PUBPOL 750: Data Analysis I

FALL 2023

**Instructor:** Justin Savoie

**Email:** savoij2@mcmaster.ca

**Lecture:** Wednesday 4:00-5:30

**Office:** NA

**Office Hours**: on Zoom Wednesday from 3pm to 5pm (email to schedule) or by appointment.

**Contents**

[Course Description 1](#_Toc1868530471)

[Course Objectives 2](#_Toc1234085907)

[Required Materials and Texts 2](#_Toc1336877630)

[Class Format 3](#_Toc649715422)

[Course Evaluation – Overview 3](#_Toc589373897)

[Homework assignment 1 (20%), due October 4 3](#_Toc157758017)

[Project 1 (30%), due November 8 3](#_Toc819873654)

[Project 2 (30%), due December 15 3](#_Toc1144279978)

[Course Evaluation – Details 3](#_Toc533952277)

[Weekly Course Schedule and Required Readings 4](#_Toc1437297015)

[Week 1 (September 13 2023) 4](#_Toc945553546)

[Week 2 (September 20 2022) 4](#_Toc184003042)

[Week 3 (September 27 2023) 4](#_Toc1297838676)

[Week 4 (October 4 2023) 5](#_Toc902589954)

[Week of October 11 2023 – NO CLASS – Mid-term recess 5](#_Toc1850059931)

[Week 5 (October 18 2023) 5](#_Toc18328619)

[Week 6 (October 25 2023) 5](#_Toc1413236476)

[Week 7 (November 1 2023) 5](#_Toc197340908)

[Week 8 (November 8 2023) 6](#_Toc1179380486)

[Week 9 (November 15 2023) 6](#_Toc469421215)

[Week 10 (November 22 2023) 6](#_Toc1462044725)

[Week 11 (November 29 2023) 6](#_Toc1841414766)

[Week 12 (December 6 2023) 7](#_Toc686966853)

[Course Policies 7](#_Toc337813914)

[Submission of Assignments 7](#_Toc373190493)

[Grades 7](#_Toc1319103822)

[Late Assignments 7](#_Toc46546698)

[Emails 7](#_Toc228492928)

[Office hours 7](#_Toc965965915)

[Courses With An On-Line Element 8](#_Toc1156672112)

[Online Proctoring 8](#_Toc1959015927)

[Authenticity / Plagiarism Detection 9](#_Toc565499722)

[Copyright and Recording 9](#_Toc717135034)

[Academic Accommodation for Religious, Indigenous or Spiritual Observances (RISO) 9](#_Toc1003817472)

[Academic Integrity Statement 9](#_Toc858484358)

[Conduct Expectations 10](#_Toc580707312)

[Academic Accommodation of Students with Disabilities 10](#_Toc417054736)

[Faculty of Social Sciences E-mail Communication Policy 11](#_Toc843054099)

[Course Modification 11](#_Toc1576608926)

[Extreme Circumstances 11](#_Toc1271626381)

[University Policies 11](#_Toc1509690745)

[Academic Integrity Statement 11](#_Toc996563866)

[Academic Accommodation of Students with Disabilities 12](#_Toc432535537)

[Faculty of Social Sciences E-mail Communication Policy 12](#_Toc623029006)

[Course Modification 12](#_Toc1294089230)

# 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.

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 the Data Analysis in the Winter semester.

# 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 2 ed) by Hadley Wickham and Garrett Gromelund Available: https://r4ds.hadley.nz/

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

An alternative resource is the notes to this course: <https://stat545.com/index.html#other-contributors>.

In most ways, the content covered in all three resources is similar, but we will focus on R4DS 2 ed.

# Class Format

The course is fully online.

Join Zoom Meeting

<https://mcmaster.zoom.us/j/96568655841?pwd=bVRCQnZuTlNGQWkzRXZxcHV4NU1vZz09>

Meeting ID: 965 6865 5841

Passcode: 292938

# 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 assignment (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.

## **Homework assignment 1 (20%), due October 4**

## **Project 1 (30%), due November 8**

## **Project 2 (30%), due December 15**

**Participation (20%)**

# Course Evaluation – Details

Homework 1 will be assigned on September 27 and due on October 4. Homework 1 is a simple coding exercise. It counts towards 20% of the final pass/fail mark.

Project 1 will be assigned on October 25 and due on November 8. 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 November 22 and due on December 15. 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 (September 13 2023)

|  |  |
| --- | --- |
| Week 1: | |
| 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 (September 20 2022)

|  |  |
| --- | --- |
| Week 2: | |
| Content | * Data visualisation |
| Readings | * R4DS Chapters 2 and 3 |
| In-class exercise | * 2.25 2.4.3 2.5.5 2.6.1 3.5 |

## Week 3 (September 27 2023)

|  |  |
| --- | --- |
| Week 3: | |
| Content | * Data transformation |
| Readings | * R4DS Chapters 4 and 5 |
| In-class exercise | * 4.2.5 4.3.5 4.5.7 5.6 |
| Assignment | * Homework 1 assigned |

## Week 4 (October 4 2023)

|  |  |
| --- | --- |
| Week 4: | |
| Content | * Importing data * Tidy data |
| Readings | * Modern Dive Chapter 4 * R4DS Chapter 6 |
| In-class exercise | * 6.2.1 |
| Assignment | * Homework 1 due |

## Week of October 11 2023 – NO CLASS – Mid-term recess

## Week 5 (October 18 2023)

|  |  |
| --- | --- |
| Week 5: | |
| Content | * Exploratory data analysis one variable |
| Readings | * R4DS Chapter 7 and 11.1-11.4 |
| In-class exercise | * 11.3.3 11.4.1 |

