New Course Proposal Template

The following information is required for all new course proposals. To facilitate the review/approval process, please use the headings below (and omit the italicized explanations below each heading).

1. **Program: Kinesiology and Health Sciences**

2. **Course Number:** XXXX

3. **Credit Value:** 3 credits

4. **Long Course Title:** Data science for health research with applications in R

5. **Short Course Title:** Health Data Science

6. **Effective Session:** Winter 2024

7. **Calendar (Short) Course Description:** This course is an introduction to applying data science methods on health data for research. Students will be introduced to the different types of health data (e.g. health administrative data, electronic health records, population health surveys, etc.) while developing key research skills for predictive modelling and causal inference.

8. **Expanded Course Description:**

Lecture: The lecture sessions will delve into the myriad types of health data, from health administrative data to electronic health records and population health surveys. Focusing on data science methods, students will be introduced to key concepts in predictive modelling and causal inference. The lectures will bridge the gap between the vast and varied world of health data and the intricacies of data science methodologies, providing students with a solid understanding of how to analyze and interpret these datasets within an interdisciplinary research context.

Lab: The labs will be hands-on sessions centred around the R programming language, a primary tool for health data analysis. During these sessions, students will engage directly with real world health datasets, learning skills in data wrangling and analysis. The first lab will involve students working with a health administrative dataset, introducing them to data wrangling, visualization, and basic modelling techniques. Subsequent labs will explore more advanced data analysis techniques, such as building predictive models and being exposed to causal inference designs using health data (e.g. difference-in-differences). Throughout these labs, students will work on regular assignments (i.e. every 2 weeks) that encompass the skills learned.

Presentation and final project: Building on their newfound knowledge, students will choose a specific health issue or condition to study using available data. This could range from topics like heart disease prevalence, diabetes trends, mental health studies, or even epidemic and pandemic data analytics. Prior to their presentation, students will share a research or review article related to their chosen topic with the class, sparking an in-depth discussion. The students will submit a final project on their chosen topic demonstrating their competence in data analytics.

9. Evaluation:
Data Analysis Assignments: 60% (10% each for 6 assignments)
Presentation on selected topic (individual): 10%
Analysis based on real world health policy + term paper (individual-based project with final report no more than 5000 words): 30%

10. Integrated Courses: N/A

11. Rationale: Demand for professionals who can navigate and derive insights from health data is growing. This course offers foundational knowledge of data science methods tailored to health data within interdisciplinary research domains. Additionally, it presents students with the opportunity to engage hands-on with diverse health data types, from health administrative records to electronic health records. Further, participants will deepen their expertise in specific predictive modelling and causal inference techniques. Currently, there's a gap in courses that intricately marry data science with health research contexts at York University.

12. Faculty Resources:

Antony Chum will instruct this course. This course will be offered every year.

13. Crosslisted Courses: N/A

14. Bibliography and Library Statement:

Harrison, E., & Pius, R. (2020). R for Health Data Science (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9780367855420 (available online as a free e-book: https://argoshare.is.ed.ac.uk/healthyr_book/)

Cunningham, S. (2021). *Causal inference: The mixtape*. Yale university press. (available online as a free e-book: https://mixtape.scunning.com/)

Fiona Imlach Gunasekara, Ken Richardson, Kristie Carter, Tony Blakely, Fixed effects analysis of repeated measures data, *International Journal of Epidemiology*, Volume 43, Issue 1, February 2014, Pages 264–269, https://doi.org/10.1093/ije/dyt221

James Lopez Bernal, Steven Cummins, Antonio Gasparrini, Corrigendum to: Interrupted time series regression for the evaluation of public health interventions: a tutorial, *International Journal of Epidemiology*, Volume 50, Issue 3, June 2021, Page 1045, https://doi.org/10.1093/ije/dyaa118

Rothbard, S., Etheridge, J.C. & Murray, E.J. A Tutorial on Applying the Difference-in-Differences Method to Health Data. *Current Epidemiology Reports* (2023). https://doi.org/10.1007/s40471-023-00327-x

Bor J, Moscoe E, Mutevedzi P, Newell ML, Bärnighausen T. Regression discontinuity designs in epidemiology: causal inference without randomized trials. Epidemiology. 2014 Sep;25(5):729-37. doi: 10.1097/EDE.0000000000000138. PMID: 25061922; PMCID: PMC4162343.

Oldenburg CE, Moscoe E, Bärnighausen T. Regression Discontinuity for Causal Effect Estimation in Epidemiology. *Current Epidemiology Reports*. 2016;3:233-241. doi: 10.1007/s40471-016-0080-x. Epub 2016 Aug 5. PMID: 27547695; PMCID: PMC4978750.

15. Physical Resources:

Free software available online will be used for lab purposes (R and Rstudio).

