



**SMU**

*Master of Science in Data Science*

# **DS 7330: File Organization and Database Management**

## **Course Syllabus**

### **Welcome to File Organization and Database Management**

Welcome to DS 7330, File Organization and Database Management. The objective of this document is to acquaint you with the administration, procedures, and policies of this course. Save this document for future reference.

### **Course Overview**

DS 7330 is designed to introduce students to the foundations of database systems, with an emphasis on the fundamental topics including data modeling, SQL, and data warehousing. This course will also introduce students to the cloud and how the cloud is affecting the lives of data scientists.

### **Course Designer**

This course was designed by Dr. Sohail Rafiqi.

### **Course Student Learning Outcomes**

This course provides an introduction to database systems. The primary goal of this course is to teach students the foundational concepts and tools of database systems. As such, the primary learning outcomes for this course are

1. An ability to design and document a data model and relational database
2. An understanding of the fundamental issues involved in designing and using databases
3. Develop a database and perform advanced SQL queries using Python
4. Understand fundamental concepts around the cloud
5. Develop a database system within the cloud environment

The general learning outcomes for the department degree programs are supported within this course. Through the various activities associated with this course, we will, to a greater or lesser degree, achieve the following departmental learning outcomes:

1. An ability to apply knowledge of mathematics, science, and engineering to software and hardware design problems
2. An ability to design and conduct experiments and to analyze and interpret data related to software and hardware design solutions
3. An ability to design a system, component, or process to meet desired needs within realistic constraints such as economic, environmental, social, political, ethical, health and safety, manufacturability, and sustainability
4. An ability to function on multidisciplinary teams using current computer engineering tools and technologies
5. An ability to identify, formulate, and solve engineering problems based on a fundamental understanding of concepts of computer engineering topics
6. An understanding of personal, professional, and ethical responsibility
7. An ability to communicate effectively both in an oral and written form
8. The broad liberal arts education necessary to understand the impact of engineering solutions in a global, economic, environmental, and societal context
9. Recognition of the need for, and an ability to engage, in lifelong learning
10. Knowledge of contemporary issues in computer engineering

11. An ability to use the techniques, skills, and modern engineering tools necessary for computer engineering practice

## Course Instruction Using Synchronous and Asynchronous Sessions

The course uses a combination of synchronous class sessions and asynchronous material and activities to teach students the course material and guide them through the learning process. Synchronous class sessions occur once per week during the course of the term. These sessions consist of lectures, discussions, problem-solving, in-class assignments, and quizzes based on the asynchronous material, including the course video lectures, assigned activities and work, and any readings assigned. It is expected that all asynchronous material will be completed—for example, videos viewed, assigned readings read, and assigned work completed and turned in—prior to the synchronous session associated with that material.

## Course Prerequisite

A student taking DS 7330 must be enrolled in the Master of Science in Data Science program at SMU.

## Course Textbook and Other Course Material

The required textbook for this course is

**Silberschatz, Abraham, Henry F. Korth, and S. Sudarshan. *IT Database System Concepts*, 6th edition. McGraw-Hill, 2011.**

**Addition Material:** Section instructors will provide additional reading materials. These will be posted to the course wall as needed.

## Course Topic Summary

Table 1 lists the basic topics covered during each week of this course.

Table 1: Topic Summary for Each Week of the Course

Week/Unit	Topic	Readings
1	Understanding Relational Database Management Systems	Chapter 1
2	RDBMS Data Models	Chapter 2
3	Designing Databases	Chapter 7
4	Introduction to SQL	Chapter 3
5	Intermediate SQL	Chapters 3–4
6	Accessing SQL	Chapter 8
7	NoSQL and OLAP	Additional Material
8	Midterm (no asynchronous content)	
9	Big Data/Data Lake	Additional Material
10	Changing Landscape with Cloud	Additional Material
11	Data Management in the Cloud	Additional Material
12	Big Data Project (no asynchronous content)	Additional Material
13	Data Warehouse and Machine Learning in the Cloud	
14	Team Project Presentations (no asynchronous content)	
15	Final Exam/Team Project Presentations (no asynchronous content)	

