PUBPOL 750 Data Analysis for Public Policy I: Descriptive Statistics

Summer 2022

**Instructor:** Justin Savoie

**Email:** savoij2@mcmaster.ca

**Lecture:** Friday 1:00-2:30

**Office:** NA

**Office Hours:** Wednesday 3-5pm

**Contents**

[Course Description 3](#_Toc102649734)

[Course Objectives 3](#_Toc102649735)

[Required Materials and Texts 3](#_Toc102649736)

[Class Format 4](#_Toc102649737)

[Course Evaluation – Overview 4](#_Toc102649738)

[Course Evaluation – Details 4](#_Toc102649739)

[Weekly Course Schedule and Required Readings 4](#_Toc102649740)

[Week 1 (May 6 2022) 4](#_Toc102649741)

[Week 2 (May 13 2022) 5](#_Toc102649742)

[Week 3 (May 20 2022) 5](#_Toc102649743)

[Week 4 (May 27 2022) 5](#_Toc102649744)

[Week 5 (June 3 2022) 5](#_Toc102649745)

[Week 6 (June 10 2022) 6](#_Toc102649746)

[Week 7 (June 17 2022) 6](#_Toc102649747)

[Week 8 (June 24 2022) 6](#_Toc102649748)

[Week 9 (July 1 2022 Reading week no class) 6](#_Toc102649749)

[Week 10 (July 8 2022) 7](#_Toc102649750)

[Week 11 (July 15 2022) 7](#_Toc102649751)

[Week 12 (July 28 2022) 7](#_Toc102649752)

[Week 13 (August 5 2022) 7](#_Toc102649753)

[Course Policies 7](#_Toc102649754)

[Submission of Assignments 7](#_Toc102649755)

[Late Assignments 8](#_Toc102649756)

[Avenue to Learn 8](#_Toc102649757)

[Turnitin.com 8](#_Toc102649758)

[Emails 8](#_Toc102649759)

[Office hours 8](#_Toc102649760)

[University Policies 9](#_Toc102649761)

[Academic Integrity Statement 9](#_Toc102649762)

[Academic Accommodation of Students with Disabilities 9](#_Toc102649763)

[Faculty of Social Sciences E-mail Communication Policy 9](#_Toc102649764)

[Course Modification 9](#_Toc102649765)

# Course Description

The MPP-DS is preparing you for a wide variety of roles. Some of you may be interested in being data scientists or analysts. Most of you will choose roles where you are examining data to make a decision, commissioning studies, or working in partnership with data scientists/analysts to address an issue.

This course is the first of a series of three where you will learn to use the R programming language for statistical computing to analyze social scientific data. The main objective of this course is to allow you to become familiar with R and be able to use it to read, visualise, transform and analyze quantitative data, and prepare you for Data Analysis for Public Policy II and III.

This course is different from a traditional graduate course. Each 90-minute class will be divided in two. The first half will consist in a short presentation of key concepts. The second half will be self-directed and will consist of hands-on exercises.

Practicing outside class is highly recommended. I don’t expect us to have enough time to complete all exercises every week, so you are responsible to finish them after the class. R, and data analysis in general, are best learned by doing and by encountering (and then fixing) problems, so writing code every week is important.

# Course Objectives

The main objective of this course is to allow you to become familiar with R and be able to use it to read, visualise, transform and analyze quantitative data, and prepare you for Data Analysis for Public Policy II and III.

# Required Materials and Texts

This course will use two freely available textbooks. One important advantage of R, in addition to it being free, is that an important ecosystem of freely available resources has developed to help learn R. In addition to books, thousands of Youtube videos/series, MOOC, blog posts and tutorials exists for almost anything that can be done. Some will be referenced in the class, but this is mostly for further learning.

The two textbooks used in the course are:

R for Data Science (R4DS) by Hadley Wickham and Garrett Gromelund Available: <https://r4ds.had.co.nz/>

Modern Dive: Statistical Inference via Data Science (Modern) by Chester Ismay and Albert Y. Kim Available: <https://moderndive.com/>

# Class Format

Course is fully online.