New Course Proposal
Suggested Resource and Support

- For assistance with process, procedure, see OSAS Director/Manager
- For assistance with alignment with academic plans see your UPD/Chair or AD Learning, Teaching, and Academic Programs
- For assistance with course design see the Moodle course "Health Curriculum Toolkit" at https://moodle.yorku.ca/moodle/course/view.php?id=148738 (requires passport York username and password to access).
- For assistance with course design, teaching and learning activities please contact either of the following:
    - Yasaman Delaviz, Faculty of Health's Educational/Curricular Developer Specialist by email at ydelaviz@yorku.ca or by phone at ext. 44086
    - Lisa Endersby, Educational Developer, by email at lendersb@yorku.ca or by phone at ext. 33047.
- For assistance with designing and developing experiential education activities see either one of the EE Coordinators Anda Petro or Paola Calderon-Valdivia (eehealth@yorku.ca)
- For assistance with technology-enhanced learning in the course, please contact the Faculty of Health e-learning specialist at Learning Technology Services (hh-help@yorku.ca).
- For assistance with integrating Indigenous content into your course, please contact the Indigenous Council (ruthkg@yorku.ca & shillier@yorku.ca).
- For assistance with determining responses to questions about resources, please contact the Operations Manager and the Chair/Director in the school/department offering the course.
- For assistance with determining resources for students see the Library and Learning Commons resources:
    - SPARK (Student Papers & Academic Research Kit)
    - LinkedIn Learning
    - Academic Research & Resources

Approval Process

- The Faculty of Health Curriculum Committee will only consider new course proposals that have been approved by the unit responsible for offering the course.
- If the proposed course is to be cross-listed, integrated, listed as a course credit exclusion with another course, or listed as a major/minor course option in an interdisciplinary program, the proposal must be accompanied by a statement from the collaborating unit signaling agreement to the proposal.

Checklist of activities to be completed:

Click or tap to enter a date.	Date reviewed course proposal draft with Director/Manager of OSAS
Click or tap to enter a date.	Date reviewed course proposal draft with Associate Dean, Learning and Teaching
Choose an item.	Attached draft course outline
Reviewed information provided on the Moodle course "Health Curriculum Toolkit" at https://moodle.yorku.ca/moodle/course/view.php?id=148738 (requires passport York username and password to access).	

New Course Proposal Form

School/Department: Kinesiology and Health Science

Course Rubric and Number: KINE 4XXX

Credit Weight: 3.00 (e.g. 3.00, 6.00)

Effective Session: (e.g. Fall 2021, F/W 2021-22)

Course Title: *The official name of the course as it will appear in the Undergraduate Calendar.*

Data Science for health research with applications in R

Short Title: *Maximum 40 characters, including punctuation and spaces. The short title appears on any documents where space is limited (*transcripts *and calendar copy).*

Health Data Science

Brief Course Description: *For editorial consistency, verbs should be in the present tense and begin the description; e.g., "Analyzes the nature and extent of...,"*

This is the official description of the course as it will appear in the Undergraduate Calendar. The course description should be carefully written to convey what the course is about. If applicable, include information regarding the language of instruction if other than English.

This course is an introduction to the application of data science methods to health data within interdisciplinary research contexts. Students will be introduced to the different types of health data (e.g. health administrative data, electronic health records, population health surveys) while developing key research skills for predictive modelling and causal inference.

List course(s) where applicable:

Prerequisites:	HH/KINE 2050 or any introductory statistics course or CD approval
Corequisites:	
Cross-listed to:	
Course Credit Exclusions*:	
Integration:**	KAHS XXXX

*Course credit exclusion is a formal status accorded to pairs of courses that are recognized as having sufficient overlap in content to warrant specifically excluding students from obtaining credit for both.

**Integrated courses are graduate courses integrated (taught with) 4000-level undergraduate courses

Include the following information only if the course is: limited to a specific group of students; closed to a specific group of students; and if there is any additional information necessary for students to know before enrolling (notes section). If the course includes experiential education, such as whether the students will work with a community partner and/or if it will involve going off-campus, please include this in the notes section.

Open to:	All students
Not open to:	
Notes:	

Science Course:	**YES**	**NO**
Denotes courses in IHST, KINE or PSYC to count as science credit for BSc degree programs	X	

Section A - Course Rationale:

1. What is the rationale for creating this course (e.g., fills a gap in the curriculum, addresses a trend in the content area)?

Demand for professionals who can navigate and derive insights from health data is growing. This course offers foundational knowledge of data science methods tailored to health data within interdisciplinary research domains. Additionally, it presents students with the opportunity to engage hands -on with diverse health data types, from health administrative records to electronic health records. Further, participants will deepen their expertise in specific predictive modelling and causal inference techniques. Currently, there's a gap in courses that intricately marry data science with health research contexts at York University.

2. Describe how this new course aligns with the School/Dept and/or Faculty and/or University Academic Plans, and the United Nations (UN) Sustainable Development Goals, as applies. For more information about these plans and the SDGs, contact your UPD, Department Chair, available online resources (i.e., SDGs, at https://www.yorku.ca/unsdgs/), and/or the Associate Dean, Learning, Teaching, & Academic Programs.

