School of Business

MIS 430: Data Warehousing

Instructor
Timothy Suh
Contact: e-mail: tsuh2@gmu.edu
Office Hours: By appointment
Class Site: https://mymasonportal.gmu.edu/

Course Materials
Required Textbook:

You are required to take class notes. Lecture slides, links to resources, video, and other reference materials will be made available using Blackboard.

Software:
Students MUST have Microsoft SQL Server 2017 installed. The recommended version to download from Microsoft Imagine is “SQL Server 2017 Enterprise”. This is widely used commercially to support an organization’s data management requirements in form of database, data warehouse, and business intelligence. Detailed instructions will be provided on how to install on your own laptop. You are required to bring your laptop to class for in-class hands-on instruction, in-class exercises, and quizzes.
THIS SOFTWARE ONLY RUNS ON WINDOWS.

In prior years, students have had significant difficult getting this software installed. Please start the process of installing this software on your laptops during the first week of class.

Minimum requirement for laptop is:
- 4GB of RAM
- 20 GB of free disk space.

Additional data analytics software may be used in the course. Instructor will provide further instructions to install these software.

Course Prerequisite
Required Prerequisites: MIS 310 or L310.
\(^c\) Requires minimum grade of C.

Course Description
Data management is critical to efficient and effective operation of all modern organizations. Businesses collect large data as part of the daily operations, it is vital to ensure that databases be suitably designed to ensure data quality as well as speed of operations required for storage as well as retrieval of this information. In the first part of this course, students will learn implementation designs and database administration policies that are used for optimizing databases for business operations. As part of this we will cover some advanced SQL for implementation of database elements as well SQL for retrieval of data using complex queries.

Traditionally, organizations undertook data collection and storage with the objective of assisting day to day operations, such as order processing, inventory management, payroll, etc. This was achieved through use of databases that supported a variety of business applications. However, these databases could not provide insight required for strategic decision making. For e.g., helping managers in deciding where to open the next store? Which customer segment to target for which types of promotion? Furthermore, unlike day-to-day operations, strategic decision making requires access to large amounts of historical information. Additionally, most of these queries are ad-hoc queries and involve large number of attributes. These criterions make transactional databases ineffectual in providing dedicated support for such decision making. Data warehousing focuses on providing just this support. Data warehousing refers not just to the design and storage of historical information but identifies the entire infrastructure involved in enabling the necessary decision making. In the second part of the course, students will learn the design principles for a data warehouse, and utilize these to create suitable data warehouse designs that meet the business requirements. Students will also learn how to create business reporting, ETL, analytic solutions, as well as other components of a modern data warehouse.

**Learning Objectives**

This course aims to provide a comprehensive understanding of database management and data warehousing, with specific focus on:

- Physical implementation of database optimized to serve business requirements.
- Apply the understanding to implement database objects such as views and indexes.
- Apply the SQL for both database implementations as well as information retrieval.
- Modeling and implementation of database using Data Modeling software.
  - Understanding the difference in the role of transaction databases and data warehouses, their objectives and how it influences the process of design and implementation of each
- Architecture of data warehouses, and associated decision support systems
- Design approaches for data warehousing, with specific focus on relational data warehouse design
- Implementation and operationalization of data warehouses
- Evaluating and tuning performance of data warehouses
- Use of data warehouses for reporting and OLAP
- Creating business intelligence reports
- Develop extract, transform, and load processes
- Create solutions that the management of structured data

**Grading**

Students must be officially registered in this course to receive a grade. It is the sole responsibility of the student to verify their own registration status. Specifically, you will not receive a grade if your name does not appear on the official class list. (Don’t wait until the end of the semester to be surprised.) Registration problems should be directed to either the SOM Office of Student Services or the Registrar’s Office. Grading for the course will be based on total
points earned by the end of the course. Final course letter grade assignments will be as follows:

