git - an Example (Part 2)
- RStudio & Github Integration
- How to use Git and GitHub with R

Week 2 (JAN 20) Univariate descriptive statistics

Due: Practice assignments 0 & 1 & Project research question \[& make sure your RStudio & GitHub workflow is set up before class.\]

Content readings:

- Lewis-Beck, Michael S. 1995. *Data Analysis: An Introduction*. Thousand Oaks: SAGE, pgs. 9-18.

R resources:

- Grolemund, Garrett, and Hadley Wickham. n.d. *R for Data Science*. O'Reilley. Chapters 4, 6, 8-9, 26-30.
- Phillips, Nathaniel D. n.d. *YaRrr! The Pirate's Guide to R*. Chapters 3-4, 9.
- Irizarry, Rafael A. n.d. *Introduction to Data Science*. Chapter 39 Git and GitHub.
- Bryan, Jennifer, and Jim Hester. n.d. What They Forgot to Teach You About R. Chapters 1-2.

Swirl:

- Swirl: 3-7
- Statistical inference: 1-3

LinkedIn:

- How to Install the Tidyverse: R Programming in Data Science: Setup and Start
- Importing Data from a Spreadsheet: Learning R
- What Are CSV Files?: Data Wrangling in R
- Importing CSV Files into R: Data Wrangling in R
- Importing Excel Files into R: Data Wrangling in R
- Loading Data Sets with Read_csv: Data Visualization in R with Ggplot2
- Recoding Variables: Learning R

- Computing New Variables: Learning R
- Computing Descriptives: Learning R
- Computing Frequencies: Learning R

Week 3 (JAN 27) Hypothesis testing and statistical significance

Content readings:

- Lewis-Beck, Michael S. 1995. *Data Analysis: An Introduction*.Thousand Oaks: SAGE, pgs. 31-41.

R resources:

- Grolemund, Garrett, and Hadley Wickham. n.d. *R for Data Science*. O'Reilley. Chapters 3, 5, 7.

Swirl:

- Getting and Cleaning Data: 1-3
- Statistical inference: 4-11

LinkedIn:

- Using the Tidyverse: Data Wrangling in R
- Variables, Observations, and Values: Data Wrangling in R
- What Is Tidy Data?: Data Wrangling in R
- Subsetting Tibbles: Data Wrangling in R
- Filtering Tibbles: Data Wrangling in R
- R Data Types: Basic Types: R for Data Science: Lunchbreak Lessons
- R Data Types: Data Frame: R for Data Science: Lunchbreak Lessons
- R Data Types: Factor: R for Data Science: Lunchbreak Lessons
- Piping Commands with %>%: Learning R
- Selecting Cases and Subgroups: Learning R

Week 4 (FEB 3) Exploratory data visualization

Due: \[should have draft started of practice assignment 2 to ask questions in class about code / content\]

Recommended content readings:

- Healy, Kieran, and James Moody. 2014. "Data Visualization in Sociology." *Annual Review of Sociology* 40 (1): 105–28.

R resources:

- Healy, Kieren. 2018. *Data Visualization*. Princeton: Princeton UP. Chapters 1, 3, 4.

- Grolemund, Garrett, and Hadley Wickham. n.d. *R for Data Science*. O'Reilley. Chapters 10-12, 14-15.
- Phillips, Nathaniel D. n.d. *YaRrr! The Pirate's Guide to R*. Chapter 11.

Swirl:

- Swirl: 8-9,12
- Exploratory Data Analysis: 1-4, 7-10

LinkedIn:

- Introducing Ggplot2: Data Visualization in R with Ggplot2
- Barplot: R for Data Science: Lunchbreak Lessons
- Bars and Columns: Data Visualization in R with Ggplot2
- Dotchart: R for Data Science: Lunchbreak Lessons
- Histogram: R for Data Science: Lunchbreak Lessons
- Histograms: Data Visualization in R with Ggplot2
- Scatterplots: Data Visualization in R with Ggplot2

Week 5 (FEB 10) Association & simple, bivariate regression

Due: Practice assignment 2
Due: **Feburary 12** Literature review

Content readings:

- Schroeder, Larry, David Sjoquist, and Paula Stephan. 1986. *Understanding Regression Analysis*. Thousand Oaks: SAGE, pgs. 12-29.
- Lewis-Beck, Michael S. 1995. *Data Analysis: An Introduction*.Thousand Oaks: SAGE, pgs. 19-30, 42-53.
- Lewis-Beck, Colin, and Michael Lewis-Beck. 2016. *Applied Regression: An Introduction*. Thousand Oaks: SAGE, pgs. 1-22.
- Kahane, Leo. 2008. *Regression Basics*. 2nd ed. Thousand Oaks: SAGE, pgs. 1-16.