Alignment with Unit and/or Faculty Plan	While introductory information on data analyses is within 2nd and 3rd year KINE courses (KINE 2050 and 3150), this course will present state-of-the-art practices in the analysis of health data using R. The emphasis of this course will be hands-on analyses with real-world data, and developing programming skills for graduate research and industry jobs. This course will expand experiential education opportunities and eLearning within the Faculty. The course is also buttressed by the faculty's newly appointed Canada Research Chair in Population Health Data Science, and the course is taught by the chairholder.
Alignment with University Academic Plan	The course supports institutional areas of research strength as laid out in York's Strategic Research Plan (SRP) including: 1) *Exploring and Interrogating the Frontiers of Science and Technology*; 2) *Building Healthy Lives, Communities and Environments*; and 3) *Advancing Fundamental Inquiry and Critical Knowledge*. It advances York's position in the areas of research opportunity identified within the SRP. Specifically, the use of big data to advance public health provision touches the *Digital Cultures* and *Healthy Individuals, Health Communities, and Global Health* themes. The digital revolution has presented an incredible opportunity to combine data sets to identify other environmental aspects that affect our health and well-being, or to evaluate the effectiveness of public health interventions. It also offers opportunities for follow-up care to vulnerable and less accessible populations through the use of apps and other social media platforms. The course will provide critical skills to support application of data science for the promotion of public health at local and international scales.
Alignment with SDG(s) (only as applies)	

3. How does this proposed course complement, align, or overlap with existing course offerings, particularly in terms of objectives and/or content? If overlap exists, please indicate the nature and extent of consultation which has taken place. If the course is to be cross-listed, integrated or listed as a course credit exclusion with another course, approval is required from all the relevant Faculties/Units.

This course will be integrated between KINE and KAHS. Approval from both the undergraduate and graduate executive committees in the School of Kinesiology and Health Sciences will be obtained.

4. What is the expected enrolment in the course? If course enrollments are below 50 please explain why.

50

Section B - Course Structure:

1. Is this course (Please select one with "X"):

	Fully face to face

X	**BLEN (blended)** – combination of virtual, asynchronous with scheduled, in-person components (instructor will define whether virtual components are synchronous or asynchronous). Note: a blended format course is usually a restructuring of class contact hours with the goal to enhance engagement and to extend access to internet-based learning opportunities" (Garrison, Vaughn, 2008).
	ONLN (online) – virtual and normally asynchronous, may include some synchronous components (instructor will define any synchronous components). Note: The Office of the Registrar supplemented this definition on Feb 1 2022 by informing staff that ONLN means "no in-person component, exams and testing will be online".
	HYFX (hyflex) – concurrent in both in-person and virtual synchronous
	Other (please describe):

2. Number of contact hours (defined in terms of hours, weeks, etc.) involved. This information is particularly important to describe for blended and online courses as it indicates whether an effective length of term is being maintained.

Approximately 1/2 (2 hours/week) of the student's learning time will be face-to-face, and 1/2 (1.5 hour/week) will be online. Online time will focus on lectures and group activities (e.g. small group discussions and case studies). Face-to-face time will focus on discussion of posted research/review articles, group presentations, and labs.

3. a) If this course is offered in a blended format, what percentage of the course will be taught online? If not blended, go to # 4.

b) In absence of scheduled contact hours (face-to-face or online), please provide an indication of the estimated time students are likely to spend engaged in learning activities online required by the course.

c) In the absence of scheduled contact hours (face-to-face or online), please describe how the course design encourages student engagement and supports students in achieving the learning outcomes.

Approximately 50% of the course will be taught online. Half of the course will be (scheduled) face-to-face and students will be able to engage with the instructor and each other during this time. In person activities will include hands-on lab activities, review of lecture materials, and presentations.

4. Indicate the planned frequency of offering and number of sections anticipated (every year, alternate years, etc.)

Every year

	YES	NO
5. Can you staff this course using current teaching capacity?	X	

If no, explain how this course will be resourced (e.g., additional hires proposed in hiring plan, etc.)

6. Please name the faculty member(s) in the school/dept who have the expertise and are willing to teach this course.

Antony Chum

7. Does the course rely on faculty from other programs to teach this course? If so, specify (proposed instructor(s) name and department and attach a letter of support from the faculty member's home school/department UPD/Chair.

No

Section C - Course Design Information:

This section provides an opportunity to describe the course, its design, and how delivery of the course content aligns with the learning outcomes, teaching activities, and assessment methods. There is also an opportunity for describing how the course applies principles of experiential education, technology enhanced learning and universal design for learning.

