
SPPH 581C Special Topics
Methods for Analyzing Routinely Collected Data
Winter 2022/2023 (Jan 2023 – March 2023)

Instructor: Sutherland, Jason M.

Abstract: It is anticipated that this course will follow regular in-person classroom instruction. Classes will be held on Monday mornings, 9am PST, unless otherwise discussed with the class.

This is a data-driven course that focuses on applying supervised and unsupervised machine learning methods and non-standard analytic problems with healthcare data. The data that will serve as the motivation will be large clinical and administrative databases commonly used in health services research in Canada, such as hospital discharge data. Students in this course will be exposed to, and apply, advanced statistical methods for analyzing sophisticated healthcare-based data problems. Permission of the instructor is required to register for the course.

Drawing from machine learning applications and traditional statistical methods, students will encounter topics including:

- Cluster Analysis
- Principal Components
- Missing Data Problems
- Classification and Regression Trees
- Mixture Models
- Spline Regression Models
- Additive Models

Acquisition of methods will be based on problem-based learning. During the course, students will be introduced to new analytic concepts, progress through the principles of advanced methods, learn the adjuvant analytic techniques with software tools (SASR), identify resources to assist with the development of their skills, and synthesize their learnings by applying the methods to observational datasets, interpreting their results and sharing their findings with their peers.

Progressing through the statistical methods, students will acquire methods for data manipulation, learn coding in SAS, data cleaning, summarizing data, preparing brief reports and presenting findings to peers. SAS will be used for at least one-half of the programming; be prepared to learn SAS. Register for SAS on Demand for academics (SPPH 581C is registered).

While there are no formal textbooks for this course, the instructor will provide readings and references to students. Students will also be responsible for identifying and evaluating resources they found helpful in learning.

Learning Objectives:

1. Importing large administrative datasets;
 2. Manipulating, merging and summarizing administrative datasets;
 3. Cleaning datasets and identifying inconsistent or incomplete variables ;
 4. Develop SAS programming experience;
 5. Expand repertoire of statistical methods commonly used in machine learning;
 6. Implementing and interpreting advanced statistical methods to complex datasets;
 7. Effective oral and written communication of the findings of analyses;
 8. Class presentations of analytic assignments.
-

Prerequisites:

- Multivariate statistical methods. SPPH 500 (or equivalent) and data analytic experience.
 - Exposure to an analytic programming language such as SAS or R.
 - Permission of the instructor.
-

Course Structure:

- Lectures
 - Assignments
 - Student presentations
-

Grade Structure:

	Marks Available	Learning Objective
Assignments	7 X 12% = 84%	1 to 6
Student participation	6%	7
Presentations	10%	7 and 8

Grading System: Numeric.

Note: No late assignments will be accepted. Assignments are due to be emailed to the instructor by 5pm of the due date.

Extensions will only be granted for extenuating circumstances. Students should be aware of what behaviours constitute plagiarism. This form of academic misconduct is subject to penalties described in the Student Discipline section of the UBC calendar (available online at: www.ubc.ca).

Course Grading Criteria:

A-level work: outstanding quality

- A+ reserved for the one or two pieces of exceptional work that far exceed or extend the quality of contributions available in the literature.
- A suggests that there is a very high level of scholarship throughout every aspect of the work. Work deserving of an A is distinguished in every aspect. This level of work demonstrates that the individual has gone well beyond what has been provided and has extended the usual ways of thinking or performing. Work of this level demonstrates outstanding comprehension of the subject and use of existing literature and research. The student shows a very high degree of engagement with the subject.
- A- suggests that there is generally a high quality throughout the work, no problems any significance and evidence of attention is given to each criterion. The work demonstrates a very good comprehension of the subject and use of existing literature and research. The student shows a very high degree of engagement with the subject.
-

B-level work: good quality with no major weaknesses

- B+ suggests there is generally very good quality throughout the work, few problems of minor significance and evidence of attention given to each criterion. The work demonstrates a good comprehension of the subject and use of existing literature and research. For the most part, the work integrates critical and creative perspectives toward the subject material and shows a fair amount of engagement with the topic.
- B suggests there is generally good quality to aspects of the work, few problems of minor significance. Attention is given to several criterion. The work demonstrates a good comprehension of the subject and use of existing literature and research. The work demonstrates few examples of integrating critical and creative perspectives towards subject material and shows a fair degree of engagement with the topic.
- B- suggests there is some aspects of good quality to the work, some problems of minor significance. Attention is given to several criterion. The work demonstrates a good comprehension of the subject and use of existing literature and research. The work demonstrates few examples of integrating critical and creative perspectives towards subject material and shows a fair degree of engagement with the topic.
-

C-level work: adequate work

C suggests there is generally adequate quality to the work, several problems of some significance. Attention given to few criterion. The work demonstrates fair comprehension of the subject and use of existing literature and research. The work demonstrates few examples of integrating critical and creative perspectives toward the subject material and minimal engagement with the topic.
