

**COURSE NAME:** DAT350 Data Analysis Tools

---

Credit Value: 4  
Total Course Hours: 56  
Prerequisite Course(s): None  
Corequisite Course(s): None

## COURSE DESCRIPTION

---

Once data has been collected, validated and manipulated, the next step in the process is to conduct the analysis that supports business decisions. Use-case analysis of Hadoop implementation will be explored to better understand how this open-source distributed processing framework supports data analysis. Students will use the R programming language to manipulate, calculate and display data in order to gain insights and make key decisions.

## PLAR INFORMATION

---

This course is eligible for Prior Learning Assessment and Recognition. Students are advised to discuss options with their program coordinator.

## COURSE LEARNING OUTCOMES

---

Upon completion of this course, the student will have reliably demonstrated the ability to:

- 1.0 Assess the strengths and limitations of proprietary and open source data analytic tools.
  - 1.1 Illustrate how specific analytic tools are used to support business decision making.
  - 1.2 Discuss the process of selecting the correct analysis tools to align with a specific problem.
  - 1.3 Determine criteria against which to assess data analysis tools.
  - 1.4 Compare common methods and tools to efficiently parse, query and display raw data sets.
  - 1.5 Identify the strengths and limitations of scripting languages that support data analysis.
  - 1.6 Discuss the opportunity cost of automating a process or using certain tools.
  - 1.7 Recommend tools for data analysis, problem solving and decision-making that would be most appropriate for different functions and industries.
- 2.0 Use Hadoop to support data analytics.
  - 2.1 Explain the challenges of working with big data that Hadoop addresses.
  - 2.2 Summarize the Hadoop Architecture, Ecosystem and the Hadoop distributed file system (HDFS).
  - 2.3 Explain how Hadoop is used for efficient data storage, processing, analysis, security and data deployment.
- 2.4 Analyze Hadoop use-cases and case studies.
- 2.5 Demonstrate Hadoop Shell Commands and cluster modes.
- 2.6 Deploy Hadoop in a data analysis project.
- 3.0 Demonstrate knowledge of how R Studio can be used in analytic projects.
  - 3.1 List important features of the R and Github ecosystems.
  - 3.2 Install R studio and its packages.
  - 3.3 Perform basic operations in R such as using variables, Arithmetic operations, and reading/writing strings.
  - 3.4 Use objects, factors, arrays and matrices in simple R operations.
  - 3.5 Perform calculations on arrays and matrices.
- 4.0 Use an integrated statistical software suite such as R to manipulate, calculate and display data in order to gain insights and make key decisions.
  - 4.1 Read data from different sources into R.
  - 4.2 Use R to clean, analyze, and visualize data.
  - 4.3 Accomplish effective data handling and storage using the statistical package.

- 4.4 Acquire, import, and export data in various formats using R.
- 4.5 Utilize R for handling missing and erroneous data.
- 4.6 Make use of R debugging tools.
- 4.7 Complete a data analysis project using R.

## GENERAL EDUCATION

---

This is not a General Education course.

## PROGRAM OUTCOMES

---

This course contributes to the following Ministry of Colleges and Universities approved program learning outcomes (PLO):

### Enterprise Analysis and Research

1. Identify, collect, organize, manipulate and analyze data to support problem solving, organizational decision-making and opportunity identification.
2. Utilize current technologies to manipulate data sets, correlate information and communicate results in order to support strategic decisions.
3. Prepare and present complex materials visually, verbally, in writing and digitally for a variety of audiences, purposes and levels of detail using specialized software.
4. Conduct data analysis and research in a respectful and ethical manner that protects privacy and maintains dignity to all involved.
5. Use statistical and predictive models that employ operational and marketing data to identify patterns and provide insights to stakeholders.
8. Apply project management principles to gather information, build solutions and sustain analytics projects.

## ESSENTIAL EMPLOYABILITY SKILLS OUTCOMES

---

This course contributes to the following Ministry of Colleges and Universities approved essential employability skills (EES) outcomes:

3. Execute mathematical operations accurately
4. Apply a systematic approach to solve problems
7. Analyse, evaluate, and apply relevant information from a variety of sources.

