



DSA 3023 Big Data Engineering

Semester TBA

Days of the Week, Time of Day, Location TBA

	Instructor	Teaching Assistant
Name	TBA	TBA
Email	TBA	TBA
Office	TBA	TBA
Office Hours	TBA	TBA

Course Description:

Students in this course will develop basic ability to design, build, and implement data pipeline systems to allow efficient access to data and databases. Several topics will be covered including data wrangling, data ingestion, and storage engines. Cloud based systems for data processing and distributed computing will also be discussed.

Student Outcomes:

1. Ability to apply the knowledge of data organization to collect datasets in structured database systems.
2. Capacity to formulate various queries on the database systems to extract and manipulate data.
3. Ability to design, develop, and implement data access methodologies on stored cloud databases.
4. Understand and apply various distributed processing paradigms.
5. Ability to use various cloud architectures and systems for data establishing data pipelines for data collection, aggregation, organization, and, processing.
6. Ability to apply various state-of-the-art big data engineering tools.

Prerequisites:

CS 1213 or CS 1313 or CS 1321 or CS 1323 or CS 1324 and DSA 3013

Required Materials:

No textbook is required for this course.

The Canvas Course Management System will be used as the main mode of communication and information for this class, in that all announcements, notes, assignments and additional class resources can be found there. For fastest response, email should be sent directly to the instructor and not through Canvas email.

Communication Plan:

E-mail communication is the primary method with me in this course. I will respond to students' e-mails within 48 hours or sooner. To ensure a prompt response, please put "DSA 3023" in your e-mail subject line as a prefix. For example, your subject line could be: DSA 3023: Question about Assignment.

Tech Support: If you cannot access canvas.ou.edu, please contact 405-325-HELP

Course Expectations:

For online students: you are expected to view all videos and complete all reading assignments. Also, you will be expected to participate in online discussions and attempt all Self-Check Quizzes.

For on-campus students: Class periods will be used for a combination of lecture, discussion and activity. Attendance is required. It is expected that all students will be prepared and will participate in classroom learning opportunities. there may be unannounced assessments. Students are also expected to engage in out-of-class learning activities including preparatory readings and online activities utilizing Canvas.

Grading:

All assignments must be uploaded to course website by the due date and time indicated for that assignment on the course website. Ensure your familiarity with the system in advance, and allot a sufficient amount of time for assignment submission *and any technical difficulties that may arise*. Assignments should be submitted in advance of an anticipated absence.

GRADING STRUCTURE

Homework & Quizzes	10%
Participation	20%
Projects	55%
Final Exam	15%

Total grade percentages in the following intervals result in the associated letter grade:
[90-100] ← A; [80-90] ← B; [70,80) ← C; [60,70) ← D; [0,60) ← F

Missed Deliverables:

Make-up homework, exams and quizzes will only be permitted for officially documented absences. All documented absences due to religious observances and officially approved trips will be guaranteed a make-up opportunity. Absences due to other unavoidable reasons (e.g., death in the family, illness) will be considered on a case-by-case basis, with appropriate documentation required. Except in the case of an emergency, the student must **always** seek instructor consent **prior to** the absence. If an emergency, you must contact the instructor as soon as feasibly possible. Make-up work will be scheduled at the discretion of the instructor and might occur before the scheduled absence. Typically, make-up work will be scheduled as close to the original assignment date as possible.



Schedule:

Week	Topic
1	Introduction/Course Overview; Introduction to Python for Programmers
2	Introduction to Python for Programmers
3	Structured and Unstructured Data
4	Data Sources and Access Via Python – CSV, JSON, SQL Databases, Text files
5	Structured Data Organization – Schemas, DataFrames, Datasets, Data Filtering, Data Views
6	Structured Data Access and Queries – Introduction to data access with Python/SQL
7	Unstructured Data Organization – Unstructured Database Systems – NoSQL databases – Key-value pair, document, column family store, and graph databases
8	Unstructured Data Access and Queries – Python/MangoDB and Python/Cassandra
9	Introductions to Cloud Platforms-1: Google Cloud Platform (GCP)
10	Introductions to Cloud Platforms-2: Amazon Web Services (AWS)
11	Introductions to Cloud Platforms-3: Azure
12	Data Cleansing, Wrangling, Anonymization, Privacy, and Security
13	Data Pipelines and Related Tools: BigQuery, EL, ELT, and ETL, Data Fusion, Cloud Composer Cloud, Dataproc, Cloud SQL
14	Batch Processing: MapReduce and Distributed File Systems
15	Apache Hadoop, Hive, HBase
16	Stream Processing: Event Streams, Databases and Streams, Processing Streams, Apache Spark, Flume, and Kafka

Professional behavior:

While many of the tools necessary for the completion of this course will be discussed in the course content or via guided assignments, nevertheless ***considerable work outside of class in learning programming tools or researching analytical techniques is expected.***

Academic Honesty:

Cheating, plagiarism, or any act of dishonesty will NOT be tolerated. This policy applies to all parties involved in the incident. Violations will be reported to the office of Academic Integrity.