- **Experiential Education** remains a top priority for York University and the Faculty of Health as it offers a range of benefits for students related to academic performance, civic engagement and employability. Note that providing and facilitating opportunities for structured, critical reflection (e.g. using iclicker/REEF polling, exit cards, journal entry) is a key component of experiential education. Course directors are invited to integrate EE into their course where possible, but it is understood that some EE activities may not be feasible in every course. Go to https://health.yorku.ca/experiential-education/faculty/ to see definitions of course focused, community focused, and work focused EE, information on the benefits of EE for students and course directors, and other details.
- The integration of tools and strategies for **technology enhanced learning** (e.g. online learning management system like Moodle, use of polling technology such as iclicker/REEF and other in class technology e.g., see https://student.computing.yorku.ca/technology-used-in-courses/) may provide useful tools for encouraging in class engagement and facilitating deeper learning. For help with online and blended learning course development go to https://lts.info.yorku.ca/health/.
- Incorporating the UN SDGs facilitates inclusive and equitable quality education and promotes lifelong learning opportunities for all. Go to https://www.yorku.ca/unsdgs/toolkit/ for options to embed any of the 17 goals in course design.
- The Faculty of Health is committed to the **universal design for learning** principles, i.e., offering and ensuring a diverse array of opportunities for all learners to engage, learn, and demonstrate their knowledge. More information about Universal Design for Learning, as well as recommendations for accommodations and inclusive teaching, can be found at: http://udlguidelines.cast.org/binaries/content/assets/udlguidelines/udlg-v2-2/udlg_graphicorganizer_v2-2_numbers-no.pdf and on the Teaching Commons website. Therefore, when designing a course, be sure to consider
  - multiple means of engagement (How will diverse students access and participate in the learning & teaching activities?)
  - multiple means of representation (How will course content be presented in a variety of different ways to support different learning needs and preferences?)
  - multiple means of action & expression (What diverse ways will students be able to demonstrate their learning?)

1. Course Topics/Theories

List the key topic areas taught in this course.

The course begins with a deep dive into health data exploration in R, emphasizing data wrangling, pre-processing, and visualization. It then covers regression techniques, with an in-depth examination of linear and logistic models (from a machine learning perspective), focusing on regression predictions and comprehensive model assessment. The 2nd half of the course focuses on causal inference techniques, including Difference-in-Differences, Regression Discontinuity Design, Instrumental Variable Analyses, and Propensity Score Methods. The program culminates with a capstone project, allowing students to reinforce and apply learned concepts.

	YES	NO
Will the course have substantial Indigenous (Aboriginal)* content?		X
Will the course include Indigenous (Aboriginal)* identity as either a module or field of study?		X
Will the course include component(s) from Aboriginal Peoples' language, history, cultural, heritage, artefacts, or traditional knowledge?		X

If you answered Yes to at least one of the questions above, provide a summary and/or list of the Indigenous (Aboriginal)* content or components you are proposing to include in your course in the box below.

N/A

*The Constitution Act, 1982, section 35(2) defines Aboriginal Peoples to include all Indigenous people of Canada – Indians (Status, Non-Status or First Nations identified), Métis and Inuit people.

2. Course Teaching Objectives

Course teaching objectives are broad goals for the course.

Examples of course teaching objectives:

- Exposes students to the various methods used for investigating the structure and function of the human brain.
- Provides students the opportunity to develop and practice skills in effective communication.

List the teaching objectives for the course below:

1. Introduce students to the diverse types of health data and the importance of ethical handling, privacy, and confidentiality in health research.
2. Equip students with practical skills in data wrangling, cleaning, and exploratory data analysis in R, tailored for health datasets.
3. Immerse students in regression techniques, emphasizing the building, interpretation, and validation of linear and logistic regression models.
4. Provide an in-depth understanding of longitudinal analysis and its relevance in health data analysis.
5. Familiarize students with essential causal inference techniques, underscoring their utility in uncovering casual mechanisms from observational health data.
6. Facilitate hands-on application of learned concepts through a capstone project, enhancing students' practical skills in health data analytics.
7. Promote effective communication skills by encouraging students to present and discuss their findings, fostering a collaborative learning environment.

3. Course Student Learning Outcomes:

Learning outcomes provide a framework for assessment by stating what the learners will be able to demonstrate after completing the course. A succinct learning outcome specifies the tasks students are expected to be able to perform and the level of competence expected for the tasks. Course Learning Outcomes are observable, measurable goals for students and their learning.

Examples of course learning outcomes:

- Students will be able to correctly identify the brain's major components and gross functional areas.
- Students will be able to accurately describe the factors that impact healthy aging.
- Students will be able to critically analyze an academic journal article to determine the merits and drawbacks of the published research.

To help describe learning outcomes, consider the key questions below:

What essential knowledge, skills, and attitudes etc. should students acquire?

- How sophisticated or complex (memorization, analysis, creation, etc.) is students learning to be?
- What will students be able to do or how will they demonstrate/articulate their level of learning?
- What information is needed to be collected to verify/demonstrate students' attainment of learning outcomes?
- How informative are each of these assessment tasks to understanding the student learning process?
- Are these clearly stated and communicated to students?

More information and additional resources can be found on the Teaching Commons website.