## Technology Requirements

## *Course Syllabus: DS 7330 File Organization and Database Management*

DS 7330 is a course taught online with both synchronous and asynchronous portions requiring the transfer of video. Students are expected to have access to a computer with reliable, high-speed internet access. Students are expected to have access to a computer with a web camera with the computer capable of running the required software to access the learning management system, to read online documents, to watch course videos, and to participate in the synchronous classes (including being on camera). Students are also expected to have access to a reliable phone connection to participate in the synchronous classes.

The DS 7330 course utilizes MySQL, MySQL Workbench, Python, and a public cloud (the section instructor will choose the public cloud) to teach the course material.

All students enrolled in SMU have an SMU email account. Notifications from the digital learning platform and from the course instructor use your SMU email account. Students are encouraged to check this email regularly.

### **Course Access**

#### **Asynchronous Videos:**

This course is accessible to registered students in the SMU MSDS program only. Course asynchronous videos are available through the 2DS digital learning platform. Access to the 2DS digital learning platform is available at <https://2ds.datascience.smu.edu/>.

Students that experience technical issues with the learning management or Zoom should contact technical support as described below.

#### **Course Documents:**

All of the course documents, including assignments, submission to assignments are available through <https://2ds.datascience.smu.edu/>.

Students will have access to only those courses and course sections in which they are currently enrolled or have been enrolled in previous terms. Access to other sections is at the discretion of the section instructor. Access to recordings of synchronous sections where the student did not participate or was not an enrolled student are prohibited to protect the privacy of the students that do attend and participate.

### **Communication and Technical Support**

Direct communications with the instructor should be made in the manner indicated by the instructor. General questions and questions that are relevant to multiple students, that is, questions that are not specific to an individual and involve that individual's private information, should be posted on the course wall.

Technical support for the learning management system and the online classroom may be reached 24 hours a day, seven days a week via

- **Chat Support:** Click **Live Support** on the lower-right-hand corner of the 2DS screen after logging in to the system to chat with a technical support representative. Chat support generally responds and engages in five minutes or less.
- **Phone:** Students should call 1-844-768-5637 (toll-free) to speak with a technical support representative.
- **Email:** [studentsupport@datascience.smu.edu](mailto:studentsupport@datascience.smu.edu) to initiate a support request with a technical support representative.

For other questions or concerns, please contact the appropriate SMU department for your questions or concerns or send email to [datascience@smu.edu](mailto:datascience@smu.edu).

It is the student's responsibility to ensure that all communications are received or acted upon.

## Course Procedures and Policies

This course has a number of policies and procedures that students should understand and follow if appropriate. The following sections present the general course policies and procedures that students must follow. Additional policies and procedures may be given by the instructor. Please discuss as early in the term as possible with the instructor any questions or concerns that you may have regarding the course procedures and policies as defined herein or any additions made by the instructor to the course procedures and policies.

## Course Grading Policy

This course consists of a number of assignments and projects that are to be completed throughout the term. It is expected that all students will put forth the effort required to earn an A letter grade for this course. Assignment grades will be determined using evaluation rubrics. You are responsible for reviewing the rubrics and raising questions or concerns related to the assignments, their rubrics, and their grading prior to the submission of each assignment. Questions regarding the grading of any assignments should be directed to the course instructor as soon as possible and in accordance with any regrading policy instituted by the instructor.

The final grade for the course will be calculated on the bases of the earned cumulative percentage and the grade received for each of the components of the cumulative percentage. This course is not graded on a curve. The required cumulative percentage needed to earn each letter grade is given in table 2.