Join Zoom Meeting [https://mcmaster.zoom.us/j/93488704412?pwd=ZjNOSG9BbEUycXk4aEpTT2djNkZYZz09](https://www.google.com/url?q=https://mcmaster.zoom.us/j/93488704412?pwd%3DZjNOSG9BbEUycXk4aEpTT2djNkZYZz09&sa=D&source=calendar&usd=2&usg=AOvVaw0VjY_UMuTWmG7W_yf-EbFV) Meeting ID: 934 8870 4412 Passcode: 936400

# Course Evaluation – Overview

This course if pass/fail and, like a professional development workshop, meant to be driven by the desire to learn rather than by a grade. Students who clearly go beyond expectations can be attributed pass with distinction. Daily work (in-class exercises) will not be collected and graded. There will be one homework (due week 4) and two projects (due week 6 and week 14) where you will apply the skills you have learned, and these will be collected, and feedback provided.

Homework and projects should be done individually.

# Course Evaluation – Details

Homework 1 will be assigned on May 20 and due on May 27. Homework 1 is a simple coding exercise. It counts towards 20% of the final pass/fail mark.

Project 1 will be assigned on June 10 and due on June 24. It will involve the univariate analysis of one or more variable and a write-up. It counts towards 30% of the final pass/fail mark

Project 2 will be assigned on July 15 and due on August 12. It will involve the bivariate analysis of one or more set of two variables and a write-up. It counts towards 30% of the final pass/fail mark.

Attendance and participation will count towards 20% of the final pass/fail mark. At the minimum, you should attend class, show some signs of having done the readings and participate when we do exercises.

# Weekly Course Schedule and Required Readings

## Week 1 (May 6 2022)

|  |
| --- |
| Week 1: May 6 2022 |
| Content | * Introduction
* Data Analysis MPP course series
* Installing R, R Studio
 |
| Readings | * R4DS Chapter 1
 |
| In-class exercise | * Installing R and R Studio (R4DS 1.4)
* Run all code snippets in R4DS Chapter 1
 |

## Week 2 (May 13 2022)

|  |
| --- |
| Week 2: May 13 2022 |
| Content | * Data visualisation
 |
| Readings | * R4DS Chapters 2 and 3
 |
| In-class exercise | * 3.2.4 3.3.1 3.5.1 3.6.1 3.7.1 3.8.1 3.9.1
 |

## Week 3 (May 20 2022)

|  |
| --- |
| Week 3: May 20 2022 |
| Content | * Data transformation
 |
| Readings | * R4DS Chapters 4 and 5
 |
| In-class exercise | * 5.2.4 5.3.1 5.4.1 5.5.2 5.6.7 5.7.1
 |
| Assignment | * Homework 1 assigned
 |

## Week 4 (May 27 2022)

|  |
| --- |
| Week 4: May 27 2022 |
| Content | * Importing data
* Tidy data
 |
| Readings | * Modern Dive Chapter 4
* R4DS Chapter 12
 |
| In-class exercise | * 12.2.1 12.3.3 12.4.3 12.5.1 12.6.1
 |
| Assignment | * Homework 1 due
 |

## Week 5 (June 3 2022)

|  |
| --- |
| Week 5: June 3 2022 |
| Content | * Exploratory data analysis one variable
 |
| Readings | * R4DS Chapter 6 and 7.1 to 7.4
 |
| In-class exercise | * 7.3.4 7.4.1
 |

## Week 6 (June 10 2022)

|  |
| --- |
| Week 6: June 10 2022 |
| Content | * RMarkdown
* Introducing project 1
 |
| Readings | * R4DS Chapter 27
 |
| In-class exercise | * 27.2.1 27.3.1
 |
| Assignment | * Project 1 assigned
 |

## Week 7 (June 17 2022)

|  |
| --- |
| Week 7: June 17 2022 |
| Content | * Exploratory data analysis two variables
 |
| Readings | * R4DS Chapter 7.5 to 7.8
 |
| In-class exercise | * 7.5.1.1 7.5.2.1 7.5.3.1
 |