List and number the learning outcomes for the course in the section below:

1. **Objective:** Introduce students to the diverse types of health data and the importance of ethical handling, privacy, and confidentiality in health research.
   - **SLO (1.1):** Students will be able to identify and categorize different types of health data.
   - **SLO (1.2):** Students will be able to describe the ethical implications, privacy concerns, and confidentiality protocols associated with health research data.
2. **Objective:** Equip students with practical skills in data wrangling, cleaning, and exploratory data analysis in R, tailored for health datasets.
   - **SLO (2.1):** Students will demonstrate proficiency in using R to wrangle, clean, and conduct exploratory analysis on health datasets.
3. **Objective:** Immerse students in regression techniques, emphasizing the building, interpretation, and validation of linear and logistic regression models.
   - **SLO (3.1):** Students will be able to construct and validate linear and logistic regression models using health data.
   - **SLO (3.2):** Students will interpret the results derived from regression analyses in a health context.
4. **Objective:** Provide an in-depth understanding of longitudinal analysis and its relevance in health data analysis.
   - **SLO (4.1):** Students will apply longitudinal analysis techniques to health datasets and interpret the results.
5. **Objective:** Familiarize students with essential causal inference techniques, underscoring their utility in uncovering causal mechanisms from observational health data.
   - **SLO (5.1):** Students will understand and apply causal inference techniques to observational health data.
   - **SLO (5.2):** Students will be able to discern causal relationships from correlations in health data scenarios.
6. **Objective:** Facilitate hands-on application of learned concepts through a capstone project, enhancing students' practical skills in health data analytics.
   - **SLO (6.1):** Students will integrate and apply the data analysis techniques learned throughout the course to a real-world health data project.
7. **Objective:** Promote effective communication skills by encouraging students to present and discuss their findings, fostering a collaborative learning environment.
   - **SLO (7.1):** Students will effectively communicate their analytical findings through both written reports and oral presentations.

4. Course Teaching Strategies and Learning Activities

What teaching strategies and learning activities (including experiential education) will take place as part of this course? What will students be doing each week in class? How will these activities help support students' learning as defined by the learning outcomes.

To help identify course learning activities that will help students work toward achieving intended learning outcomes, reflect on these key questions:

- How will students receive or gain the information necessary for achieving the course intended learning outcomes?
- What experiential education activities will students engage in?
- What opportunities will or could students be provided to practice the skills they will develop?
- How and when will students engage with each other, with the instructor, and/or with course content?
- If technology-enhanced learning is incorporated into the course, what activities will the students engage in?

Examples:

(This is not an exhaustive list, but rather a summary of the strategies an instructor may use to encourage and facilitate meaningful learning throughout the course)

- In class discussions
- Lecture
- Online discussion forums (e.g. in Moodle)
- Active learning strategies (e.g. think, pair, share; structured debates)
- Wikis (contribute to and curate collaborative content)
- Experiential Education (EE)- Classroom Focused Activities (e.g. guest speakers, role playing, visual media, case studies, simulations, workshops and laboratory, course-based research, etc.)
- EE- Community Focused EE Activities (e.g. community-based learning; community-based research, community service learning)

List the teaching strategies and learning activities that will be included in this course:

Laboratory demonstration and data analysis Synchronous and asynchronous presentation of material (lecture format) In class discussion of lectures and readings in groups and individually

Section D - Course Mapping and Constructive Alignment

This section is designed to help you demonstrate the connections between your student learning outcomes, teaching and learning activities, and assessment strategies. For each teaching and learning activity, please i) identify the learning outcome it will help the students achieve and ii) if the activity will include a formal, graded assessment of student learning. For EE activities, also identify iii) how you will engage students in reflection around the activity (i.e. critically examining the experience), and iv) the type of EE strategy the activity corresponds to.

			For EE Activities Only	
Teaching and Learning Activity	**Which course learning outcome/s will this activity help student achieve?**	**Will this activity include a formal, graded assessment of student learning? (Y/N)** *A detailed description of assessment and evaluation strategies will be provided in the next section.*	**How you will engage students in reflection around this activity?**	**Corresponding EE Strategy** **1-** Classroom Focused **2-** Community Focused **3-** Work Focused
Example: *1. Guest Speaker representing a community-focused agency*	*Example: Identify and critically evaluate challenges to implementing equity-informed health policies* *OR* *Learning Outcome #3*	*Example:* *N*	*Example: Think-Pair-Share- In pairs, students will discuss two key questions, and share responses with the class.*	1
In-class discussion of readings and lectures	SLO 1.1 - 7.1	*N*	*Think-Pair-Share, small group discussions, case studies*	1

Synchronous and asynchronous presentation of material (lecture format)	SLO 1.1, 1.2, 5.1, 5.2	*Y*	*Final exam*	1
Laboratory demonstration and data analysis	SLO 2.1 - 6.1	*Y*	6 data analysis assignments will be completed	1
In-class presentations	SLO 3.2, 5.2, 6.1, 7.1	*Y*	Undergraduate students will conduct group presentations and graduate students will conduct an individual presentation	1

1. If the course will not include any type of experiential education, please comment below on the rationale for not incorporating experiential education into the course.