R resources:

- Phillips, Nathaniel D. n.d. *YaRrr! The Pirate's Guide to R*. Chapters 13-15.4.
- Grolemund, Garrett, and Hadley Wickham. n.d. *R for Data Science*. O'Reilley. Chapters 17-18, 22-23.
- McConville, Chester Ismay and Albert Y. Kim Foreword by Kelly S. 2019. *Statistical Inference via Data Science*. New York: CRC Press. Chapter 5.

Swirl:

- Regression models: 1-4

LinkedIn:

- Computing a Linear Regression: Learning R
- Computing Contingency Tables: Learning R
- Computing Correlations: Learning R

Week 6 (FEB 17) Mid-term recess, NO CLASS

Week 7 (FEB 24) Regression assumptions & statistical inference

(intro to multivariate)

Content readings:

- Lewis-Beck, Michael S. 1995. *Data Analysis: An Introduction*.Thousand Oaks: SAGE, pgs. 54-74.
- Lewis-Beck, Colin, and Michael Lewis-Beck. 2016. *Applied Regression: An Introduction*. Thousand Oaks: SAGE, pgs. 23-54.
- Schroeder, Larry, David Sjoquist, and Paula Stephan. 1986. *Understanding Regression Analysis*. Thousand Oaks: SAGE, pgs. 30-53.
- Kahane, Leo. 2008. *Regression Basics*. 2nd ed. Thousand Oaks: SAGE, pgs. 17-58.

R resources:

- Healy, Kieren. 2018. *Data Visualization*. Princeton: Princeton UP. Chapter 6.

Swirl:

- Regression models: 5

Week 8 (MAR 3) Multiple, multivariate regression

Due: Practice assignment 3

Content readings:

- Lewis-Beck, Colin, and Michael Lewis-Beck. 2016. *Applied Regression: An Introduction*. Thousand Oaks: SAGE, pgs. 55-74.
- Kahane, Leo. 2008. *Regression Basics*. 2nd ed. Thousand Oaks: SAGE, pgs. 59-78.

R resources:

- Grolemund, Garrett, and Hadley Wickham. n.d. *R for Data Science*. O'Reilley. Chapters 24-25.

- McConville, Chester Ismay and Albert Y. Kim Foreword by Kelly S. 2019. *Statistical Inference via Data Science*. New York: CRC Press. Chapter 6.

Swirl:

- Regression models: 5 (repeat), 9

Week 9 (MAR 10) Categorical independent variables, non-linear relationships, & interactions

Due: Project research design diagram

Content readings:

- Lewis-Beck, Colin, and Michael Lewis-Beck. 2016. *Applied Regression: An Introduction*. Thousand Oaks: SAGE, pgs. 75-96.
- Schroeder, Larry, David Sjoquist, and Paula Stephan. 1986. *Understanding Regression Analysis*. Thousand Oaks: SAGE, pgs. 54-65.
- Kahane, Leo. 2008. *Regression Basics*. 2nd ed. Thousand Oaks: SAGE, pgs. 79-102.
- Hardy, Melissa A. 1993. *Regression with Dummy Variables*. Thousand Oaks: SAGE, pgs. 18-29.
- Fox, John. 1991. *Regression Diagnostics*. Thousand Oaks: SAGE, pgs. 53-65.

R resources:

- Grolemund, Garrett, and Hadley Wickham. n.d. *R for Data Science*. O'Reilley. Chapters 15, 23.4.

Week 10 (MAR 17) Outliers and predicted outcomes

Due: Project data description

Content readings:

- Lewis-Beck, Michael S. 1995. *Data Analysis: An Introduction*.Thousand Oaks: SAGE, pgs. 54-72.
- Lewis-Beck, Colin, and Michael Lewis-Beck. 2016. *Applied Regression: An Introduction*. Thousand Oaks: SAGE, pgs. 55-74.
- Schroeder, Larry, David Sjoquist, and Paula Stephan. 1986. *Understanding Regression Analysis*. Thousand Oaks: SAGE, pgs. 54-65. (repeat)
- Fox, John. 1991. *Regression Diagnostics*. Thousand Oaks: SAGE, pgs. 21-40.

