



Course Syllabus

Course Information

Course Number:	ALY6110
Course Name:	Data Management and Big Data
CRN:	80524
Quarter / Year:	Spring 2021 CPS Quarter – Second Half
Term Length	6 weeks
Start / End Dates	05/24/2021 – 07/03/2021
Where	Hayden Hall, Room 012
When	Mondays, 5:50pm - 8:30pm EST
Course Format:	Hybrid NUFlex

Instructor Information

Full Name:	Valeriy Shevchenko
NEU Email Address:	v.shevchenko@northeastern.edu
Office Hours	Virtual, will reply within 12 hours

NuFLEX Requirements

Should this course fall under the Hybrid NuFlex policy, please refer to any and all supplemental materials pertaining to class attendance, participation, and other aspects impacting student and/or instructor engagement. The NUFlex course modality is the same as on-ground, and has been designed to reflect varying student density in class and the flexibility for students to have synchronous learning from a remote location given the current circumstances resulting from the COVID-19 outbreak. For additional information, please refer to your advisor.

Technical Requirements

Courses are available on Northeastern University's **Canvas** at the following link:

<http://canvas.northeastern.edu>. Canvas Technical support and resources including 24/7 phone (1-833-450-3937), and chat can be found on the help icon in Canvas. Be sure to click the Help link in Canvas for technical requirements and support resources. Northeastern Technical Support can be reached at 617-373-4357 (xHELP) or help@northeastern.edu.

Each student is responsible for his or her access to the internet for purposes of this course and for research. Internet access is a required component of this course and will not be accepted as an excuse for missed work. If you know that you will be traveling, then make sure you plan accordingly.

Note regarding e-mail/voicemail: If you e-mail me, please include **class number and CRN number into the email subject**. Please allow up to 24 hours for an email reply. All class correspondence will go to your assigned NEU email account. Any emails being sent to the whole class will go to your NEU email. It is mandatory you use your NEU/Husky email account when contacting your instructor. If you leave me a voicemail, please remember to include your name, class title, CRN Number and phone number.

Course Prerequisites

There are no prerequisites for this class.

Course Description

In this course you will see the core concepts of data collection and management. Topics include systems for collecting data and implications for practice; types of data (textual, quantitative, qualitative, etc.); storing data with privacy and security issues in mind; including a discussion of business process

outsourcing. Course offers students an opportunity to obtain a high-level understanding of big data technologies for data accessibility, efficiency, and security of data management at scale, including big data storage and computing technologies and big data analytics applications. Students create a working system for data acquisition and management using publicly available data.

Course Materials

Reading

We will not use a textbook for this course. Instead, we will assign weekly readings from varied sources. See Canvas for more details.

Software & Tools:

You might need to use virtualization software (e.g., VMWare's Workstation/Player or Oracle's VirtualBox or Microsoft Hyper-V Manager) on a computer with at least 4GB of RAM (it is recommended to have 8GB for better performance). You will have an access to Microsoft Azure Windows VM provided by NEU to perform assignments and final project as needed.

Program Student Learning Outcomes (PLOs)

Specialized Knowledge	Broad and Integrative Knowledge	Applied and Collaborative Learning	Civic and Global Learning	Experiential Learning
PLO1	PLO2	PLO3	PLO4	PLO5
Integrate the major theories, tools, and approaches in data analytics to identify and successfully communicate data-driven insights for informed decision making.	Articulate and effectively defend the significance and implications of the work in data analytics in terms of challenges and trends in a local, national or global context.	Apply the principles, tools and methods of analytics to a comprehensive real world problem or project related to data analyses for tactical and/or strategic decision making; present data, information and/or analytical insights and recommendations for successful implementation of the project.	Propose an effective path to resolution of an analytical problem that may be complicated by the competitive environment, opposing interests, divergent or uncertain data and information.	Apply the principles, tools and methods of analytics to a project within a sponsoring organization to successfully assist with the extraction, development, delivery, and/or translation /implementation of data analysis for tactical and/or strategic decision making in organizations.

In this course students are *developing* skills that relate to Program SLOs 1-5.

