



DSA - 8660 - 001
Data Management & Warehousing
Spring 2025

Course Description

"Focuses on business advantages and potential of data assets. Emphasis is on case studies and corporate data to explore operational and business intelligence in managerial decision making."

- 2024-2025 Clemson University Graduate Catalog

Instructor Information

He Li

Email: hl3@clemson.edu

Title: Assistant Professor

Office Hours:

2:00pm - 3:00pm, T/Th

Office Location: <https://clemson.zoom.us/j/6789923551>

Course Learning Objectives

Data management is prevalent to evidence-based decision-making by efficiently storing, integrating, processing, and accessing large-scale data in diverse types. This course expands knowledge of data modeling techniques and tools in the contexts of relational

databases, data warehousing, and big data (NoSQL and Hadoop). It intends to balance theoretical concepts, hands-on skills, and business applications of data management for analytics.

Upon completion of the course, students will be able to:

- Describe and explain important data management concepts, including
 - the prevalence of data management and a holistic view of data management framework;
 - Extract, Transform, and Loading (ETL) processes in data warehousing;
 - the components, mechanisms, and applications of big data (NoSQL and Hadoop);
- Design relational databases based on business rules, including
 - building a conceptual data model using Entity-Relationship (ER) and Enhanced ER (EER) models;
 - converting ER and EER models to relations;
 - normalizing relations;
- Query and manipulate data in various types of databases, including
 - Structural Query Language (SQL) in a relational DBMS environment;
 - ETL process using Tableau Prep Builder;
 - data visualization using Tableau Desktop;
 - document-based databases using MongoDB;
- Solve real-world problems and identify business opportunities using effective data management through a final hands-on project;

Course Materials

The instructor will provide required lecture notes and hands-on session instructions. The scopes of exams and course expectations are within the lecture notes assigned by the instructor. While textbooks are NOT required, some of the course materials are based on chapters in the following textbooks:

- Jeff Hoffer, Ramesh Venkataraman, Heikki Topi, Modern Database Management, 13th Edition, Pearson (Print ISBN: 9780134773650, 0134773659; eText ISBN: 9780134792293, 0134792297).
- Carlos Coronel, Steven Morris, Database Systems: Design, Implementation, & Management, 13th Edition, Cengage (ISBN-10: 1337627909; ISBN-13: 9781337627900)

Course Policies

Receiving Grades & Instructor Feedback: Assignment grades and feedback are provided generally 3 days after the assignment is due and always before an assignment of the same type is due. Unless otherwise stated, grades and feedback will be available via the Grades area of the online course site.

Accepting Late Work: Late work will be accepted up to one week from the original deadline but may be subject to a grade penalty. Late work submitted more than one week from the original deadline may receive a failing grade. All work must be submitted by the last day of the course; no extensions or late work will be accepted beyond that date. Please plan ahead.

Communicating with Your Instructor: You can communicate with your instructor via email through your Clemson email account at any time OR Zoom meetings by appointment in advance via email. I am here to help you, so please ask questions and seek clarification as early and as often as needed. Delay will only hinder your learning.

Response Time: Instructor response time is 24 hours for questions sent via email. This response times excludes weekends, official University closures, and other times as noted by the instructor. Should you need live assistance, email me to arrange an online consultation via Zoom.

Course Components

This course is organized around chapters, and usually focuses on the content of one chapter in 1-2 weeks. Each chapter consists of some, or all components listed below.

Component	Description
Course Materials	In each week, you often learn concepts through either PowerPoint presentations (with detailed video scripts) or tutorial videos. You also need to access other course materials made available to you on the online course site.
Zoom Meetings	<p>The instructor will have regular Zoom meetings for the following purposes: (1) answer any questions you may have regarding learning materials; (2) hands-on practices or demo; (3) clarify assignments and learning activities due next week. You are highly encouraged to attend those weekly meetings.</p> <p>All Zoom meetings will be recorded and posted on Canvas afterwards.</p>
Quiz and Exercises (25%)	<p>After studying the topic, you need to complete a short quiz. The quiz usually covers the concepts you learned from the tutorial and other course materials.</p> <p>You will also participate in one online discussion to better understand the database approach.</p> <p>For some practical topics such as designing ERD or relational tables, you will need to complete exercises. The detailed instructions and tasks of these exercises will be posted on the online course site in the corresponding week.</p>
Hands-On Practice	In this course, we will survey several data management tools such as relational database (e.g.,