N/A

	YES	NO
2. Will the course engage Indigenous (Aboriginal) communities (including reserves, territories, departments, or community organizations, etc) on experiential education?		*X*

If yes, please comment below on how you will or might engage Indigenous (Aboriginal) communities in experiential education

N/A

Learning/Teaching with Technology:

3. How are learning or teaching technologies incorporated into the course?

The course will utilize an eClass site as the "home base" for all lectures, course materials and connections. The course will also be teaching students to program in the R language.

4. If the course does not include any type of technology enhanced learning, please comment below on the rationale for not incorporating learning or teaching technologies in the course.

N/A

5. If the proposed course employs technology-enhanced forms of delivery (e.g., replacing in-class time with online learning activities), please identify how the integrity of the learning evaluation will be maintained (e.g., using online quizzes that randomly selects questions from a test-bank; specified time length of the test, “on-site” examinations will be required, etc.)

Course exams will be held in-person.

Assessment and Evaluation Strategies:

1. How will student learning be assessed? Please list each graded component of the proposed course including the type and percentage value of each component. Indicate which learning outcome(s) are evaluated by which assessment component.

Assessment Strategy	Percentage (%) of Final Grade	Evaluated Learning Outcome(s)
Example: Final Exam	40%	1, 2
Example: In Class Quizzes	4/10% each	1
Example: Teaching & Learning Activity #1 (Reflection) (1%)	1%	3
Data analyses assignments x6	10% x 6 = 60%	1.1, 1.2, 2.1, 3.1, 3.2, 4.1, 5.1, 5.2
Presentation	10%	3.2, 5.2, 6.1, 7.1
Analysis based on a real-world health policy (group project for undergrad up to 2000 words, individual project for graduate students up to 5000 words)	15% (worth 30% for graduate students)	2.1, 3.1, 3.2, 4.1, 5.1, 5.2, 6.1, 7.1
Final exam (only application for undergrad students)	15%	1.1 - 5.2

2. Formative feedback is just in time feedback to the students during the course that does not always count toward the final grade. This formative feedback can help the students and instructor progress towards the intended learning outcomes by providing ongoing, low stakes feedback at key points in a lesson or at milestones toward completing a major assignment.

Some examples of formative feedback include:

a) a pre-test or quiz that asks students to share what they already know about a topic
b) a think-pair-share exercise where students explore and discuss key course concepts individually, in pairs, and as part of a larger in class discussion
c) exit cards following a lecture or lesson where students are asked to indicate what they have learned and questions they still have about the topic

List the formative assessment strategies that will be used in this course below.

Group discussions about readings and student presentations.

3. If the course is to be integrated (i.e., graduate/undergraduate), please list the additional evaluation requirements for graduate students.

Graduate students will be required to present an individual presentation to the class while facilitating class discussion pertaining to their chosen topic. Graduate students will also need to write a term paper on their chosen topic which will be assessed as 30% of their final grade (taken from the weighting of the final exam and analysis based on a real-world health policy).

Bibliography:

4. Please list the required readings for the course (include ebooks, online readings, and open access resources). The reading list must contain complete bibliographical information (full name of author, title, year of publication, etc.).

Harrison, E., & Pius, R. (2020). R for Health Data Science (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9780367855420 (available online as a free e-book: https://argoshare.is.ed.ac.uk/healthyr_book/)

5. Please list any suggested readings for the course (include ebooks, online readings, and open access resources). The reading list must contain complete bibliographical information (full name of author, title, year of publication, etc.)

Cunningham, S. (2021). *Causal inference: The mixtape*. Yale university press. (available online as a free e-book: https://mixtape.scunning.com/)

6. If the course is to be integrated (graduate/undergraduate), a list of the additional readings required of graduate students must be included. If no additional readings are required, a rationale should be provided.

Cunningham, S. (2021). *Causal inference: The mixtape*. Yale university press. (available online as a free e-book: https://mixtape.scunning.com/)

Fiona Imlach Gunasekara, Ken Richardson, Kristie Carter, Tony Blakely, Fixed effects analysis of repeated measures data, *International Journal of Epidemiology*, Volume 43, Issue 1, February 2014, Pages 264–269, https://doi.org/10.1093/ije/dyt221

James Lopez Bernal, Steven Cummins, Antonio Gasparrini, Corrigendum to: Interrupted time series regression for the evaluation of public health interventions: a tutorial, *International Journal of Epidemiology*, Volume 50, Issue 3, June 2021, Page 1045, https://doi.org/10.1093/ije/dyaa118

Rothbard, S., Etheridge, J.C. & Murray, E.J. A Tutorial on Applying the Difference-in-Differences Method to Health Data. *Current Epidemiology Reports* (2023). https://doi.org/10.1007/s40471-023-00327-x

Bor J, Moscoe E, Mutevedzi P, Newell ML, Bärnighausen T. Regression discontinuity designs in epidemiology: causal inference without randomized trials. Epidemiology. 2014 Sep;25(5):729-37. doi: 10.1097/EDE.0000000000000138. PMID: 25061922; PMCID: PMC4162343.