R resources:

- Grolemund, Garrett, and Hadley Wickham. n.d. *R for Data Science*. O'Reilley. Chapters 23.3

- Healy, Kieren. 2018. *Data Visualization*. Princeton: Princeton UP. Chapters 5-6.

LinkedIn:

- Lines and Smoothers: Data Visualization in R with Ggplot2
- Plot to File: R for Data Science: Lunchbreak Lessons

Week 11 (MAR 24) Collinearity, F-tests, adjusted \(R^2\), & model specification

Due: Practice assignment 4

Content readings:

- Schroeder, Larry, David Sjoquist, and Paula Stephan. 1986. *Understanding Regression Analysis*. Thousand Oaks: SAGE, pgs. 66-80.
- Fox, John. 1991. *Regression Diagnostics*. Thousand Oaks: SAGE, pgs. 1-20.
- Kahane, Leo. 2008. *Regression Basics*. 2nd ed. Thousand Oaks: SAGE, pgs. 119-142.

R resources:

- Grolemund, Garrett, and Hadley Wickham. n.d. *R for Data Science*. O'Reilley. Chapters 24-25.

Swirl:

- Regression models: 10-11

Week 12 (MAR 31) Model fitting & assumptions

Due: Project results

Content readings:

- Fox, John. 1991. *Regression Diagnostics*. Thousand Oaks: SAGE, pgs. 40-53.
- Achen, Christopher H. 2005. "Let's Put Garbage-Can Regressions and Garbage-Can Probits Where They Belong." *Conflict Management and Peace Science* 22 (4): 327–39.
- Schrodt, Philip A. 2014. "Seven Deadly Sins of Contemporary Quantitative Political Analysis." *Journal of Peace Research* 51 (2): 287–300.

R resources:

- Grolemund, Garrett, and Hadley Wickham. n.d. *R for Data Science*. O'Reilley. Chapter 25.

Week 13 (APR 7) Logistics regression & other advanced models

Due: Practice assignment 5

Content readings:

- Kahane, Leo. 2008. *Regression Basics*. 2nd ed. Thousand Oaks: SAGE, pgs. 143-46.
- Menard, Scott. 2002. *Applied Logistic Regression Analysis*. 2nd ed. Thousand Oaks: SAGE, pgs. 1-67.

R resources:

- Phillips, Nathaniel D. n.d. *YaRrr! The Pirate's Guide to R*. Chapter 15.5.
- Healy, Kieren. 2018. *Data Visualization*. Princeton: Princeton UP. Chapter 8.

Swirl:

- Regression models: 11-12

LinkedIn:

- Using Colors in R: Learning R and next 4 modules are all about customizing your figures

Week 14 (APR 14) Research Project Presentations

Week 15 (APR 21) Final report due

Course Policies

Submission of Assignments

[Insert policy on format of assignments and how to be submitted]

Grades

Grades will be based on the McMaster University grading scale:

MARK	GRADE
90-100	A+
85-90	A
80-84	A-
77-79	B+
73-76	B
70-72	B-
69-0	F

Late Assignments

- *No* late *practice* assignments are accepted because these are marked based on completeness and are reviewed in class the day they are due.
- Project assignments *may* be turned in after the deadline without an evaluation penalty. However, given the scaffolded nature of class assignments, late project assignments will not be marked before additional project assignments are due. In some cases, late project assignments may not be marked until the end of term, depending on how many students turn in late work.
- Assignments will be marked in the order that they are received.

Absences, Missed Work, Illness

Regular attendance is crucial to your success in this course and is expected of all graduate students. Though attendence and participation are not formal parts of learning evaluation in this class, students who have missed even one class have had trouble catching up with the material, and students who have missed more than one class usually have had significant trouble completing the final project to their satisfaction.

Course with an On-Line Element

Some courses may use on-line elements (e.g. e-mail, Avenue to Learn (A2L), LearnLink, web pages, capa, Moodle, ThinkingCap, etc.). Students should be aware that, when they access the electronic components of a course using these elements, private information such as first and last names, user names for the McMaster e-mail accounts, and program affiliation may become apparent to all other students in the same course. The available information is dependent on the technology used. Continuation in a course that uses on-line elements will be deemed consent to this disclosure. If you have any questions or concerns about such disclosure please discuss this with the course instructor.