(20%)	<p>SQL), data processing (e.g., Tableau Prep), data visualization (e.g., Tableau Desktop), and NoSQL databases (e.g., MongoDB). You will learn how to manage data stored in different types of databases based on business needs and data types. For learning purposes, the instructor will try to minimize the installation of software packages on your computer. There will be a detailed Hands-On Practice study guide for each assignment.</p>
Final Project (20%)	<p>There is one final project requiring you to pack all skills you learned from this course (and possibly other DSA courses) to solve a real business problem using effective data management and analytics. You will practice how to solve real-world problems and generate insights using data management tools. Given that coordinating groups can be a challenge, you will have the flexibility to either work individually or work in a group of up to THREE.</p> <p>You have two main options and more flexibility within each option for the project.</p> <p>Option 1: This year, we provide two data analytics topics that you can choose from – healthcare IT or Airbnb listings.</p> <ul style="list-style-type: none"> • Option 1A: The central theme of this year's healthcare IT group project is: How does IT transform the healthcare sector? The instructor has secured a complete dataset of HIMSS 2017,

which is a detailed survey of several thousands of healthcare entities' various IT systems in the North America area in 2017. Each group will identify a relevant business problem in this context. Then, you will need to identify other datasets that can complement the provided HIMSS dataset, merge and process the data, analyze it, and provide findings of your study. You can use any data management tools that you are familiar with, although the tools we learned in this class are recommended.

- **Option 1B:** The Airbnb Listings project uses Airbnb listings data to analyze questions related to Airbnb. The Inside Airbnb website (<http://insideairbnb.com/get-the-data/>) provides detailed data of Airbnb listings of several major cities, including the listings, calendar, and review information. Utilizing the Inside Airbnb data as the base, you can analyze how Airbnb listings compete and survive, factors determining Airbnb growth, the impacts of Airbnb on related industries and societal issues, etc. These are just examples of questions you can explore. You are encouraged to brainstorm interesting and novel questions to study.

Option 2: You can contact a small business such as a coffee shop or a local restaurant. These small businesses may not have a database but would need one as they grow. You can perform a consulting project

	to help them structure their raw data into a well-designed database using techniques we learn in this class. Then, identify a business area that the small business can apply data management or analytics to improve performance and build a demo for them.
Exams (35%)	There will be two mid-term exams to test your learning outcome.

Grading Scale

Total Percentages	Letter Grade
94 – 100	A
90 – 93.99	A-
87 – 89.99	B+
83 – 86.99	B
80 – 82.99	B-
75 – 79.99	C+
70 – 74.99	C
65 – 69.99	C-
0 – 64.99	F

Note: There will be no rounding up of final grades. Discrepancies with grades must be given via email within two days after grades are assigned on Canvas.

Course Schedule

ACTIVITY	POINTS
Week 1: January 8-12 (Due: Sunday, January 12, 11:59 PM)	
Course Syllabus	

Zoom Meeting (January 9, 7:00 PM Eastern Time)	
Week 2: January 13-19 (Due: Sunday, January 19, 11:59 PM)	
Watch Videos: Ch. 1 Introduction to Data Management	
Complete Ch. 1 Quiz	2
Complete Ch. 1 Online Discussion	2
Week 3: January 20-26 (Due: Sunday, January 26, 11:59 PM)	
Watch Videos: Ch. 2 Conceptual Data Modeling	
Watch Video: How to Use Lucidchart	
Complete Ch. 2 Quiz	2
Zoom Meeting (January 23, 7:00 PM Eastern Time)	
Complete Ch. 2 EER Exercise	5
Week 4: January 27-February 2 (Due: Sunday, February 2, 11:59 PM)	
Watch Videos: Ch. 3 Relational Data Model	
Zoom Meeting (January 30, 7:00 PM Eastern Time)	
Complete Ch. 3 Relational Model Exercise	3
Week 5: February 3-9 (Due: Sunday, February 9, 11:59 PM)	
Complete Ch. 3 Quiz	2
Zoom Meeting (February 6, 7:00 PM Eastern Time)	
Complete Ch. 3 Normalization Exercise	5
Week 6: February 10-16 (Due: Sunday, February 16, 11:59 PM)	
Watch Videos: Ch. 4 SQL – Single Tables	
Setup SQL Hands-on Environment	
Practice Exemplar Queries in the Lecture	
Zoom Meeting (February 13, 7:00 PM Eastern Time)	
Complete Hands-on Practice (SQL)	5
Week 7: February 17-23 (Due: Sunday, February 23, 11:59 PM)	
Watch Videos: Ch. 4 SQL – Joining Tables and Subqueries	
Practice Exemplar Queries in the Lecture	
Zoom Meeting (February 20, 7:00 PM Eastern Time)	
Complete Hands-on Practice (SQL)	5
Week 8: February 24-March 2 (Due: Sunday, March 2, 11:59 PM)	
Exam I	20
Week 9: March 3-9 (Due: Sunday, March 9, 11:59 PM)	
Watch Videos: Ch. 5 Data Warehousing and Integration	