Never take credit for anyone else's intellectual property, be it on an exam or homework assignment. **This includes, but is not limited to, copying from another student's paper, copying from a paper from a previous semester, using forbidden information on exams, and copying from published writings.** Students are responsible for knowing the requirements of the Academic Misconduct Code, available at <http://integrity.ou.edu/>.

Reasonable Accommodation Policy:

Students requiring academic accommodation should contact the Disability Resource Center for assistance at (405) 325-3852 or TDD: (405) 325-4173. For more information please see the Disability Resource Center website <http://www.ou.edu/drc/home.html>. Any student in this course who has a disability that may prevent him/her from fully demonstrating his/her abilities should contact me personally as soon as possible so we can discuss accommodations necessary to ensure full participation and facilitate your educational opportunities.

Title IX Resources and Reporting Requirement:

For any concerns regarding gender-based discrimination, sexual harassment, sexual assault, dating/domestic violence, or stalking, the University offers a variety of resources. To learn more or to report an incident, please contact the Sexual Misconduct Office at 405/325-2215 (8 to 5, M-F) or smo@ou.edu. Incidents can also be reported confidentially to OU Advocates at 405/615-0013 (phones are answered 24 hours a day, 7 days a week). Also, please be advised that a professor/GA/TA is required to report instances of sexual harassment, sexual assault, or discrimination to the Sexual Misconduct Office. Inquiries regarding non-discrimination policies may be directed to: Bobby J. Mason, University Equal Opportunity Officer and Title IX Coordinator at 405/325-3546 or bjm@ou.edu. For more information, visit <http://www.ou.edu/eoo.html>.

Religious Observance:

It is the policy of the University to excuse the absences of students that result from religious observances and to reschedule examinations and additional required classwork that may fall on religious holidays, without penalty.

Adjustments for Pregnancy/Childbirth Related Issues:

Should you need modifications or adjustments to your course requirements because of documented pregnancy-related or childbirth-related issues, please contact your professor or the Disability Resource Center at 405/325-3852 as soon as possible. Also, see <http://www.ou.edu/eoo/faqs/pregnancy-faqs.html> for answers to commonly asked questions.

Final Exam Preparation Period:

Pre-finals week will be defined as the seven calendar days before the first day of finals. Faculty may cover new course material throughout this week. For specific provisions of the policy please refer to OU's Final Exam Preparation Period policy (<https://apps.hr.ou.edu/FacultyHandbook#4.10>).

Emergency Protocol:

During an emergency, there are official university procedures that will maximize your safety.

Severe Weather: If you receive an OU Alert to seek refuge or hear a tornado siren that signals severe weather 1. LOOK for severe weather refuge location maps located inside most OU buildings near the entrances 2. SEEK refuge inside a building. Do not leave one building to seek shelter in another building that you deem safer. If outside, get into the nearest building. 3. GO to the building's severe weather refuge location. If you do not know where that is, go to the lowest level possible and seek refuge in an innermost room. Avoid outside doors and windows. 4. GET IN, GET DOWN, COVER UP. 5. WAIT for official notice to resume normal activities.

Armed Subject/Campus Intruder:

If you receive an OU Alert to shelter-in-place due to an active shooter or armed intruder situation or you hear what you perceive to be gunshots: 1. GET OUT: If you believe you can get out of the area WITHOUT encountering the armed individual, move quickly towards the nearest building exit, move away from the

building, and call 911. 2. HIDE OUT: If you cannot flee, move to an area that can be locked or barricaded, turn off lights, silence devices, spread out, and formulate a plan of attack if the shooter enters the room. 3. TAKE OUT: As a last resort fight to defend yourself. For more information, visit <http://www.ou.edu/emergencypreparedness.html>

Fire Alarm/General Emergency:

If you receive an OU Alert that there is danger inside or near the building, or the fire alarm inside the building activates: 1. LEAVE the building. Do not use the elevators. 2. KNOW at least two building exits 3. ASSIST those that may need help 4. PROCEED to the emergency assembly area 5 ONCE safely outside, NOTIFY first responders of anyone that may still be inside building due to mobility issues. 6. WAIT for official notice before attempting to re-enter the building.

Mental Health Support Services:

If you are experiencing any mental health issues that are impacting your academic performance, counseling is available at the University Counseling Center (UCC). The Center is located on the second floor of the Goddard Health Center, at 620 Elm Rm. 201, Norman, OK 73019. To schedule an appointment call (405) 325-2911. For more information please visit: <http://www.ou.edu/ucc>

All students are responsible for any changes/additions to this syllabus announced in class.