<table>
<thead>
<tr>
<th>Grade</th>
<th>Percentage Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>93% - 100%</td>
</tr>
<tr>
<td>B+</td>
<td>87% - 89.99%</td>
</tr>
<tr>
<td>C+</td>
<td>76% - 79.99%</td>
</tr>
<tr>
<td>D</td>
<td>64% - 69.99%</td>
</tr>
<tr>
<td>A‐</td>
<td>90% - 93%</td>
</tr>
<tr>
<td>B‐</td>
<td>80% - 82.99%</td>
</tr>
<tr>
<td>B</td>
<td>83% - 87.99%</td>
</tr>
<tr>
<td>C</td>
<td>70% - 75.99%</td>
</tr>
<tr>
<td>F</td>
<td>below 64%</td>
</tr>
</tbody>
</table>

Mid Term Exam = 25%
Final Exam = 35%
Quizzes = 20%
Assignments = 20%
Total = 100%

Quizzes/Exercises/Exams: There will in-class quizzes and two exams - final and midterm - focused on the key principles of data warehousing covered in class. They will consist of true/false, multiple choice, short answer or similar type questions. Database design and other exercises may also be included.

Assignments: There are assignments meant to reinforce key concepts of database and data warehousing. The goal of these assignments is to provide early hands-on experience with subjects later used in the project. The assignments are graded for content and effort. Late submissions are subject to penalties up to 50%. As part of the assignments, there will be a series of hands-on lab to to give students a practical experience developing your very own data warehouse about a particular subject area. You will be given a dataset to work with in order to design the model, load data into the warehouse, and visualize the results.

A Note on Collaboration and Plagiarism: The presentations and project for this class will be group assignments so, obviously, you may collaborate on those within your group. However, please be cautious of plagiarism. While I expect you to do outside research, please be sure to incorporate your findings into your work in a proper manner. Quote and reference other’s work correctly when using it in your own. Please refer to the Honor Code. Results of instances of plagiarism or cheating range from receiving no credit for an assignment/group project to failure in the course.

Final cumulated scores of the class will be analyzed to determine if an adjustment or curve is necessary. However, a curve will only help you. Therefore, if you have a 90% or more you will get an A in the class - I won’t curve to raise the minimum requirements for grades, only lower them if necessary. Depending on the distribution of the grades I may or may not use the +/- system.

Attendance
Class Attendance is your responsibility, and you are solely responsible for all assignments, material presented/provided and announcements made in class. You must get any information missed from a class from another student. Assignments due to absence cannot be made up or turned in outside of class.

Classroom Etiquette
It is expected that you are courteous and professional to both your instructor and classmates. This includes, but is not limited to, turning off your cell phone during class. There is a zero tolerance policy for texting or any other cell phone use in class. Cell phones may be left on vibrate for emergency notification purposes. If you expect an important phone call, please inform me before class and quietly excuse yourself when you receive it. For every instance of texting or other cell phone use that the instructor observes during lecture, 3% will be deducted from the final grade. I reserve the right to ask you to leave the classroom if I consider your behavior disruptive to the delivery of the lecture or exam.
During class you are allowed to use a laptop or a tablet only for note-taking purposes or class materials related work. If during the class you are seen using your laptop or tablet for other than note-taking purposes you will be asked to turn off your device. Photos or video may not be taken in class without prior permission. (Personal use only allowed in an emergency or during a break).

One of the most important things is to start the class on time. Students should make every effort to be at class on time and, if late, find a seat quickly and disturb the class as little as possible. In order to arrive on time, please make sure that you allow 30 minutes between this class and classes on main campus.

I will try to learn your names as fast as I can. To help me, please put a name tent on your desk in every class. In order to make this process quicker, I strongly encourage using it in every class and participating in the class discussion. As mentioned earlier, the latter will also help you if you are borderline between two grades at the end of the semester.

**Make up Exams/Quizzes/Assignments**
Barring extenuating circumstances no exceptions will be made for absence. *Final exam due date/time will not be rescheduled.* Adequate proof needs to be provided to prove extenuating circumstances. Work related time conflict does not constitute extenuating circumstance. If you have any work-related time conflict, you need to resolve it in advance.

**Disability**
All academic accommodations due to disability must be arranged through the Disability Resource Center (DRC). If you are a student with a disability and you require academic accommodations, please contact the DRC at 993-2474. I will cooperate fully with the DRC to accommodate a student’s special needs.

**Honor Code**
GMU students are expected to be familiar with the Honor Code of George Mason University and with its specific application to exams, assignments, and class work required by faculty in the program. If you are in doubt about how the honor system applies to a particular assignment or class, it is your responsibility to clarify the requirements with the professor. Concerns about breaches of the honor system may be discussed with the professor, or with the Associate Dean of the School of Management. More detail on honor code provided on class website.