<i>Enter 5 for Central, 4 for Significant, 3 for Moderate, 2 for Minimal, 1 for Potential, or 0 for None</i>					<i>Enter 1 for Passive Engagement, 2 for Active Engagement, or 3 for Generative Engagement</i>
Social Consciousness & Commitment	Global Mindset	Intellectual Agility	Personal & Professional Effectiveness	Well-Being	Level of Engagement
2	2	5	4	1	3

Course Learning Outcomes (CLOs)

CLO1	CLO2	CLO3	CLO4	CLO5	CLO6	CLO7	CLO8
Design a Hadoop File System (HFS) for a large data mart to meet the needs for modern commercial operations.	Evaluate how Big Data solutions that utilize data lakes to store unstructured data.	Discuss methodologies utilized to analyze big data sets and determine insights.	Assess the use of big data storage and computing technologies using big data analytics application.	Analyze the use of big data technologies for data accessibility, efficiency, and security of data management at scale.	Develop a working process for data acquisition and management ensuring that data is consistent, secure, reliable and well linked, with an end to end integration.	Compare and contrast business solutions for constructing a system in house or outsourcing.	Introduction to the capabilities of Spark

Expectations

- One (1) academic credit requires 50 minutes a week of classroom or faculty instruction and about two hours of out of class student work for a 15-week course; 100 minutes a week of classroom or direct faculty instruction and about 3.5 hours of out of class student work for a 7.5-week course.
- For a three-credit course, students should expect 2.5 hours a week of classroom or faculty instruction and a minimum of 5 hours of out of class student work for a 15-week course; 5 hours of classroom or direct faculty instruction and a minimum of 10 hours of out of class student work for a 7.5-week course.
- APA citations

Attendance Policy

Student will attend six classes online or on-site at dates and times as specified in Course Information section. Attendance will be recorded and stored using CPS Attendance Tracking Tool.

Policy on late work

Any delays in submission without prior notification will negatively affect your score. However, not all prior notifications will be accepted by the instructor, therefore, avoid requesting extra time unless absolutely necessary. Any submissions after assignment availability date or for assignments that are already graded will not be considered unless prior arrangements had been discussed and agreed upon.

Course Methodology

Each week, you will be expected to:

1. Review the week's learning objectives.
2. Complete all assigned readings.
3. Complete all lecture materials for the week.
4. Participate in the Discussion Board.
5. Complete and submit all assignments and tests by the due date.

Participation/Discussion Board

Discussion board minimum requirements are:

- At least 1 (one) meaningful primary response is due by 5:00 PM EST on or before the day of the next class.
- At least 2 (two) meaningful secondary responses are due by 5:00 PM EST on or before the day of the next class.

To facilitate interaction, students are expected to review the online postings on a regular basis even after they have posted their own minimum required postings. Please treat your classmates and the instructors with the utmost respect. Inappropriate posts will be removed immediately. The instructor reserves the right to penalize students for repeated violations of the participation policy (and/or Academic Integrity Policy) within a course. In the discussion board and in class, high quality contributions advance the class discussions and do not simply summarize the material that was assigned. Quality contributions take into account not only the instructor's questions but also your classmates' contributions. Please be mindful that the Discussion Board is a space for academic exchanges. As a result, students are accountable for using proper and exacting punctuation, spelling, and grammar. In addition, you may be required to reference all outside sources in correct citation format. It is crucial that all participants maintain a high regard for proper decorum in the Discussion Board.

The following are the expectations for successful grades for the discussion board assignments:

- Participation in the discussion board: each week students must post **at least one primary response** (new thread to the discussion board) and **two secondary responses** (responses to other students' posts).
- The primary response should consist of a minimum of 250 words, and replies should contain at least 50 words.
- Last minute postings that are empty of substance and essence will result in significant point deductions.
- When responding to the Discussion forum, learners should support their comments with logical reasoning and with the techniques of data analysis. Simply stating that agree or disagree without further describing why will not be accepted as a valid discussion.
- Copying and pasting from any source into the discussion board is unacceptable.
- Attachments in discussion board only allowed as a references and cannot substitute primary or secondary response content

Assignment Submission and Communication

In the Assignments folder, click on the View/Complete Assignment link to view and each assignment. Attach your completed assignments here and click Submit to turn them in to me. Once your assignment has been graded, you will be able to view the grade and feedback I have provided by clicking on My Grades in the Tools module from the Northeastern University Online Campus tab. You are also encouraged to post content oriented communication in the Discussion Board instead of in e-mail so that all students can benefit from each other's learning. Your PDF or WORD submission file **must** contain at least the following sections:

- Title page (dedicated page)
 - Your name (as it is registered with this class).
 - Class and CRN number.
 - Class name.
 - Assignment name.
- Summary section explaining what this assignment is about.
- Introduction section explaining:
 - Input data.
 - Proposed methods and/or techniques.
 - Goals of the research.
 - Expected outcomes.
- Any results or outputs provided by research and /or code (screen shots, charts, graphs, etc.).
- Each results must have a meaningful explanation directly related to its purpose with this research.
- Final conclusions and explanation of the goals and findings from the data standpoint.

Final Project

Students will work on a project during the course and will present the results on the last class. You will choose the most appropriate **big data** technology to perform the analysis and execute it. The project will consist of:

- (i) A project description and implementation plan in word format
- (ii) A data manipulations using Big Data Management system only
- (iii) A basic data analysis implementation and report
- (iv) A presentation of your project

Evaluation Standards

The course grade will be determined on the following basis:

Assignment	Due Date	Total	Total %
5 Quizzes (100 points)	Week 1 through week 5: <ul style="list-style-type: none">Available from the beginning of the class @ 5:50PM EST. Due by the end of the class @ 8:30PM EST	100	10%
5 Weekly Discussions (40 points each)	WEEKLY: <ul style="list-style-type: none">One primary posting for each discussion thread due on or before the beginning of the next class on Monday @ 05:00 PM ESTTwo Secondary Responses for each discussion due on or before the beginning of the next class on Monday @ 05:00 PM EST	200	30%
5 Assignments (100 points each)	WEEKLY: <ul style="list-style-type: none">Due before the beginning of the next class on or before Monday @ 05:00 PM EST	500	30%
1 Collaborative Project (200 points)	<ul style="list-style-type: none">Project Presentation file submission due on or before the beginning of Week 6 class on 06/28/2021 @ 5PM EST (50 points)Week 6 Discussion board main project post with a summary and list of group members (one per group) due on or before the beginning of Week 6 class 06/28/2021 @ 5:00PM EST (25 points)Week 6 Discussion board secondary replies for each presented project main thread due before the end of the Week 6 class 06/28/2021 @ 8:30PM EST (25 points)Final Project Paperwork submission due on Monday 06/28/2021 @ 11:59 PM EST (100 points)	200	30%
Total		1000	100%

Grading Scale:

The final grade reported will be broken down by the following scale:

95-100%	A	87-89.9%	B+	77-79.9%	C+	69.9% or below	F
		84-86.9%	B	74-76.9%	C		
90-94.9%	A-	80-83.9%	B-	70-73.9%	C-		

Course Schedule

Week	Date	Topic	Assignments	COs/SLOs
1	05/24 – 05/30	Introduction to Big Data	Quiz due by the end of the class (worth 20 points), Module 1 Discussion board (due before the next class on Monday @ 5PM EST, worth 40 points), Module 1 Assignment (due before the next class on Monday @ 5PM EST, worth 100 points)	PLO1, CLO1
2	05/31 – 06/06	Big Data Projects Methodology	Quiz due by the end of the class (worth 20 points), Module 2 Discussion board (due before the next class on Monday @ 5PM EST, worth 40 points), Module 2 Assignment (due before the next class on Monday @ 5PM EST, worth 100 points)	PLO2, CLO2
3	06/07 – 06/13	Methods Used in Big Data Analysis	Quiz due by the end of the class (worth 20 points), Module 3 Discussion board (due before the next class on Monday @ 5PM EST, worth 40 points), Module 3 Assignment (due before the next class on Monday @ 5PM EST, worth 100 points)	PLO3, CLO3, CLO4
4	06/14 – 06/20	Introduction to Spark	Quiz due by the end of the class (worth 20 points), Module 4 Discussion board (due before the next class on Monday @ 5PM EST, worth 40 points), Module 4 Assignment (due before the next class on Monday @ 5PM EST, worth 100 points)	PLO3, CLO7
5	06/21 – 06/27	Big Data Storage	Quiz due by the end of the class (worth 20 points), Module 5 Discussion board (due before the next class on Monday @ 5PM EST, worth 40 points), Module 5 Assignment (due before the next class on Monday @ 5PM EST, worth 100 points)	PLO4, CLO5
6	06/28 – 07/03	Data Lakes, Final Project	Group Project: 1. Discussion Board Project main thread post with Project summary due on or before the beginning of the last class @ 5 PM EST (worth 25 points). 2. Secondary Post for each presented project is due before the end of the last class @ 8:30 PM EST (worth 25 points). 3. Presentation Paperwork due before the beginning of the last class @ 5PM 4. Live Presentation of the project to the class due before the end of the last class @ 8:30 PM EST (worth 50 points). 5. Final Project Paperwork due on Wednesday 06/28/2021 @ 11:59PM EST (worth 100 points)	PLO5, CLO6, CLO7, CLO8