## EXTERNAL COURSE ACCREDITATIONS AND CONDITIONS

---

## COURSE EVALUATION

---

- 40% In-Class Exercises
- 30% Tests
- 30% Projects

## PROGRAM SPECIFIC GRADING

---

Per College Grading System

### GRADING SYSTEM

---

A+:	90-100%	B+:	77-79%	C+:	65-69%	D:	50-54%	S - Satisfactory
A:	85-89%	B:	73-76%	C:	60-64%	F:	0-49%	I - Incomplete
A-:	80-84%	B-:	70-72%	D+:	55-59%			F- Repeat Course, included in GPA
								FS- Failure Supplemental
								FR- Repeat course, excluded from GPA

\*For a complete chart of grades and descriptions, please see the Grading Policy.

## LEARNING RESOURCES

---

Other Resources:

Required:

Hadoop Illuminated

Author: Mark Kerzner, Sujee Maniyam, 2016

Creative Commons Attribution-NonCommercial-Share-Alike 3.0

[http://hadoopilluminated.com/hadoop\\_illuminated/hadoop-illuminated.pdf](http://hadoopilluminated.com/hadoop_illuminated/hadoop-illuminated.pdf)

R Notes for Professionals

<https://books.goalkicker.com/RBook/RNotesForProfessionals.pdf>

The R Notes for Professionals book is compiled from Stack Overflow Documentation, the content is written by the beautiful people at Stack Overflow. Text content is released under Creative Commons BY-SA. 2018.

Online Resources

Cognitive Class - Hadoop 101 (Free Online Course)

<https://cognitiveclass.ai/courses/introduction-to-hadoop>

Learning Hadoop using LinkedIn videos

<https://www.linkedin.com/learning/learning-hadoop>

Resources listed on the course outline support the achievement of learning outcomes, and may be used throughout the course to varying degrees depending on the instructor's teaching methodology and the nature of the resource.

## LEARNING ACTIVITIES

---

Use Case Analysis, Lectures, Lab Exercises, Group Work

## DELIVERY MODE

---

This course may be delivered, in whole or in part, in a number of modalities, including in class, online, hybrid, in a synchronous or asynchronous manner or a combination thereof, as per accreditation and/or regulatory standards where appropriate.

## ACADEMIC POLICIES

---

Canadore College is committed to the highest standards of academic integrity, and expects students to adhere to these standards as part of the learning process in all environments. The College's Academic Integrity policy seeks to ensure that all students understand their rights and responsibilities in upholding academic integrity and that students receive an accurate and fair assessment of their work. Please review the Academic Integrity policy (A-18) and other academic policies found on our website:  
<https://www.canadorecollege.ca/about/policies>.

## COLLEGE POLICIES

---

- Protecting human rights in support of a respectful college community

For college policies please see: <http://www.canadorecollege.ca/about-us/college-policies>.

## STUDENT SUCCESS SERVICES - Your Success Matters!

---

Student Success Services provides student-focused services to facilitate students' success in their studies. Staff provide support by reducing and/or removing educational-related barriers through individualized accommodations and supports to students with disabilities.

Please visit our webpage to learn more: <https://www.canadorecollege.ca/support/student-success-services> or look for our events on social media.

To connect with Student Success Services email [studentsuccessnow@canadorecollege.ca](mailto:studentsuccessnow@canadorecollege.ca) or call 705.474.7600 ext 5205.

## FIRST PEOPLES' CENTRE:

A culturally safe environment offering CONFIDENTIAL student focused services, drop in or make an appointment to access:

- One on one counselling

- Elder in residence program
- Peer tutoring
- Peer mentorship
- Lunch & learn workshops on study skills, self-care, life skills
- Learning Resource Centre

Drop by our offices at C254 College Drive, E101 Commerce Court or call 705 474 7600 Ext. 5961 College Drive / 5647 Commerce Court.

<https://www.canadorecollege.ca/experience/indigenous-student-experience>

### **WAIVER OF RESPONSIBILITY**

---

Every attempt is made to ensure the accuracy of this information as of the date of publication. The college reserves the right to modify, change, add, or delete content.

### **HISTORICAL COURSE OUTLINES**

---

Students use course outlines to support their learning. Students are responsible for retaining course outlines for future use in applications for transfer of credit to other educational institutions.