Table 2: Cumulative Percentage Required to Reach Each Letter Grade

Cumulative Percentage	Earned Grade
100–92	A
92–90	A–
90–88	B+
88–82	B
82–80	B–
80–78	C+
78–72	C
72–70	C–
70–60	D
< 60	F

The cumulative percentage for the course is determined by the course assignment components with their corresponding percentages defined in table 3.

Table 3: Grade Components and Weightings of the Cumulative Percentage

Percentage of Cumulative Percentage	Component
15%	Asynchronous/Synchronous Participation
20%	Mini Projects (Labs) – 6 Projects
20%	Midterm
25%	Final Project Presentation and Documentation (includes First Draft, Project Proposal, Final Paper, Final Presentation)
20%	Final Exam

You will receive a grade for each component. If you earn less than 60 percent (grade of D) in any one of these components, you will receive a final grade of F for this course.

A course grade of *Incomplete* (I) will be given only in the case of extraordinary circumstances that prevent the student from finishing the semester. Students must have completed at least 50 percent of the course with a passing grade to be eligible for an *Incomplete* grade.

## **Graded Assignment Details**

### **Asynchronous/Synchronous Participation**

To receive full credit for asynchronous/synchronous participation, you must complete all asynchronous questions, post to the weekly discussion boards, and participate in all live sessions, while on-camera. Students will receive one grade for their overall participation in the course.

### **Mini Projects (Labs)**

The Mini Projects (labs) grade will consist of five mini projects as well as the Big Data project. Each project will be weighted equally. Collaboration is expected and encouraged; however, each student must hand in their own homework assignment. You will have one to two weeks to complete each project, depending upon the assignment. Due dates for mini projects will be shared on the course wall by your instructor. Instructor will share password for access to the mini projects. All Mini Projects will be weighted equally.

### **Midterm**

Your midterm will cover weeks 1–7. The midterm will consist of multiple choice and essay questions. You will receive the password to access the midterm at the start of the week 7 live session. It is due by the start of the week 8 live session.

### **Final Project Presentation and Documentation**

This is due by the start of week 14 live session. Details regarding the final project presentation are in the Toolbox. The final project is divided into four parts.

1. Project proposal (by the start of week 5 live session)
2. First draft (by the start of the week 10 live session)
3. Final paper (by the start of the week 14 live session)
4. Presentation (during the week 14 live session)

### **Final Exam**

The final exam will be comprehensive. You will work with AWS to develop an end to end configuration on the cloud. You will receive the password to access the final exam at the start of the week 14 live session. It is due by the start of the week 15 live session.

## **Grade Grievance Policy**

Students are responsible for saving all graded materials as evidence in case of a discrepancy with the assigned grades. Students are responsible for ensuring that all grades are correctly reflected on the grade store. Any identified discrepancies should be brought to the attention of the instructor as soon as the discrepancy is found.

Refer to the university catalogue for the university policy and process for grade grievances.

## **Assignment and Collaboration Policy**

Data science is an inherently collaborative subject, and learning often occurs best when subjects are taught both to and from peers. Collaboration is expected to occur both in learning the course material and in performing the course work. However, each student must hand in their own work performed by themselves unless explicitly allowed by written directions given by the instructor. Collaboration means helping one another learn the material. Collaboration does not mean copying answers from one another.

Assignment submissions that contain substantially the same answers shall receive a grade of zero on the first instance and a course grade of F upon a second instance. To mitigate potential issues and questions of similarity, peers with whom a student collaborates should be clearly identified by that student in their submissions.

## **Scholarly Expectations**

Work submitted at the graduate level is expected to demonstrate critical and creative thinking skills and be of significantly higher quality than work produced at the undergraduate level. To achieve this expectation, all students

are responsible for giving and receiving peer feedback of their work. Students are also expected to resolve technical issues, be active problem-solvers, and embrace challenges as positive learning opportunities. Data science professionals must be able to teach themselves and teach others to fill in any gaps in their knowledge or to find a way of learning new material that is most conducive to their learning style. Data science professionals must also be able to work cooperatively and collaboratively with others—skills that students are expected to practice in this course. Students are expected to ask questions and ask for help when they need it and to offer help when others are in need.