**Inclement Weather & Campus Emergencies**
Information regarding weather-related changes in the University's schedule (e.g., closing or late opening) will be provided on the GMU website and via Mason Alert. Students sign up for the Mason Alert system to provide emergency information of various sorts at https://alert.gmu.edu. If campus is closed, please check Blackboard for announcements from the professor.

**Communications**
All communications from me to you will be directed via e-mail. Students must use their GMU email account to receive important University information, including communications related to this class. *I will not respond to messages sent from or send messages to a non-Mason email address.* I will only address and reply to all of e-mails from/to your @gmu.edu e-mail address for concerns of privacy and confidentiality. If you use another e-mail account as your primary e-mail, please be sure to forward your GMU e-mail to that account.

Your communications with me – e-mail is the preferred channel of communication, since it ensures your accessibility to me irrespective of where I may be. *You should always feel free to send me an e-mail*, no matter if it is a question, comment, concern, something interesting you came across (related to class or otherwise), etc. Basically – when in doubt, decide in favor of clicking the “send” button.
Notice of mandatory reporting of sexual assault, interpersonal violence, and stalking

As a faculty member, I am designated as a “Responsible Employee,” and must report all disclosures of sexual assault, interpersonal violence, and stalking to Mason’s Title IX Coordinator per University Policy 1412. You may seek assistance from Mason’s Title IX Coordinator, Jennifer Hammat, by calling 703-993-8730 or email cde@gmu.edu. If you wish to speak with someone confidentially, please contact one of Mason’s confidential resources, such as Student Support and Advocacy Center (SSAC) at 703-993-3686 or Counseling and Psychology Services (CAPS) at 703-993-2380. The 24-hour Sexual and Intimate Partner Violence Crisis Line for Mason is 703-380-1434.

Other Useful Resources

- University Catalog: https://catalog.gmu.edu/
- University Policies: https://universitypolicy.gmu.edu/
- Calendar of Religious Holidays and Observations: https://ulife.gmu.edu/religious-holiday-calendar/
- Office of Diversity, Inclusion and Multicultural Education: https://odime.gmu.edu/
- University Registrar (Calendars): http://registrar.gmu.edu/calendars/
- Counseling and Psychological Services: http://caps.gmu.edu/
- Learning Services: http://caps.gmu.edu/learning-services/
- University Career Services: http://careers.gmu.edu/
- The Writing Center: http://writingcenter.gmu.edu/
- University Libraries: http://library.gmu.edu/
- InfoGuides (Business): http://infoguides.gmu.edu/business
- School of Business, George Mason University: http://business.gmu.edu/
- ISOM Area, School of Business, George Mason University: http://business.gmu.edu/academic-departments/isom/

HONOR CODE

To promote a stronger sense of mutual responsibility, respect, trust, and fairness among all members of the George Mason University community and with the desire for greater academic and personal achievement, we, the student members of the University Community have set forth this:

Student members of the George Mason University community pledge not to cheat, plagiarize, steal and/or lie in matters related to academic work. The Honor Code of George Mason University deals specifically with cheating and attempted cheating, plagiarism, lying, and stealing.

A. Cheating
Cheating encompasses the unauthorized use of, access to, or provision of academic work in an attempt to misrepresent a student’s actual efforts. This
includes submitting another individual’s work for a grade, soliciting solutions/assignments from online websites, unauthorized collaboration, or failing to adhere to requirements (verbal and written) established by the professor of the course. Subcategories of cheating include:

- Use of unauthorized material
- Use of unauthorized assistance
- Duplicate use of student’s work
- Providing or attempting to benefit from unauthorized academic material
- Submission of another individual’s work
- Violation of course requirements regarding integrity

B. Plagiarism

Plagiarism is defined as using another individual’s ideas or words without attribution or credit. It also includes using prior work that has been submitted for credit or published in another venue as a new submission without citation. Using the ideas of others without proper attribution or citation is unethical and a violation of the Honor code. Subcategories of plagiarism include but not limited to:

- Self-plagiarism
- Inadequate citation
- False citation
- Failure to quote sources or material

Plagiarism does not include mistakes in the format of a citation as long as the student has clearly indicated the materials quoted or relied upon and the source of the materials.