In this course we will be using GitHub.com to distribute course content and host a class discussion board, and GitHub Classroom to submit assignments. Site content will be restricted to course members, and student assignments will all be only visible to the instructor (and TA when there is a TA).

Copyright and Recording

Students are advised that lectures, demonstrations, performances, and any other course material provided by an instructor include copyright protected works. The Copyright Act and copyright law protect every original literary, dramatic, musical and artistic work, **including lectures** by University instructors

The recording of lectures, tutorials, or other methods of instruction may occur during a course. Recording may be done by either the instructor for the purpose of authorized distribution, or by a student for the purpose of personal study. Students should be aware that their voice and/or image may be recorded by others during the class. Please speak with the instructor if this is a concern for you.

Academic Accommodation for Religious, Indigenous or Spiritual Observances (RISO)
Students requiring academic accommodation based on religious, indigenous or spiritual observances should follow the procedures set out in the RISO policy. Students should submit their request to their Faculty Office ***normally within 10 working days*** of the beginning of term in which they anticipate a need for accommodation or to the Registrar's Office prior to their examinations. Students should also contact their instructors as soon as possible to make alternative arrangements for classes, assignments, and tests.

Academic Integrity Statement

You are expected to exhibit honesty and use ethical behaviour in all aspects of the learning process. Academic credentials you earn are rooted in principles of honesty and academic integrity.

Academic dishonesty is to knowingly act or fail to act in a way that results or could result in unearned academic credit or advantage. This behaviour can result in serious consequences, e.g. the grade of zero on an assignment, loss of credit with a notation on the transcript (notation reads: “Grade of F assigned for academic dishonesty”), and/or suspension or expulsion from the university.

It is your responsibility to understand what constitutes academic dishonesty. For information on the various types of academic dishonesty please refer to the Academic Integrity Policy**.**

The following illustrates only three forms of academic dishonesty

- Plagiarism, e.g. the submission of work that is not one’s own or for which other credit has been obtained.
- Improper collaboration in group work.
- Copying or using unauthorized aids in tests and examinations.

Conduct Expectations

As a McMaster student, you have the right to experience, and the responsibility to demonstrate, respectful and dignified interactions within all of our living, learning and working communities. These expectations are described in the Code of Student Rights & Responsibilities (the “Code”). All students share the responsibility of maintaining a positive environment for the academic and personal growth of all McMaster community members, **whether in person or online**.

It is essential that students be mindful of their interactions online, as the Code remains in effect in virtual learning environments. The Code applies to any interactions that adversely affect, disrupt, or interfere with reasonable participation in University activities. Student disruptions or behaviours that interfere with university functions on

online platforms (e.g. use of Avenue 2 Learn, WebEx or Zoom for delivery), will be taken very seriously and will be investigated. Outcomes may include restriction or removal of the involved students' access to these platforms

Academic Accommodation of Students with Disabilities

Students with disabilities who require academic accommodation must contact Student Accessibility Services (SAS) at 905-525-9140 ext. 28652 or sas@mcmaster.ca to make arrangements with a Program Coordinator. For further information, consult McMaster University's Academic Accommodation of Students with Disabilities policy.

Faculty of Social Sciences E-mail Communication Policy

Effective September 1, 2010, it is the policy of the Faculty of Social Sciences that all e-mail communication sent from students to instructors (including TAs), and from students to staff, must originate from the student's own McMaster University e-mail account. This policy protects confidentiality and confirms the identity of the student. It is the student's responsibility to ensure that communication is sent to the university from a McMaster account. If an instructor becomes aware that a communication has come from an alternate address, the instructor may not reply at his or her discretion.

Course Modification

The instructor and university reserve the right to modify elements of the course during the term. The university may change the dates and deadlines for any or all courses in extreme circumstances. If either type of modification becomes necessary, reasonable notice and communication with the students will be given with explanation and the opportunity to comment on changes. It is the responsibility of the student to check his/her McMaster email and course websites weekly during the term and to note any changes.

Extreme Circumstances

The University reserves the right to change the dates and deadlines for any or all courses in extreme circumstances (e.g., severe weather, labour disruptions, etc.). Changes will be communicated through regular McMaster communication channels, such as McMaster Daily News, A2L and/or McMaster email.