End-of-Course Evaluation Surveys

Your feedback regarding your educational experience in this class is very important to the College of Professional Studies. Your comments will make a difference in the future planning and presentation of our curriculum.

At the end of this class, please take the time to complete the evaluation survey at the NEU Evaluation Kit website: <https://new.evaluationkit.com>. Your survey responses are **completely anonymous and confidential**. Surveys will be open for the last two weeks of the class. An email will be sent to your HuskyMail account notifying you when surveys are available.

If you do not pass the systems check or have any questions or concerns, you can contact Examity's® technical support team 24/7 via email at support@examity.com or phone at (855) 392-6489. Please tell your instructor immediately if your computer/equipment does not meet the standard to use online proctoring.

Academic Integrity

A commitment to the principles of academic integrity is essential to the mission of Northeastern University. The promotion of independent and original scholarship ensures that students derive the most from their educational experience and their pursuit of knowledge. Academic dishonesty violates the most fundamental values of an intellectual community and undermines the achievements of the entire University.

As members of the academic community, students must become familiar with their rights and responsibilities. In each course, they are responsible for knowing the requirements and restrictions regarding research and writing, examinations of whatever kind, collaborative work, the use of study aids, the appropriateness of assistance, and other issues. Students are responsible for learning the conventions of documentation and acknowledgment of sources in their fields. Northeastern University expects students to complete all examinations, tests, papers, creative projects, and assignments of any kind according to the highest ethical standards, as set forth either explicitly or implicitly in this Code or by the direction of instructors.

Go to <http://www.northeastern.edu/osccr/academic-integrity-policy/> to access the full academic integrity policy.

CPS expects independent and original work. Dishonesty and/or carelessness violates fundamental values of an intellectual and professional community and will be reported to the Office of Student Conduct and Conflict Resolution (OSCCR) as a violation of the Academic Integrity Policy. To safeguard the integrity of assignments and programs, your course may use systems such as TurnItIn, which checks written work, and Examity, an online exam proctoring system. Work that contains academic integrity violations (AIV) will be graded on a case by case basis. Work displaying AIV may earn a failing grade, a zero, or even result in the instructor assigning the student a failing grade for the class. Students may not withdraw from a class to avoid the grade penalty for serious academic integrity violations.

For additional information or for answers to frequently asked questions, please visit: <https://cps.northeastern.edu/academic-resources/academic-integrity>

You may also email CPS-AcademicIntegrity@northeastern.edu.

Student Accommodations

Northeastern University and the Disability Resource Center (DRC) are committed to providing disability services that enable students who qualify under Section 504 of the Rehabilitation Act and the Americans with Disabilities Act Amendments Act (ADAAA) to participate fully in the activities of the university. To receive accommodations through the DRC, students must provide the Disability Resource Center (DRC) with appropriate documentation that demonstrates a current substantially limiting disability.

For more information, visit <http://www.northeastern.edu/drc/getting-started-with-the-drc/>.

Library Services

The Northeastern University Library is at the hub of campus intellectual life. Resources include over 900,000 print volumes, 206,500 e-books, and 70,225 electronic journals.

For more information, visit <http://library.northeastern.edu/>.

Wyzant Tutoring Services

Northeastern University's College of Professional Studies is covering the cost of 4 hours of tutoring from Wyzant to help ensure your success in challenging courses. Tutors on Wyzant have helped thousands of students in 300+ subjects ranging from College Algebra to Academic Writing, Statistical Analysis to Microbiology. [Click here](#) to access your free tutoring through Northeastern University's Wyzant account.