## Week 8 (June 24 2022)

|  |
| --- |
| Week 8: June 24 2022 |
| Content | * Strings
* Factors
* Dates, time
 |
| Readings | * R4DS Chapters 14.1 to 14.3 15 and 16
 |
| In-class exercise | * 14.2.5 14.3.1.1 14.3.5.1 15.3.1 15.4.1 15.5.1
 |
| Assignment | * Project 1 due
 |

## Week 9 (July 1 2022 Reading week no class)

|  |
| --- |
| Week 9: July 1 2022 |
| Content | Reading week |
| Readings |  |
| In-class exercise |  |

## Week 10 (July 8 2022)

|  |
| --- |
| Week 10: July 8 2022 |
| Content | * Functions
* Loops
 |
| Readings | * R4DS Chapter 19
* R4DS Chapter 21
 |
| In-class exercise |  |

## Week 11 (July 15 2022)

|  |
| --- |
| Week 11: July 15 2022 |
| Content | * Working with weighted data
 |
| Readings | * Handout on weighted data
 |
| In-class exercise |  |
| Assignment | Project 2 assigned |

## Week 12 (July 28 2022)

|  |
| --- |
| Week 12: July 28 2022 |
| Content | * Working on final project
 |
| Readings | None |
| In-class exercise |  |

## Week 13 (August 5 2022)

|  |
| --- |
| Week 13: August 5 2022 |
| Content | * Working on final project
 |
| Readings | None |
| In-class exercise |  |

# Course Policies

## Submission of Assignments

All assignments are to be summitted online through Avenue to learn.

## Late Assignments

Since the course is pass/fail, I don’t have a late policy. You will not fail because your work is late. If you can’t submit either the homework or the projects on time, let me know by email. Simply try to treat all this professionally. If you have unjustified late submissions, it will reduce the probability to pass with distinction.

## Avenue to Learn

In this course we will be using Avenue to Learn. Students should be aware that, when they access the electronic components of this course, 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 this course will be deemed consent to this disclosure. If you have any questions or concerns about such disclosure please discuss this with the course instructor.

## Turnitin.com

In this course we will be using a web-based service (Turnitin.com) to reveal plagiarism. Students will be expected to submit their work electronically to Turnitin.com and in hard copy so that it can be checked for academic dishonesty. Students who do not wish to submit their work to Turnitin.com must still submit a copy to the instructor. No penalty will be assigned to a student who does not submit work to Turnitin.com. All submitted work is subject to normal verification that standards of academic integrity have been upheld (e.g., on-line search, etc.). To see the Turnitin.com Policy, please to go [www.mcmaster.ca/academicintegrity](http://www.mcmaster.ca/academicintegrity).

## Emails

I will answer emails in 48h maximum, except on Saturday and Sunday. After 48h, if I missed the email, feel free to send me a reminder.

## Office hours

I will be available to meet on Zoom Wednesday from 3pm to 5pm. Write me an email to schedule. I am quite flexible, and we can also meet by Zoom outside this time.

Zoom office hours link: https://mcmaster.zoom.us/j/9489966590

# University Policies

## Academic Integrity Statement

You are expected to exhibit honesty and use ethical behavior 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 behavior 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, located at [www.mcmaster.ca/academicintegrity](http://www.mcmaster.ca/academicintegrity).

The following illustrates only three forms of academic dishonesty:

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

## Academic Accommodation of Students with Disabilities

Students who require academic accommodation must contact Student Accessibility Services (SAS) to make arrangements with a Program Coordinator. Academic accommodations must be arranged for each term of study. Student Accessibility Services can be contacted by phone 905-525-9140 ext. 28652 or e-mail sas@mcmaster.ca. For further information, consult McMaster University’s Policy for [Academic Accommodation of Students with Disabilities.](http://www.mcmaster.ca/policy/Students-AcademicStudies/AcademicAccommodation-StudentsWithDisabilities.pdf)

## 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.