



NORTHWEST
MISSOURI STATE UNIVERSITY
MARYVILLE | KANSAS CITY

School of Computer Science and Information Systems

**44-517-01 and 44-517-02, Big Data, 3 credit hours
Fall 2022**

Instructor: Denise Case

Email: dcase@nwmissouri.edu

Office hrs: MWF 9-9:50am, 3:00-3:50pm, Tu 11-11:50am

Office: CH 2280

Phone: 660-562-1136

Prerequisites: Undergraduate prerequisites: MATH 17230 or MATH 17316, both with a grade of C or better and CSIS 44242 with a grade of C or better. Graduate prerequisite: CSIS 44542 with a grade of B or better, or concurrent enrollment in CSIS 44542, or consent of instructor. (F)

Textbook and supplementary materials: There is no required textbook. Free resources will be required as designated by the instructor.

Course description: An introduction to the design of data-intensive, reliable, scalable, and maintainable systems. Includes an introduction to current and relevant tools, technologies, design principles, and frameworks. This may include concepts such as parallel programming, distributed computing, distributed file systems, MapReduce, regular expressions, and the ingesting and processing of data at rest and data in motion. Tools used may include Hadoop, HDFS, Pig, Hive, Spark, Storm, Kafka, Mahout, MLlib, etc.

Student learning outcomes and assessment methods:

Competency	Program Outcome	Assessment
Managing Information	DSI students will access, generate, and reorganize information using contemporary technologies.	Selected assignment(s)
Teamwork	DSI Students will work as a team to design, implement, and deliver solutions to problems using best practices with contemporary technologies.	Selected assignment(s)

Additionally, after successful completion of this course, students should be able to:

1. Describe the characteristics of big data systems
2. Describe the history of distributed data processing including Hadoop, the HDFS file system, the MapReduce programming model, and Google PageRank.
3. Describe big data integration platforms such as Apache Beam.
4. Compare choices for data processing engines including frameworks such as Apache Spark, Apache Flink, and managed services such as Google Cloud Dataflow
5. Describe key processing concepts including sources, sinks, pipelines, and transformations.
6. Describe common options for programming languages.
7. Describe common options for I/O connectors.
8. Create and execute big data jobs using Apache Beam.
9. Function effectively on teams to accomplish a common goal

Instructional methods: Instructional methods may include lectures, class discussions, individual work, small group work, learner presentations, guest speakers, and collaborative development.

Course Expectations: This course includes the development of unique, creative applications. Students should expect to spend at least 2-3 hours on out-of-class work for every hour of instruction.

Course Email: Professional email practices are required. **Email must include course number and section (e.g., 44-517-01) in the subject line.** Email should be professional, politely written, and use proper grammar, capitalization, and spelling.

Course Conduct Guidelines: Students will behave in a professional manner in all course interactions. Devices including cell phones, earphones, gadgets, and other items should be silent and put away during class. Unauthorized use of devices during class may result in reduction in scores. Be ready to work when the class is scheduled to begin. Be polite and respectful to others in the class. Do not interrupt someone during a class discussion, and respect others' opinions/comments as you expect them to respect yours. Follow classroom policies as listed here and in the course site.

Course citation requirements: Cite ALL Sources. You must provide proper citation for any content including text, images, videos, and code that you did not personally develop. Do not copy and paste content for any submission. Failing to acknowledge assistance including internet photos, clipart, slide content, course files, shared solutions, etc. can be expected to result in a 0 on the assignment, a one-half letter grade reduction, a 0 in the course, or failing the course with an academic integrity violation.

Course late work policy: Graded activities are expected to be completed by due date. If you must miss a due date, contact the instructor in advance or as early as possible. Due dates get you ready for coming material. Keeping up is critical. On rare occasions, special challenges or obligations may make it difficult to submit on time. You typically have a one-week grace period where the work may still be submitted and graded. Late penalties will be described in the rubric. **No late submissions will be accepted during the last two weeks of class.**

Grading: Grades will generally be posted within 2-7 days of the due date. If you wish to petition for additional points, please wait 24 hours. Petitions must be emailed, and must include the specific changes you are requesting, along with your reasoning on why the higher score is warranted. Challenges must be raised within seven days of the grade being posted; after seven days, grades are considered final.

