Course Outline

DATA3102
Data Wrangling
Fall 2023 - Current

Course Outline

DATA3102 Data Wrangling

COURSE DESCRIPTION

Data wrangling is an incredibly important step in a data analysis project. Once data has been collected, it must be prepared – reviewed, cleaned, structured, and enriched – prior to analysis. In this course learners profile a dataset, reshape the data structure, identify and clean data issues (such as missing data, duplicates, outliers, coding issues), and enrich data through augmentation, aggregation, or calculation. The result of data wrangling is a finalized dataset that is well documented and ready for analysis. Additional Performance Standards: This course, along with Data Acquisition, make up the Data Acquisition and Wrangling competency; competency assessments in both courses must be successfully completed to be deemed competent. In order to be successful in this course, learners must be competent in the Math for Data Analytics and Data Programming course outcomes.

REQUISITES	None
EQUIVALENTS	None
CREDITS	3
HOURS	45
ELIGIBLE FOR PLAR	No
ZERO TEXTBOOK COST	Yes

COURSE COMPETENCY

COMPETENCY TITLE
Data Management and Analytics – Acquiring and Wrangling Data

COMPETENCY STATEMENT

Data Management and Analytics professionals must select, acquire, analyze and prepare datasets that meet project requirements and quality standards

COMPETENCY DESCRIPTION

The quality of a data analysis project is limited by the quality of the data used. Therefore, the acquisition of relevant data and its subsequent transformation into a usable, reliable dataset is a key job function for a data analyst or scientist. In order to meet project objectives, a data analyst must first select and locate meaningful and relevant data, and then perform a quality analysis to ensure that it is accurate, reliable, and effectively structured. With the emergence of big data and related technologies, there is an abundance of easily accessible data, but not all data is created (or managed) equally. Data analysts and scientists must apply a range of techniques to select, acquire and wrangle structured and unstructured data to produce datasets that are ready for analysis. Data is becoming an increasingly important driver of operations across

many high-value sectors and industries making data acquisition and wrangling a high-value job function for data professionals.

The Acquiring and Wrangling Data competency makes up part of the Data Management and Analytics job profile along with:

- Designing Data Projects
- Performing Data Analysis
- Visualizing and Presenting Insights

COURSE LEARNING OUTCOMES

Bow Valley College is committed to ensuring our graduates can demonstrate their abilities in key areas that will make them effective citizens and encourage their development as lifelong learners. In addition to the discipline-specific skills that learners acquire in their programs, the College has identified ten learning outcomes.

College-Wide Outcomes:

1. Communication
2. Thinking Skills
3. Numeracy and Financial Literacy
4. Working with Others
5. Digital Literacy
6. Positive Attitudes and Behaviours
7. Continuous Learning
8. Health and Wellness Awareness
9. Citizenship and Intercultural Competence
10. Environmental Sustainability

#	COURSE LEARNING OUTCOME(S)	COLLEGE WIDE OUTCOMES SUPPORTED
1	Wrangle data into a finalized, fit-for-purpose dataset by demonstrating judgement and problem-solving skills.	1, 2, 3, 5, 6

LEARNING PATHWAY

The time it takes learners to demonstrate competencies will vary. An example of a suggested schedule for learning and development is shown below. Learners will need to plan out their assessment attempts

within their course. For additional information, please consult the Course Offering Information in Brightspace.

WEEK/HOURS LEARNING AND DEVELOPMENT PLAN

Week 1	Review Performance Demonstration Competency Assessment, and, Learning Pathway
Week 2	Create Success Plan; Performance Demonstration
Week 3	Competency Assessment; Determine data wrangling requirements
Week 4	Restructure data to facilitate analysis
Week 5	Clean data to address data issues
Week 6	Clean data to address data issues
Week 7	Enrich data through augmentation, aggregation, or calculation
Week 8	Enrich data through augmentation, aggregation, or calculation
Week 9	Reading Week
Week 10	Enrich data through augmentation, aggregation, or calculation
Week 11	Enrich data through augmentation, aggregation, or calculation
Week 12	Validate data
Week 13	Finalize dataset
Week 14	Performance Demonstration
Week 15	Competency Assessment

COURSE MODULES AND SCHEDULE

Course schedule subject to change, depending on delivery mode and term of study. For exact dates, please consult the Course Offering Information in Brightspace.

WEEK/HOURS	MODULES
1	Evaluate data sources
2	Evaluate data sources
3	Clean a data set
4	Clean a data set
5	Clean a data set
6	Restructure a data set
7	Restructure a data set
8	Time series data
9	Reading week
10	Time series data
11	Text data
12	Text data
13	Documentation
14	Web scraping and API's
15	Project

ASSESSMENT

This course follows an assessment-first approach, in which learners will be assessed, and receive structured feedback, and a personalized learning plan. Learners will also receive differentiated support from an instructor based on their individual needs.

Learners will have a variety of ways to demonstrate they have met the required competency through the demonstration of learning outcomes and criteria as laid out in the rubric. Learners will have multiple (but not unlimited) attempts to prove competency. It is suggested that learners plan out their assessment attempts within their course.