Absent questions or requests for assistance, instructors must assume that students understand the material being covered and are able to complete the assignments. It is primarily through your questions that the instructor learns where the students are struggling to understand and on which topics more time needs to be spent for the students' benefit.

### **Timeliness**

Because a 15-week term goes by quickly, assignments must be submitted by the designated due dates. Full credit cannot be earned by late or incomplete assignments. Assignments may lose up to 10 percent of their possible value each day late if submitted after the posted due date/time (e.g., assignments can lose all of their value at 10 days past due). When a project incorporates peer review, it is imperative that all projects be available at the beginning of the review period and that reviews are completed by the end of the review period so that others may incorporate feedback into project revisions. You will have plenty of notification and time to complete course assignments. If you know you are going to be out of town, involved in a special event/project, or unable to access a computer, please plan ahead. Also ensure that you have a backup plan ready in the event you might lose power, internet access, or your available technology.

### **Time Commitment**

As a technical graduate-level course, it is expected that students will spend between three and four hours beyond course instruction for each hour spent in instruction. MSDS courses are designed to have approximately three hours of course instruction, or contact hours, per week of the course. Therefore, it is expected that students will spend between 12 and 15 hours per week on this course.

### **Attendance Policy**

Attendance and on-camera participation at the weekly synchronous sessions in this course is mandatory. Students with more than three unexcused absences will receive a final grade of F for this course. It is the student's responsibility to notify the instructor if a synchronous session will be missed for either an excused or unexcused reason at least 24 hours, or as soon as reasonably possible, prior to the synchronous session.

### **Drop Policy**

Refer to the university drop policy for a complete description of the drop and withdrawal policies for this course.

### **Campus Concealed Carry**

Concealed handguns are prohibited on the Southern Methodist University campus. Pursuant to section 30.06, Penal Code (Trespass by License Holder with a Concealed Handgun), a person licensed under subchapter H, chapter 411, Government Code (Handgun License Law), may not enter SMU property with a concealed handgun. Report violations to the Southern Methodist University Police Department by dialing 9-1-1 or 214-768-3388 (nonemergency) or 214-768-3333 (emergency).

### **Americans With Disabilities Act**

Disability Accommodations: Students needing academic accommodations for a disability must first be registered with Disability Accommodations & Success Strategies (DASS) to verify the disability and to establish eligibility for accommodations. Students may call 214- 768-1470 or visit <http://www.smu.edu/alec/dass> to begin the process. Once registered, students should then schedule an appointment with the professor to make appropriate arrangements. (See University Policy No. 2.4.)

## **Religious Observance**

Religiously observant students wishing to be absent on holidays that require missing class should notify their professors in writing at the beginning of the semester, and should discuss with him, in advance, acceptable ways of making up any work missed because of the absence. (See University Policy No. 1.9.) Failure to notify your professor prior to your absence will result in an unexcused absence and possibly a grade of zero for any assignments.

## **Excused Absences for University Extracurricular Activities**

Students participating in an officially sanctioned, scheduled University extracurricular activity should be given the opportunity to make up class assignments or other graded assignments missed as a result of their participation. It is the responsibility of the student to make arrangements with the instructor prior to any missed scheduled examination or other missed assignment for making up the work.

## **Academic Integrity**

It is the philosophy of Southern Methodist University that academic dishonesty is a completely unacceptable mode of conduct and will not be tolerated in any form. All persons involved in academic dishonesty will be disciplined in accordance with University regulations and procedures. Discipline may include suspension or expulsion from the University.

Scholastic dishonesty includes but is not limited to cheating, plagiarism, collusion, the submission for credit of any work or materials that are attributable in whole or in part to another person, taking an examination for another person, or any act designed to give unfair advantage to a student or the attempt to commit such acts.