Oldenburg CE, Moscoe E, Bärnighausen T. Regression Discontinuity for Causal Effect Estimation in Epidemiology. *Current Epidemiology Reports*. 2016;3:233-241. doi: 10.1007/s40471-016-0080-x. Epub 2016 Aug 5. PMID: 27547695; PMCID: PMC4978750.

Section E - Resources Requirement:

This section may need to be filled in with the help of your Chair/Director and operations manager:

1. Computing:

- Indicate the expected hardware, software and need for student access to computing labs, including the number of student access hours needed (e.g. access to teaching computer lab with SPSS installed; students required to bring their own device). Provide cost of software, where possible. Indicate, what the cost will be for students, if any?

Access to computer lab with R and Rstudio installed (both are open source software). Students may also install these programs to the personal computers at no cost. All in-person sessions should be held in the computer lab.

2. Classroom:

- Indicate the expected specialized classroom needs (e.g. moveable table and chairs; audio/visual equipment; WIFI to support students with bringing their own device)

Regular presentation/projector equipment will be needed for all labs and presentations.

3. Teaching Support:

• Does the course require technical support? (e.g. lab technician; UIT support). If yes, specify:	**YES**	**NO**
		x
• Does the course require a tutorial or lab in addition to lecture/seminar hours? If yes, specify and provide expected group size:	**YES**	**NO**
	x	
25		
• Does the course require marker/grader, teaching assistant, lab demonstrator etc. support above those normally allocated by the department/school offering the courses?	**YES**	**NO**
If yes, specify why and for what duties/tasks the extra support is needed:		x
• If the course includes off campus practicums/placements or field experiences, such as students working with a community partner, indicate:		
○ Will the instructor need to travel to visit the off-campus community partner(s)?	**YES**	**NO**
		X
○ Will the Experiential Education Coordinator be required to support and maintain the experiential education component while the course is being offered? If yes, please specify:	**YES**	**NO**
		X
○ Is the placement intended to be domestic or international, or both?	**Domestic**	
	International	
	Both	
• If the course is blended or online, indicate whether the support of the eLearning specialist is required?	**YES**	**NO**
If yes, please specify the type of eLearning supports you need:		**X**

4. Statements of Support (please attach these to the proposal)

For new course proposals with resource implications please provide a supporting statement from your Chair/Director of your program. The Chair/Director should indicate how resourcing will be addressed e.g., through a reallocation of existing resources,

with new/additional resources, etc.

For course proposal with impact on other programs (in the Faculty or out of the Faculty), please provide evidence of consultation and supporting statement from the other program(s).

Learning Technology Services (LTS) Statement:
If there is a technology-enhanced component to the course, a statement is required from the Learning Technology Services indicating whether resources are adequate to support the course. Requests for statements can directed to Rob Finlayson (rfinlays@yorku.ca) and Helen Brennagh (brennagh@yorku.ca). Please note, it will take two weeks to get a statement of support.

Library Support Statement:
Proposals for new courses must include a **library support statement** from the Bibliographer responsible for the relevant discipline to indicate whether resources are adequate to support the course. To request a support statement, see the list of subject and liaison librarians at http://www.library.yorku.ca/web/about-us/contact-us/liaison-librarians/.

Senate Committee on Academic Standards, Curriculum & Pedagogy

Faculty of Health
School of Kinesiology and Health Sciences

Course: KINE 4XXX - Data Science for Health Research with Applications in R
Course Webpage: None

Term: Fall 2024

Prerequisite / Co-requisite: HH/KINE 2050 or CD approval

Course Instructor

Antony Chum, PhD
301 Stong
chuma@yorku.ca Course consultation hours: By appointment only

Time and Location

Lectures	Online – Zoom address to be provided
Labs	In-Person - Computer Lab

Expanded Course Description

Lecture: The lecture sessions will delve into the myriad types of health data, from health administrative data to electronic health records and population health surveys. Focusing on data science methods, students will be introduced to key concepts in predictive modelling and causal inference. The lectures will bridge the gap between the vast and varied world of health data and the intricacies of data science methodologies, providing students with a solid understanding of how to analyze and interpret these datasets within an interdisciplinary research context.

Lab: The labs will be hands-on sessions centred around the R programming language, a primary tool for health data analysis. During these sessions, students will engage directly with real world health datasets, learning skills in data wrangling and analysis. The first lab will involve students working with a health administrative dataset, introducing them to data wrangling, visualization, and basic modelling techniques. Subsequent labs will explore more advanced data analysis techniques, such as building predictive models and being exposed to causal inference designs using health data (e.g. difference-in-differences). Throughout these labs, students will work on regular assignments (i.e. every 2 weeks) that encompass the skills learned.

Presentation and Final Project: Building on their newfound knowledge, students will choose a specific health issue or condition to study using available data. This could range from topics like heart disease prevalence, diabetes trends, mental health studies, or even epidemic and pandemic data analytics. Prior to their presentation, students will share a research or review article related to their chosen topic with the class, sparking an in-depth discussion. The students will submit a final project on their chosen topic demonstrating their competence in data analytics.