## Week 6 (October 25 2023)

|  |  |
| --- | --- |
| Week 6: | |
| Content | * Quarto * Introducing project 1 |
| Readings | * R4DS Chapter 29 |
| In-class exercise | * 29.2.1 29.3.1 29.4.1 29.5.5 29.6.3 29.7.1 29.8.1 |
| Assignment | * Project 1 assigned |

## Week 7 (November 1 2023)

|  |  |
| --- | --- |
| Week 7: | |
| Content | * Exploratory data analysis two variables |
| Readings | * R4DS 11.5-11.7 |
| In-class exercise | * 11.5.1.1 11.5.2.1 11.5.3.1 |

## Week 8 (November 8 2023)

|  |  |
| --- | --- |
| Week 8: | |
| Content | * Strings * Factors * Dates, time |
| Readings | * R4DS Chapters 15, 17, 18 |
| In-class exercise | * 15.2.4 15.3.4 17.3.1 17.4.1 18.2.5 18.3.4 |
| Assignment | * Project 1 due |

## Week 9 (November 15 2023)

|  |  |
| --- | --- |
| Week 10: | |
| Content | * Functions * Loops |
| Readings | * R4DS Chapter 26 * R4DS Chapter 27 |
| In-class exercise |  |

## Week 10 (November 22 2023)

|  |  |
| --- | --- |
| Week 10: | |
| Content | * Working with weighted data |
| Readings | * Handout on weighted data |
| In-class exercise |  |
| Assignment | Project 2 assigned |

## Week 11 (November 29 2023)

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

## Week 12 (December 6 2023)

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

## Week 13 (December 13 2023)

# Course Policies

## Submission of Assignments

All assignments are to be submitted online through Avenue to Learn.

## Grades

Grades will be based on the McMaster University graduate grading scale:

|  |  |  |
| --- | --- | --- |
| **Mark** | **Grade** | **Pass/Fail** |
| 90-100 | A+ | P+ |
| 85-89 | A | P |
| 80-84 | A- |
| 77-79 | B+ |
| 73-76 | B |
| 70-72 | B- |
| 69-0 | F | F |

## 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.

## 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: \*\*\*

## **Courses 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.

**Generative AI**

**[INTRUCTORS can decide whether or not to allow use of this technology. Here is some sample text to use (pick one of the two paragraphs) if you do or do not allow GenAI tools, based on the McMaster guidelines available here <https://mi.mcmaster.ca/generative-artificial-intelligence-in-teaching-and-learning/#tab-content-provisional-guidelines>]** Students are not permitted to use generative AI in this course. In alignment with [McMaster academic integrity policy](https://secretariat.mcmaster.ca/app/uploads/Academic-Integrity-Policy-1-1.pdf), it “shall be an offence knowingly to … submit academic work for assessment that was purchased or acquired from another source”. This includes work created by generative AI tools. Also state in the policy is the following, “Contract Cheating is the act of “outsourcing of student work to third parties” (Lancaster & Clarke, 2016, p. 639) with or without payment.” Using Generative AI tools is a form of contract cheating. Charges of academic dishonesty will be brought forward to the Office of Academic Integrity.

#### *OR Some Use Permitted*

Students may use generative AI in this course in accordance with the guidelines outlined for each assessment, and so long as the use of generative AI is referenced and cited following citation instructions given in the syllabus. Use of generative AI outside assessment guidelines or without citation will constitute academic dishonesty. It is the student’s responsibility to be clear on the limitations for use for each assessment and to be clear on the expectations for citation and reference and to do so appropriately.

## **Online Proctoring**

Some courses may use online proctoring software for tests and exams. This software may require students to turn on their video camera, present identification, monitor and record their computer activities, and/or lock/restrict their browser or other applications/software during tests or exams. This software may be required to be installed before the test/exam begins.

## **Authenticity / Plagiarism Detection**

Some courses may use a web-based service (Turnitin.com) to reveal authenticity and ownership of student submitted work. For courses using such software, students will be expected to submit their work electronically either directly to Turnitin.com or via an online learning platform (e.g. A2L, etc.) using plagiarism detection (a service supported by Turnitin.com) so it can be checked for academic dishonesty.

Students who do not wish their work to be submitted through the plagiarism detection software must inform the Instructor before the assignment is due. No penalty will be assigned to a student who does not submit work to the plagiarism detection software. **All submitted work is subject to normal verification that standards of academic integrity have been upheld** (e.g., on-line search, other software, etc.). For more details about McMaster’s use of Turnitin.com please go to [www.mcmaster.ca/academicintegrity](http://www.mcmaster.ca/academicintegrity).

## **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](https://registrar.mcmaster.ca/wp-content/uploads/2019/04/RISO-Form-Examinations.pdf) 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. **It is your responsibility to understand what constitutes academic dishonesty.**

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. For information on the various types of academic dishonesty please refer to the [Academic Integrity Policy](https://secretariat.mcmaster.ca/university-policies-procedures-%20guidelines/), located at <https://secretariat.mcmaster.ca/university-policies-procedures-> guidelines/

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](https://secretariat.mcmaster.ca/app/uploads/Code-of-Student-Rights-and-Responsibilities.pdf) (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](https://sas.mcmaster.ca/) (SAS) at 905-525-9140 ext. 28652 or [sas@mcmaster.ca](mailto:sas@mcmaster.ca) to make arrangements with a Program Coordinator. For further information, consult McMaster University’s [Academic Accommodation of Students with Disabilities](https://secretariat.mcmaster.ca/app/uploads/Academic-Accommodations-Policy.pdf) 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.

# 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](mailto: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.