Complete Ch. 5 Quiz	2
Watch Tableau Prep Hands-On Tutorial Video	
Complete Hands-on Practice (Tableau Prep)	5
Week 10: March 10-16 (Due: Sunday, March 16, 11:59 PM)	
Watch Videos: Ch. 6 Big Data	
Complete Ch. 7 Quiz	2
Watch MongoDB Hands-On Tutorial Video	
Start Hands-on Practice (MongoDB)	
Week 11: March 17-23	
No Class, Spring Break	
Week 12: March 24-30 (Due: Sunday, March 30, 11:59 PM)	
Complete Hands-on Practice (MongoDB)	5
Zoom Meeting (March 27, 7:00 PM Eastern Time) – Final Project Details	
Weeks 13 & 14: March 31- April 13	
Supplemental Material: Tableau Desktop	
Work on Final Project; Zoom Meeting with Instructor as Needed	
Week 15: April 14-20 (Due: Sunday, April 20, 11:59 PM)	
Mid-Term Exam II	15
Week 16: April 21-27 (Due: Sunday, April 27, 11:59 PM)	
Submit Record Final Project Presentation on Canvas Discussion Board Participate in Discussion by Commenting/Liking Others' Posted Work	20

University Policies

You can find further information about campus wide policies, including student affairs information and accessibility services, by clicking on [Clemson's University Policies Page](#). There, you can find information about academic integrity, access and equity (including student accessibility and Title IX info), student financial services, emergency planning, and more. My expectation is that you will review these policies carefully and be responsible for them this semester.

Student Accessibility Services

Clemson University values the diversity of our student body as a strength and a critical component of our dynamic community. Students with disabilities or temporary injuries/conditions may require accommodations due to barriers in the structure of facilities, course design, technology used for curricular purposes, or other campus resources. Students who experience a barrier to full access to this class should let the instructor know and make an appointment to meet with a staff member in Student Accessibility Services as soon as possible. You can make an appointment by calling 864-656-6848, by emailing studentaccess@lists.clemson.edu, or by visiting Suite 239 in the Academic Success Center building. Appointments are strongly encouraged – drop-ins will be seen, if at all possible, but there could be a significant wait due to scheduled appointments. Students who have accommodations are strongly encouraged to [request, obtain, and send these](#) to their instructors via SAS as early in the semester as possible so that accommodations can be made in a timely manner. It is the student's responsibility to follow this process each semester.

You can access further information at the [Student Accessibility website](#). Additional accessibility information can be found at the university's [Accessibility Portal](#).

Title IX

The Clemson University Title IX statement: Clemson University is committed to a policy of equal opportunity for all persons and does not discriminate on the basis of race, color, religion, sex, sexual orientation, gender, pregnancy, national origin, age, disability, veteran's status, genetic information or protected activity in employment, educational programs and activities, admissions and financial aid. This includes a prohibition against sexual harassment and sexual violence as mandated by Title IX of the Education Amendments of 1972. This [Title IX policy](#) is available online. Ms. Alesia

Smith is the Clemson University Title IX Coordinator, and the Executive Director for Equity Compliance. She can be reached at (864) 656-3181 or via email at alesias@clemson.edu. Remember, email is not a fully secured method of communication and should not be used to discuss Title IX issues.

Academic Integrity

As members of the Clemson University community, we have inherited Thomas Green Clemson's vision of this institution as a "high seminary of learning." Fundamental to this vision is a mutual commitment to truthfulness, honor, and responsibility, without which we cannot earn the trust and respect of others. Furthermore, we recognize that academic dishonesty detracts from the value of a Clemson degree. Therefore, we shall not tolerate lying, cheating, or stealing in any form. Using materials generated using artificial intelligence (AI) that are turned in without attribution is considered plagiarism.

All infractions of academic dishonesty by undergraduates must be reported to Undergraduate Studies for resolution through that office. In cases of plagiarism instructors may use the Plagiarism Resolution Form. See the following resources:

- [Undergraduate Academic Integrity Policy](#)
- [Graduate Student Handbook](#)

Online Testing

By enrolling in a course with online assessments, the student agrees and consents to the use of an online test proctoring service or software program as described above. The student also agrees to allowing the student, their activity, and surrounding workspace to be recorded by video and audio and then analyzed by the test proctoring system, the course instructor, and

others at Clemson University. For more information see the Academic Regulations section of the Undergraduate Catalog.

Tech Support

If you have trouble with Canvas or another university system, check here first: [Clemson System Status](#). There, you can see if there is a current issue and when it might be resolved. If there's no current issue listed, try logging out and quitting your browser before trying again. You can also look in the [Canvas Help Guides](#) for more information.