Course Objectives

This course is an introduction to the application of data science methods to health data within interdisciplinary research contexts. Students will be introduced to the different types of health data (e.g. health administrative data, electronic health records, population health surveys) while developing key research skills for predictive modelling and causal inference.

Demand for professionals who can navigate and derive insights from health data is growing. This course offers foundational knowledge of data science methods tailored to health data within interdisciplinary research domains. Additionally, it presents students with the opportunity to engage hands-on with diverse health data types, from health administrative records to electronic health records. Further, participants will deepen their

expertise in specific predictive modelling and causal inference techniques. Currently, there's a gap in courses that intricately marry data science with health research contexts at York University.

Course Text / Readings

Harrison, E., & Pius, R. (2020). R for Health Data Science (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9780367855420 (available online as a free e-book: https://argoshare.is.ed.ac.uk/healthyr_book/)

Cunningham, S. (2021). Causal inference: The mixtape. Yale university press. (available online as a free e-book: https://mixtape.scunning.com/)

Additional Course Text / Readings for Graduate Students

Fiona Imlach Gunasekara, Ken Richardson, Kristie Carter, Tony Blakely, Fixed effects analysis of repeated measures data, *International Journal of Epidemiology*, Volume 43, Issue 1, February 2014, Pages 264–269, https://doi.org/10.1093/ije/dyt221

James Lopez Bernal, Steven Cummins, Antonio Gasparrini, Corrigendum to: Interrupted time series regression for the evaluation of public health interventions: a tutorial, *International Journal of Epidemiology*, Volume 50, Issue 3, June 2021, Page 1045, https://doi.org/10.1093/ije/dyaa118

Rothbard, S., Etheridge, J.C. & Murray, E.J. A Tutorial on Applying the Difference-in-Differences Method to Health Data. *Current Epidemiology Reports* (2023). https://doi.org/10.1007/s40471-023-00327-x

Bor J, Moscoe E, Mutevedzi P, Newell ML, Bärnighausen T. Regression discontinuity designs in epidemiology: causal inference without randomized trials. Epidemiology. 2014 Sep;25(5):729-37. doi: 10.1097/EDE.0000000000000138. PMID: 25061922; PMCID: PMC4162343.

Oldenburg CE, Moscoe E, Bärnighausen T. Regression Discontinuity for Causal Effect Estimation in Epidemiology. *Current Epidemiology Reports*. 2016;3:233-241. doi: 10.1007/s40471-016-0080-x. Epub 2016 Aug 5. PMID: 27547695; PMCID: PMC4978750.

Evaluation:

Undergrad evaluation:
Data Analysis Assignments: 60% (10% each for 6 assignments)
Presentation on selected topic (in groups of 2-3): 10%
Analysis based on real world health policy (in groups of 2-3): 15% (Note: covering 1 research question, length of report ≃2000 words)
Final Exam: 15%

Graduate Evaluation:
Data Analysis Assignments: 60% (10% each for 6 assignments)
Presentation on selected topic (individual): 10%
Analysis based on real world health policy + term paper (individual): 30% (Note: covering 2-3 research questions, length of report ≃5000 words)

Graduate students will be required to present an individual presentation to the class while facilitating class discussion pertaining to their chosen topic. Graduate students will also need to write a term paper on their chosen topic which will be assessed as 30% of their final grade (taken from the weighting of the the final exam and analysis based on a real-world health policy).

Grading:

The grading scheme for the course conforms to the 9-point grading system used in undergraduate programs at York (e.g., A+ = 9, A = 8, B+ - 7, C+ = 5, etc.). Assignments and tests* will bear either a letter grade designation or a corresponding number grade (e.g. A+ = 90 to 100, A = 80 to 90, B+ = 75 to 79, etc.)
(For a full description of York grading system see the York University Undergraduate Calendar - http://calendars.registrar.yorku.ca/2012-2013/faculty_rules/AP/grading.htm

"Final course grades may be adjusted to conform to Program or Faculty grades distribution profiles."

Missed Exams:

Students with a documented reason for missing a course test, such as illness, compassionate grounds, etc., which is confirmed by supporting documentation (e.g. doctor's letter) may request accommodation from the Course Instructor. Further extensions or accommodation will require students to submit a formal petition to the Faculty.

Important Course Information:
All students are expected to familiarize themselves with the following information, available on the Senate Committee on Curriculum & Academic Standards webpage (see Reports, Initiatives, Documents) http://secretariat-policies.info.yorku.ca/.

Search for:

1. Academic Honesty (Policy)
2. Research Involving Human Participants (Policy)
3. Academic Accommodation for Students with Disabilities (Policy)
4. Disruptive and /or Harassing Behaviour in Academic Situations (Policy and Procedures)
5. Religious Observance Accommodation (https://w2prod.sis.yorku.ca/Apps/WebObjects/cdm.woa/wa/regobs)

NOTE: ANY derogatory emails are regarded as harassment and will be directed to the Office of Student Conduct. No cell phones, iPods, iPads, earphones, etc. are permitted during the exams.