Learners will have flexibility in how they satisfy course learning outcomes while still adhering to the criteria found in the rubric and the Course Offering information. Please refer to the Course Offering Information and the rubric in Brightspace for additional information.

COURSE LEARNING OUTCOMES	ASSESSMENT
1	Performance Demonstration

ASSESSMENT

COURSE LEARNING OUTCOME(S)	ASSESSMENT	WEIGHT
1	Assessment Projects (Minimum of 5)	60%
1	Learning Activities	20%
1	Discussions	20%

Important: For details on each assignment and exam, please see the Course Offering Information.

PERFORMANCE STANDARDS

A minimum grade of D is required to pass this course. However, a program may require a higher grade in this course to progress in the program or to meet specific program completion requirements.

Please consult with the program area or contact the program chair for further details. A minimum Grade Point Average of 2.0 is required for graduation.

GRADING SCHEME

Grade	Percentage	Grade Point	Description
A+	95-100	4.0	Exceptional: superior knowledge of subject matter
A	90-94	4.0	Excellent: outstanding knowledge of subject matter
A-	85-89	3.67	
B+	80-84	3.33	
B	75-79	3.0	Very Good: knowledge of subject matter generally mastered
B-	70-74	2.67	
C+	67-69	2.33	

			Satisfactory/Acceptable: knowledge of subject matter adequately mastered
C	64-66	2.0	
C-	60-63	1.67	
D+	57-59	1.33	
D	50-56	1.0	Minimal Pass
F	Less than 50	0.0	Fail: an unsatisfactory performance

REQUIRED LEARNING RESOURCES

Additional learning resources may be found in the Course Offering Information or in Brightspace.

ADDITIONAL INFORMATION

Additional information may be found in the Course Offering Information or in Brightspace.

ACADEMIC ACCOMMODATIONS

Learners with a disability (learning, physical, and/or mental health) may qualify for academic and exam accommodations. For more information, or to apply for accommodations, learners should make an appointment with Accessibility Services in the Learner Success Services (LSS) Department. Accessibility Services can also assist learners who may be struggling with learning but do not have a formal diagnosis. To make an appointment visit LSS on the first floor of the south campus or call 403-410-1440. It is the learner's responsibility to contact Accessibility Services and request academic accommodations. For more information, please visit our website at http://www.bowvalleycollege.ca/accessibility.

INSTITUTIONAL POLICIES

Bow Valley College is committed to the highest standards of academic integrity and honesty. Learners are urged to become familiar with and uphold the following policies: Academic Integrity (500-1-7), Learner Code of Conduct, Procedures and Guidelines (500-1-1), Learner Appeals (500-1-12), Attendance (500-1-10), Grading (500-1-6), Academic Continuance and Graduation (500-1-5), and Electronic Communications (300-2-13). Audio or video recording of lectures, labs, seminars, or any other teaching and learning environment

by learners is allowed only with consent of the instructor as part of an approved accommodation plan. Recorded material is to be used solely for personal study and is not being used or distributed without prior written consent from the instructor.

Turnitin:

Students may be required to submit their course work to Turnitin, a third-party service provider engaged by BVC. Turnitin identifies plagiarism by checking databases of electronic books and articles, archived webpages, and previously submitted student papers. Students acknowledge that any course work or essays submitted to Turnitin will be included as source documents in the Turnitin.com reference database, where it will be used solely to detect plagiarism. The terms that apply to a student's use of Turnitin are described on Turnitin.com.

Online Exam Proctoring:

Examinations for this course may require proctoring through an online proctoring service. Online proctoring enables online exam taking within a controlled and monitored environment, thereby enhancing academic integrity. Online proctoring may occur through a variety of methods, including but not limited to:

a. live online proctoring where a remote invigilator authenticates identity and observes completion of an exam using specialized software and recordings;
b. automated proctoring where the exam session is recorded and AI (artificial intelligence) analyzed;
c. browser lockdown that limits access to other applications, websites, copying, printing, screen capture and other functions; or
d. a combination of both live/automated proctoring and browser lockdown.

Course instructors will review recordings, analyses, and data obtained through online proctoring for academic integrity infractions. It is the student's responsibility to meet the technical, software, location, and identity verification requirements necessary to enable online proctoring.

Further details of these policies are available in the Academic Calendar and on the Bow Valley College website, bowvalleycollege.ca.

Learners are encouraged to keep a copy of this course outline for future reference.

Collection of Personal Information:

This course, including your image and voice, may be recorded and made available to you and other students taking the course section. By attending the class(es) online or in person, you consent to the collection of your personal information. If you do not wish to be recorded, please contact your instructor before starting the course/class to discuss alternative arrangements.

You may use the recordings only for educational purposes and you must not copy, share, or use the recordings for any other purpose without the instructor's express permission.

Your personal information is collected in accordance with section 33(c) of the Freedom of Information and Protection of Privacy Act (Alberta) to deliver academic programming, support learner flexibility, promote

universal design for learning principles, and for purposes consistent with the course activities and outcomes. If you have any questions about the collection, disclosure, use, or protection of this information, please contact the College's Access and Privacy Officer at foip@bowvalleycollege.ca.