

Sketch of a Data Science minor at UBC

DSCI Minor committee: Gabriela Cohen-Freue, Mike Gelbart, Rachel Pottinger, Tiffany Timbers

Rationale

Data Science is a growing field and is becoming increasingly popular with students. For example, the UBC Master of Data Science program has received over 1000 applicants this cycle. Currently, we do not have an undergraduate Data Science program, except for the newly minted DSCI 100: Introduction to Data Science. We aim to put together a data science minor. The target audience is any UBC undergraduate student.

Disclaimer: this document is meant as a first pass at one *possible* data science minor. The goal is to get feedback. It should not be construed anything like a final version, nor as a commitment on either the Computer Science or Statistics departments to offer or create such a minor, nor can either department agree to create or offer any additional courses to additional students without additional resources.

Course requirements

The prerequisites for the DSCI minor are 1st year data science, 2nd year programming and 2nd year statistics and 1st year calculus. These are shown in the table below, under “Prerequisite” type. In addition to the prerequisites, the minor requires 6 courses (18 credits) of upper-level courses. These 6 courses are broken down into 2 required areas and 4 elective areas; for each area, there may be multiple UBC courses that fulfill the requirement.

The two required areas are statistical modeling and machine learning. These are shown in the table below, under “Required” type. For the remaining 4 courses, students must complete a course in 4 different areas, such as Visualization, Databases, Ethics, etc. Currently, there are 6 elective areas listed, but we may add more in the future.

Area	Recommended courses	Example acceptable alternatives	Type
Data science	DSCI 100		Prerequisite
Programming	CPSC 103, 203	CPSC 110 (or 107) and CPSC 210	Prerequisite
Statistical inference	STAT 2xx (new course)	STAT 302 and STAT 200	Prerequisite
Math	One of MATH 101, MATH 103, MATH 105, MATH 121, SCIE 001		Prerequisite
Statistical modeling	STAT 3xx (new course)	STAT 306	Required
Machine learning	CPSC 330	CPSC 340 or STAT 406	Required
Visualization	DSCI 3xx.a (new course)	CPSC 4xx (new vis course for majors, requires CPSC 310) or STAT 545 & 547	Counts toward minor
Databases	CPSC 3xx (new course)	CPSC 304	Counts toward minor
Reproducible Workflows	DSCI 3xx.b (new course)		Counts toward minor
Ethics	DSCI 4xx (new course)	CPSC 430	Counts toward minor
Cloud Computing	CPSC 4xx (new course)		Counts toward minor
In-discipline data science courses	(we will compile a list of acceptable courses across the University)	n/a	Counts toward minor

Note that for all of the new courses, we are not saying that the code necessarily be DSCI. We are deferring that to a later conversation. The course numbering is only to differentiate the courses, not to imply anything about the eventual ordering.

In general, we expect students to be taking courses from the “Recommended course” column. However, we recognize that we will not be able to develop all of these new courses at once. Therefore, for each area we have made a list of other courses (“Example acceptable alternatives” column) that students can take in the meantime. Over time we aim to shift the program more and more towards the recommended courses.

In some cases, these alternative courses will be inaccessible to most students. For example, the alternative Visualization course (CPSC 3xx.a) will require CPSC 310 as a prerequisite, and so is mainly just accessible to CS majors. For non-CS majors in the Data Science minor, if the recommended course has not yet been created and the acceptable alternatives are totally inaccessible, we expect students to not pursue that particular area.

As an example path, consider a student in Psychology who is interested in the Data Science minor. For the sake of the example, let's say new courses have been created in statistical inference, statistical modeling, reproducible workflows, and databases. Then, this example student might take the following courses:

- Prerequisite (data science): DSCI 100
- Prerequisite (programming): CPSC 103 and 203
- Prerequisite (statistical inference): STAT 2xx (new course)
- Prerequisite (math): MATH 103
- Statistical modeling: STAT 3xx (new course)
- Machine learning: CPSC 330
- Databases: CPSC 3xx (new course)
- Ethics: CPSC 430
- Workflows: DSCI 3xx.b (new course)
- In-discipline: PSYC xxx (new or existing course to be approved for inclusion into the minor)

Examples of potential in-discipline data science courses:

EOSC 442: Climate Measurement and Analysis

PSYC 359: Advanced Research Methods in Behavioural Sciences

MICB 405: Bioinformatics