Graded course requirements: All submissions must be your own, creative, unique work. You may be asked to explain your work during class, in discussions, or in a recorded submission.

Graded contributions:

Category	Contribution
Engagement	10%
Group Big Data Project	10%
Quizzes	30%
Projects and exercises	50%
All	100%

Grading scale:

Undergraduate	Graduate	
Percent Range	Percent Range	Grade
88-100%	90-100%	A
>= 78% and < 88%	>= 80% and < 90%	B
>= 68% and < 78%	>= 70% and < 80%	C
>= 58% and < 68%	>= 60% and < 70%	D
below 58%	below 60%	F

Course outline/major topics studied:

Module	Content
1	Course intro, big data, shell processing, history
2	Python, map-filter-reduce, architectures
3	Big data engines and Java apps
4	Apache Beam - Choose Runner/Lang/Group
5	Big Data Pipelines, Transforms, Sources, Sinks
6	Group Beam projects
7	Kafka, optional certs, final reflection

Course schedule:

Week	Start	Notes
1	Wed 1-12	Course starts Wed, assignments due Day 1, Module 1
2	Mon 1-17	No class Monday, Module 1
3	Mon 1-24	Module 2
4	Mon 1-31	Module 2
5	Mon 2-7	Module 3
6	Mon 2-14	Module 3
7	Mon 2-21	Module 4
8	Mon 2-28	Module 4
9	Mon 3-7	No class, spring break
10	Mon 3-14	Module 5
11	Mon 3-21	Module 5
12	Mon 3-28	Module 6
13	Mon 4-4	Module 6
14	Mon 4-11	Module 6
15	Mon 4-18	Module 7
16	Mon 4-25	Module 7
17	Mon 5-2	Module 7 - Finals week (no late submissions)

Note: Course schedule is subject to change with instructor notification and students will be responsible for abiding by these changes.

Attendance: Students are expected to attend all classes as specified in the course syllabi for each course. It is the responsibility of the student to promptly notify his or her instructor when unable to attend class. Please refer to the university policy on attendance at

<https://www.nwmissouri.edu/policies/academics/Attendance.pdf>

Final exams: If an emergency occurs that prevents the administration of a course scheduled final examination, the final course grades will be calculated based on the work in the course completed to that point in time and the faculty member's considered judgment. Final exams will not be rescheduled, and a grade of "I" will not be given as a result of an institutional cancellation of a final examination. This final exam policy does not apply to online courses.

Administrative drop: An instructor may request the Office of the Registrar delete a student from a course roster if the student has not met the prerequisite for the course as stated in the catalog, or as a result of non-attendance in the course.

Academic Administrative withdrawal: When it is in the best interest of Northwest Missouri State University for a student to withdraw, a student will be given a W, put on administrative hold, and given notice that they are about to be withdrawn. This action will result in removal of all credits associated with courses that have yet to be completed in the semester in question and the student will be administratively withdrawn from the University. An Administrative Withdrawal does not affect the student's grade point average. Please refer to Northwest Missouri State University's Academic Administrative Withdrawal Policy at:

<https://www.nwmissouri.edu/policies/academics/Adding-Dropping-Withdrawals.pdf>

University communications: Students are expected to use their Northwest student email account for any electronic correspondence within the university. Students are also strongly advised to check their email and CatPAWS accounts on a regular basis.

Academic integrity policy: The students, faculty, and staff at Northwest endeavor to sustain an environment that values honesty in academic work, that acknowledges the authorized aid provided by and intellectual contributions of others, and that enables equitable student evaluation. Please refer to Northwest Missouri State University's Academic Integrity Policy at <http://www.nwmissouri.edu/policies/academics/Academic-Integrity.pdf>

Learning or Living Accommodations Request Process: Northwest Missouri State University complies with Section 504 of the Rehabilitation Act of 1973 and the Americans with Disabilities Act of 1990 [ADA] and the ADA Amendments Act of 2008 [ADAAA]. If a student has a documented disability that qualifies under the ADA/ADAAA and requests accommodations, they should review the Accessibility and Accommodations webpage at <https://www.nwmissouri.edu/titleixequity/accessibility/index.htm> for guidance, including the accommodations application and supporting documentation requirements. Contact ada@nwmissouri.edu or 660.562.1873 for further assistance. For the university policy on disability accommodation refer to <http://www.nwmissouri.edu/policies/student/Disability-Accommodation.pdf>