Students caught being academically dishonest shall receive a grade of F for this course.

## **University Honor Code**

When you signed your letter of intent to enroll in the MSDS program, you initialed the following statement:

“I have read and agree to abide by the SMU Honor Code available online at:  
<https://www.smu.edu/StudentAffairs/StudentLife/StudentHandbook/HonorCode>”

The Honor Code is taken seriously at all levels within the university. Students that are found to have violated the honor code will be disciplined which often includes expulsion from the university.

## **Plagiarism**

Plagiarism is the “practice of taking someone else’s work or ideas and passing them off as one’s own” (this definition is from Google Dictionary). An example of plagiarism is as follows:

A regression is a statistical analysis assessing the association between two variables. It is used to find the relationship between two variables.

The following is NOT plagiarism:

“A regression is a statistical analysis assessing the association between two variables. It is used to find the relationship between two variables” (<https://www.easycalculation.com/statistics/learn-regression.php>).

The difference is in the punctuation and the attribution. Note that one can self-plagiarize. If you are using something that you wrote (e.g., a blog or a previously published article), please reference yourself.

DO NOT PLAGIARIZE. If you have any question as to what is and what is not plagiarism, ask your instructor. As a general rule, always use your own words and cite your source.



The consequence for being caught plagiarizing is to earn at least a zero on the identified assignment and may include earning a course grade of F and a referral to the SMU Honor Council for your Honor Code violation.

## **Best Practices for Success in the Course**

*Attendance.* Take responsibility for your commitment. Attendance means not only being there for synchronous sessions but also participating in asynchronous work.

*Citizenship.* You need to be actively engaged to succeed in this class. Talking on cell phones, texting, “Facebooking,” tweeting, and leisure web browsing are prohibited in class. I consider these to be a disruption (not to mention rude).

*Integrity.* A lot of the graded work occurs outside of class, so I expect honesty and integrity in what you submit for evaluation. Evidence of academic dishonesty will minimally result in zeros for all involved parties and perhaps University-level disciplinary action. Don’t risk your career.

*Humility.* Don’t get lost! Ask questions in class. If something isn’t clear to you, it probably isn’t clear to others either. Questions may arise because I haven’t made a connection clear or have inadvertently left out an important point. Your question gives me a chance to explain more clearly. Don’t be proud or shy.

*Organization.* Don’t procrastinate! This is a technology-driven course. Count on your computer failing or your wireless connection breaking the night before a due date. Start early and give yourself a chance to succeed.

*Deadlines.* You will generally have a week to complete an assignment. Due dates and times will be clearly indicated. Late submissions will be penalized, but it is much better to turn in work late than not at all (or to turn in incomplete/sloppy work). Work turned in after solutions have been posted to the course website will receive no credit.

*Getting help.* If questions arise while doing assignments/exams, do your best to resolve these questions before the assignment is due, first by taking time to seek answers yourself, next by asking questions on the wall, and finally via email to your instructor or other students. I encourage you and expect you to seek help. For questions during exams, please email the live session instructor directly.

*Collaboration.* I encourage the formation of study groups and collaboration with your fellow students in tackling the assignments. Working together in groups on homework is permitted, even encouraged. However, every student should write up and complete his or her homework independently. Talking about problems with other people does help in learning, but just copying the solutions from one another doesn’t help!

*Looks do matter!* All assignments must be NEATLY executed and organized. You risk a zero on any assignment submitted in a sloppy manner. See submission guidelines for more detail.

*Have Fun!* Learning is meant to be a fun activity. Although it can be difficult, time consuming, frustrating, and sometimes disappointing, always seek to find the fun in what you are doing and learning. The gratification from learning complex concepts and applying them to solve hard problems is what we are all striving to achieve. Having fun while we are learning, and teaching others just makes the learning easier and friendships better.