The CCIT Support Center at Clemson offers a wide range of support options and hardware repair with several contact methods to help you answer your questions as quickly as possible:

- Phone: (864) 656-3494
- Email: ITHelp@clemson.edu
- Chat: [Live Online Chat](#)
- Web Form: [Help Request Form](#)
- Troubleshoot: [Knowledge Base](#)
- Everything CCIT Does: [Browse Services](#)

Officially, Canvas supports all the major web browsers: Chrome, Firefox, Edge, and Safari. However, Safari is not fully compatible; images often do not show up in Safari, so Clemson Online recommends against using it with Canvas.

Canvas also has mobile apps available for teachers, students, and parents, all of which can be found in your mobile device's app store. These apps are convenient, but do not necessarily display or connect to every component of a Canvas site. They are good for assignment reminders, announcements,

and messaging, but not as good for reading materials, taking exams, or building elements.

Emergency Preparedness Statement

Emergency procedures have been posted in all buildings and on all elevators. Students should be reminded to review these procedures for their own safety. All students and employees should be familiar with guidelines from the Clemson Police Department. Visit the [Clemson University Emergency Management site](#) for information about safety.

Academic Success Center

The [Academic Success Center](#) (ASC) offers a variety of free learning and success services for all undergraduate students that include:

Mastery of course content:

- Tutoring – students can expect a 1:1 meeting with a trained undergraduate peer leader (who made an A or B in the course and was recommended by a faculty member) during which the student can share specific questions they have about course content with the tutor focused on helping the student, through questioning techniques and identification of helpful learning strategies, and master course concepts. Tutors do not help with homework or other class assignments.
- Peer-Assisted Learning (PAL) – students can expect collaborative and active group learning and study sessions focused on mastery of course content and learning strategies that is facilitated by a trained undergraduate peer leader (who made an A or B in the course and was recommended by a faculty member). PAL leaders do not help with homework or other class assignments.

Learning and Success Strategies:

- Academic coaching – students can expect a 1:1 meeting with a trained professional academic coach during which the coach helps students see themselves, their skills, and their study habits from a fresh perspective through one-on-one sessions focused on learning and personal success strategies.
- Success strategy workshops – students can expect 30-45 minute workshops on college success skills, time management and organizational skills, test-taking strategies, study strategies, finals preparation, life skills, and academic resources.
- College success skills course (CU 1010) – students experiencing academic difficulty can expect a course focused on academic and personal skill building taught by instructors who wish to work with this student population.

ASC services are designed to equip students with strategies and resources they can use to:

- Succeed in their courses
- Become more confident, independent, and skillful learners
- Engage in more productive and effective study and learning strategies
- Manage their time more effectively

Location: The Class of 1956 Academic Success Center building is located in the center of campus adjacent to Cooper Library and the Watt Family Innovation Center.

Clemson Writing Lab

[The Writing Lab](#) can help you with any kind of writing assignment. You can schedule a face to face, Zoom, or email appointment with a trained peer

consultant. Contact the Writing Lab at: writinglab@clermson.edu

Clemson Libraries

If you need help researching or finding resources, the library is here to help. You can contact a Librarian through the [library website](#), or ask in person at Cooper Library.

Online Course Etiquette ("Netiquette")

Even if you have experience communicating in an online environment, online courses may require an extra effort to be civil. An online classroom is still a classroom and we should all behave professionally while in it. Please read and follow the guidelines below.

While in our class, please remember:

- Be respectful. Consider the person on the receiving end of your comments. If you were sitting in front of that person, would you say the same thing to their face?
- Humor and sarcasm don't always translate well in online discussions. Always consider your tone and whether your words can be clearly understood by your classmates.
- Only reveal information that you are comfortable sharing with your classmates.
- Please spell-check your responses; you can even compose your work in a different program (such as Word), and then paste it into the online discussion.
- Posts don't have to be perfect, but they should be as clear and readable as you can make them.

- The law still applies on the Internet. Do not commit illegal acts online, such as libeling or slandering others, and do not joke about committing illegal acts.

Please do not:

- Use offensive language, even in jest. Any comments that can be construed as racist, sexist, etc., may be removed.
- Avoid using caps lock to post messages or responses to posts. All-caps is considered SHOUTING AND IS VERY RUDE.
- Post large, unbroken blocks of text. These can be difficult to read. Try to break up your post into paragraphs.
- Post a message more than once; if you are having trouble seeing or submitting posts, contact me and we will try to fix it.
- Re-post or forward any communication that is sent to you unless you receive explicit permission to do so. Respect the response you get if your request is denied.