Non-discrimination and anti-harassment policy: Northwest Missouri State University is committed to maintaining an environment for all faculty, staff, students, and third parties that is free of illegal discrimination and harassment. Please refer to the Non-Discrimination and Anti-Harassment Policy at <https://www.nwmissouri.edu/TitleIXEquity/titlevi/>

Family Education Rights and Privacy Act (FERPA) policy: Family Educational Rights and Privacy Act of 1974, as amended (commonly known as the Buckley Amendment), is a federal law which provides that colleges and universities will maintain the confidentiality of student education records. Please refer to the Family Educational Rights and Privacy Act (FERPA) Policy at <http://www.nwmissouri.edu/policies/academics/Family-Educational-Rights-and-Privacy-Act.pdf>

COVID-19 Classroom Mitigation: Northwest is committed to the health and safety of the University community. Face coverings are required in all instructional spaces (classrooms/labs/studios) for all students and faculty until campus vaccination and infection data allow the university to direct otherwise. Should a non-vaccinated faculty member or student make close contact with a symptomatic, COVID-19 positive student, they will be asked to quarantine and not enter University facilities (unless otherwise instructed by University Wellness) or attend class until that quarantine period is over. Faculty will provide means for students missing class because of COVID-19 quarantine or isolation to keep up with course work, but students must communicate with faculty and adhere to the quarantine/isolation start and end dates. All other illnesses or absences for personal reasons will be handled as they were before the pandemic. Northwest further asks all students and faculty to practice good hygiene and not enter University facilities or attend face-to-face classes when they feel sick. Students who do not comply with these requirements will be subject to standard disciplinary procedures according to the Northwest Student Code of Conduct (i.e. verbal and written warnings followed by a hearing, if necessary). We thank you for doing your part to maintain our learning environment and to protect the health of fellow Bearcats.

Change in Course Delivery:

It is our goal as a University to continue all courses as planned on campus. However, due to unforeseeable impacts of COVID-19 and its variants, faculty and students must be prepared to adjust course formats or mitigation strategies anytime during the semester, either permanently or for a short term. The University and faculty will communicate with students in the event such action is deemed necessary to preserve the health and safety of students and employees.

Quarantine checklist:

YOUR to do List:

- Call 911 if you experience difficulty breathing or a change in your level of consciousness.
- Contact Wellness Services at 660.215.8573 with questions. You may text or call between 7 a.m. and 7 p.m. If you have an emergency outside of these hours, contact UPD at 660.562.1254.
- Notify all of your instructors of your absence and predicted date of return. They will assist you in your coursework by providing information on their strategies for distance learning. Please ensure that you have your laptop, notes, textbooks, and other necessary course materials when you enter into quarantine. Lack of engagement in your coursework during quarantine will likely result in lower grades or failure. When communicating with your instructors, you may provide them with the reason you cannot be in the classroom, but you are not required to do so.
- Stay away from all people.
- You must talk with a Wellness Services staff member every few days, either by phone, text, Zoom or Skype. If you develop COVID-19 symptoms, you MUST notify Wellness Services immediately. It is important for your wellbeing that Northwest maintains contact with you. We want to know you are doing okay, answer any questions you have, check for symptoms and assist with your needs.

Behavior and Wellness: Northwest focuses on student success—every student, every day.

The Wellness Center, 660.562.1348, offers free counseling for students coping with depression, anxiety, alcohol or drug misuse, relationships, and other emotional, social, and academic stressors. In addition, faculty, staff, and students who are concerned with student wellbeing can report their concerns, including anonymously, to the Behavioral Intervention Team, so that the student can be offered relevant support, at this link: [Concerning Behavior Reporting Form](#). If you are concerned about the immediate safety of a student, please call the University Police, who have specialized training in intervention, at 660.562.1254.

Disclaimer: Course schedule is subject to change and you will be responsible for abiding by any such changes. Your instructor will notify you of any changes.