C. Stealing

Stealing from an academic perspective means obtaining and/or providing unauthorized access to educational materials. These materials might be tests or quizzes from faculty members, or they may be the work product of another student. Subcategories of stealing include:

- Removing an exam or other academic work from a classroom without authorization
- Taking photos of exams/academic work without authorization or permission
- Taking someone else’s work without their knowledge

D. Lying

Lying in an academic context refers to providing information known to be false as a way to bypass classroom expectations or gain an unfair advantage in completing academic work. Subcategories of lying include:

- Falsifying sources, data, or information
- Providing a false excuse for missing a test or class
- Providing false information, including identifying information
- Falsifying official correspondence

Source: George Mason University Honor Code 2018-2019 final
<table>
<thead>
<tr>
<th>Week No</th>
<th>Date</th>
<th>Topics (Tentative)</th>
<th>DUE (subject to change)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WEEK 1</td>
<td>Jan 23</td>
<td>Syllabus &amp; Introduction</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Refresher of Database Concepts I (Ch 1 - 6)</td>
<td></td>
</tr>
<tr>
<td>WEEK 2</td>
<td>Jan 30</td>
<td>Refresher of Database Concepts II (Ch 1 - 6)</td>
<td></td>
</tr>
<tr>
<td>WEEK 3</td>
<td>Feb 6</td>
<td>Refresher of Database Concepts II (Ch 1 - 6)</td>
<td>Homework 1 - Database Concepts</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Data Warehousing Concepts (Ch 7)</td>
<td></td>
</tr>
<tr>
<td>WEEK 4</td>
<td>Feb 13</td>
<td>Data Warehousing Concepts (Ch 7)</td>
<td>Quiz 1 (ch 1 – 6)</td>
</tr>
<tr>
<td>WEEK 5</td>
<td>Feb 20</td>
<td>Data Warehouse and Data Mart Modeling (Ch 8)</td>
<td></td>
</tr>
<tr>
<td>WEEK 6</td>
<td>Feb 27</td>
<td>Data Warehouse and Data Mart Modeling (Ch 8)</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>--------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>WEEK 7</td>
<td>Mar 6</td>
<td>Data Warehouse and Data Mart Modeling (Ch 8)</td>
<td>Quiz 2 (ch 7)</td>
</tr>
<tr>
<td>WEEK 8</td>
<td>Mar 13</td>
<td>Spring Break</td>
<td></td>
</tr>
<tr>
<td>WEEK 9</td>
<td>Mar 20</td>
<td>Data Warehouse and Data Mart Modeling (Ch 8)</td>
<td></td>
</tr>
<tr>
<td>WEEK 10</td>
<td>Mar 27</td>
<td><strong>Midterm Exam (ch 7 – 8)</strong></td>
<td>Homework 2 - Dimensional Modeling</td>
</tr>
<tr>
<td>WEEK 11</td>
<td>Apr 3</td>
<td>Data Warehouse Implementation and Use (Ch 9)</td>
<td></td>
</tr>
<tr>
<td>WEEK 12</td>
<td>Apr 10</td>
<td>Data Warehouse Implementation and Use (Ch 9)</td>
<td>Introduce BI Tool - Tableau</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quiz 3 (ch 8 &amp; 9)</td>
<td></td>
</tr>
<tr>
<td>WEEK 13</td>
<td>Apr 17</td>
<td>Hands On Lab - ERDPlus &amp; SQL Server</td>
<td>Homework 3 - OLAP Exercises</td>
</tr>
<tr>
<td>WEEK 14</td>
<td>Apr 24</td>
<td>Hands On Lab - SSMS</td>
<td>Quiz 4 (ch 9)</td>
</tr>
<tr>
<td>WEEK 15</td>
<td>May 1</td>
<td>Hands On Lab - Tableau</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Review for Exam</td>
<td>Homework 4 - Lab Results</td>
</tr>
<tr>
<td>WEEK 16</td>
<td>May 8</td>
<td><strong>Final Exam (Cumulative)</strong></td>
<td></td>
</tr>
</tbody>
</table>