Access your free tutoring using your Northeastern login credentials. Indicate your program, time zone, the course you'd like support in, and the specific topics or materials you'd like help with.

You'll receive an activation email from Wyzant for Higher Ed. (If you don't see it, check your spam or junk folders.) From the email, click on the activation button.

Potential tutors will begin reaching out to you. Their messages will be sent to your Northeastern email inbox and will be accessible via your Wyzant account dashboard.

Respond back to the tutors you believe might be a good fit to schedule an online session when it's convenient for you — even late at night!

24/7 Canvas Technical Help

For immediate technical support for Canvas, call 1-833-450-3937. You can chat and report an issue to Canvas technical support right from within Canvas through the Help icon located in the Canvas global navigation.

myNortheastern, e-mail, and basic technical support

Visit the [Information Technology Services \(ITS\) Support Portal](#)

Email: help@northeastern.edu

ITS Customer Service Desk: 617-373-4357 (help)

myNortheastern, e-mail, and basic technical support

Visit the [Information Technology Services \(ITS\) Support Portal](#)

Email: help@northeastern.edu

ITS Customer Service Desk: 617-373-4357

Catalog

The College of Professional Studies Undergraduate Catalog is a reference/resource with information about curricula, resources, and academic and student policies.

For more information, visit <http://www.cps.neu.edu/student-resources/>.

Diversity and Inclusion

Northeastern University and the Office of Institutional Diversity and Inclusion (OIDI) is committed to equal opportunity, affirmative action, diversity and social justice while building a climate of inclusion on and beyond campus. In the classroom, member of the University community work to cultivate an inclusive environment that denounces discrimination through innovation, collaboration and an awareness of global perspectives on social justice.

Please visit <http://www.northeastern.edu/oidi/> for complete information on Diversity and Inclusion

Title IX

Title IX of the Education Amendments of 1972 protects individuals from sex or gender-based discrimination, including discrimination based on gender-identity, in educational programs and activities that receive federal financial assistance.

Northeastern's **Title IX Policy** prohibits sex and gender-based discrimination, including sexual harassment, sexual assault, sexual exploitation, relationship or domestic violence, and stalking. The Title IX Policy refers to sex and gender-based discrimination as "Prohibited Offenses." The Title IX Policy applies to the entire Northeastern community, including students, faculty and staff of all gender identities.

Allegations of Prohibited Offenses can be reported to the Title IX Coordinator within **The Office for University Equity and Compliance** at: titleix@northeastern.edu and/or through **the Northeastern University Police Department (NUPD)** by phone: for an Emergency 617.373.3333; for Non-Emergency 617.373.2121. Reporting to NUPD does **NOT** commit the victim/affected party to future legal action.

Faculty members are considered "responsible employees" at Northeastern University, meaning they are required to report all reports of and information about alleged Prohibited Offenses to the Office for University Equity and Compliance.

If you or someone you know has experienced a Prohibited Offense, **confidential** support and guidance can be found through **ViSION Resource Center (VRC)** staff

(<http://www.northeastern.edu/vision/visionresourcecenter/>), **University Health and Counseling Services (UHCS)** staff (<http://www.northeastern.edu/uhrs/>) and the **Center for Spiritual Dialogue and Service (CSDS) clergy members** (<http://www.northeastern.edu/spirituallife/>). Employees within the VRC, UHCS, and CSDS are not required to report allegations of Prohibited Offenses to the Office for University Equity and Compliance.

In case of an emergency, please call 911 OR NUPD'S Emergency line: 617-373-3333.

Please visit www.northeastern.edu/titleix for a complete list of reporting options and **resources, both on- and off-campus.**

Northeastern University Online Copyright Statement

Northeastern University Online is a registered trademark of Northeastern University. All other brand and product names are trademarks or registered trademarks of their respective companies.

This course material is copyrighted and Northeastern University Online reserves all rights. No part of this publication may be reproduced, transmitted, transcribed, stored in a retrieval system, or translated into any language or computer language, in any form or by any means, electronic, mechanical, magnetic, optical, chemical, manual, or otherwise, without the express prior written permission of Northeastern University Online.

Copyright 2021 © by Northeastern University Online
All Rights Reserved

The instructor reserves the right to amend this syllabus, both online and the document itself, during the term and will notify students of the change(s). The revised syllabus is the official record of class policies and schedule of due dates.