DATA MANAGEMENT FOR ANALYTICS

MSA 8040
CRN 95019 (Blended) or CRN 95147 (Online)
Thursday, 6:00pm—8:30pm, Room 1108 @ 55 Park Place
Dr. Houping Xiao

Course Description

Welcome to MSA 8040: Data Management for Analytics! We’re going to have a great time this semester exploring database and some unstructured data analytics tools together. My name is Houping Xiao and I’ll be your professor. You can call me Houping. I’ve been teaching this course for 2 years. I always enjoy teaching Data Management for Analytics because I am able to learn new technology about database and data analytics. Here’s a little about me and how I got into data analytics. I got a Ph.D. in computer science with focuses on data mining and machine learning and join Robinson college right after graduation. One of my major research interests in Robinson is to apply data mining and machine learning techniques to solve realistic business problems. I’m looking forward to working with all of you this semester.

The course, Data Management for Analytics, provides an introduction to the database management and data analytics. The course emphasizes the understanding of the fundamentals of database design, ranging from the relational database model to entity relationship modeling, to normalization and to implementation, including the structured query language (SQL). The course also provides an understanding of new developments and trends such as No-SQL and database on AWS. Finally, the course will deliver data analytics tools, including web scraping, Topic Modeling and Sentiment Analysis for unstructured data. The course uses a problem-based approach to learning.

Course Goals

1. The students will be confident to solve Database problems in their job interviews
2. The students will be able to build or query Database in their future jobs
3. The students will be well prepared for MSA 8770 Text Analytics, offered in Fall 2021
4. The students will work effectively as a group leader or member
5. The students will develop advanced presentation skills in presenting their works

Student Learning Outcomes (SLO)

After finishing this course,

1. SLO1: When asked about the concepts in database, students should be able to:
   • SLO1.1: Explain in own words concepts in both SQL and NO-SQL
   • SLO1.2: Differentiate between SQL and NO-SQL
2. SLO2: When asked about the concepts in database, students should be able to:
   • SLO2.1: Build a SQL, NoSQL database in their devices or on AWS
• SLO2.2: Implement query in both SQL, NO-SQL databases, and database on AWS

3. SLO3: When asked about the data analytics, students should be able to:

• SLO3.1: Differentiate between structured and unstructured data
• SLO3.2: Extract useful information for unstructured data using Web Scraping
• SLO3.3: Differentiate between unsupervised and supervised learning framework, and be able to implement topic model and sentiment analysis for textual data

Assessments
I’ve designed a variety of assessments to help you practice your disciplinary thinking and skills. More detailed instructions for course assessments can be found in the Welcome To The Course Module in iCollege.

<table>
<thead>
<tr>
<th>Class Participation</th>
<th>10</th>
<th>15 or 15% of Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mid-term Exam</td>
<td>1</td>
<td>25 or 25% of Grade</td>
</tr>
<tr>
<td>Projects</td>
<td>3</td>
<td>60 or 60% of Grade</td>
</tr>
</tbody>
</table>

Grade Breakdown
This is how this course awards grades for courses:

<table>
<thead>
<tr>
<th>Percentage</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>96.5 – 100</td>
<td>A+</td>
</tr>
<tr>
<td>92.5 – 96.4</td>
<td>A</td>
</tr>
<tr>
<td>89.5 – 92.4</td>
<td>A-</td>
</tr>
<tr>
<td>85.5 – 89.4</td>
<td>B+</td>
</tr>
<tr>
<td>82.5 – 85.4</td>
<td>B</td>
</tr>
<tr>
<td>79.5 – 82.4</td>
<td>B-</td>
</tr>
<tr>
<td>69.5 – 79.4</td>
<td>C+</td>
</tr>
<tr>
<td>59.5 – 69.5</td>
<td>C</td>
</tr>
<tr>
<td>55.0 – 59.4</td>
<td>C-</td>
</tr>
<tr>
<td>50.0 – 54.9</td>
<td>D</td>
</tr>
<tr>
<td>0 – 49.9</td>
<td>F</td>
</tr>
</tbody>
</table>

COURSE HELP

Note that this is a Blended course for CRN 95019 or 100% online course for CRN 95147. Here are a few tips to get you started:

How Do I Contact You?
I prefer to be contacted in the following way(s):

- iCollege Email
- GSU Email: hxiao@gsu.edu
- Online Office Hours, (details will be posted on iCollege)

How Do I Access My Course?
You can login to your course via iCollege. If you need more help, you can review the Welcome to iCollege help-guide. If you have problems accessing your course, please contact Ian Schonberg, Assistant Director of Graduate Advisement or me.

What Are The Required and Optional Materials?
Textbooks:

Recommended References:

Are There Any Required Meetings?
- **Blended Section:** Each week, we will only allow 25% of the room capacity (no more than 12 students due to the capacity of 50 at Room 1108, 55 Park Place) to meet face-to-face. For other students, we will meet virtually.
- **Online Section:** We will meet virtually for every week. The information about how to join virtually will be released soon in iCollege.

How Do I Succeed in this Course?
The course materials delivered in both blended and online sections are exactly the same. Both sections are synchronized. You will join live to go through lecture notes with me. The course consists of both theoretical and application parts. I will encourage students to practices online frequently. The practices websites will be informed in the class.

SCHEDULE

Although this is a blended or online course, we do have a set schedule. Please note that deviations may become necessary as the semester progresses. You'll want to refer to the calendar below frequently as we work together. I've also designed the iCollege course in such a way to help us all stay on track, including building in Slides, Code, and References, Etc.). If this is your first time taking a course in iCollege, you'll want to review the Welcome to iCollege help-guide. Please note that deviations may become necessary as the semester progresses.

This course is divided into 13 topics, along with 3 projects. You'll have around 4 of weeks to complete each project. For detailed information about what's required for each project, visit iCollege and check out the project Introductions. While you're working, I'll also be working hard to give you quality feedback and grade your assessments by the dates indicated below.

So, how much time do you need to spend working on this course? Well, since this is a 3-Credit Hour course GSU recommends that you spend around 3 hours or more per week interacting with readings, videos, and other sorts of content and then 3 hours per credit hour per week completing activities and assessments.
Please talk to your instructor and your advisor before withdrawing from this course. We care about your success and are here to discuss your options with you. The withdrawal period is SEP 02 – OCT 13.

**Tentative Course Schedule (Topics)**

The course syllabus provides a general plan for the course; deviations may be necessary.

<table>
<thead>
<tr>
<th>Week</th>
<th>Date</th>
<th>Topic</th>
<th>Corresponding Readings</th>
<th>Project</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>08/27/2020</td>
<td>Introduction to Database: Database and design concepts, Relational database model</td>
<td>Chapter 1, 2, &amp; 9</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>09/03/2020</td>
<td>SQL: ER Model, ER diagrams, normalization, and SQL</td>
<td>Chapter 3, 4, &amp; 6</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>09/10/2020</td>
<td>MySQL: Introduction to MySQL Server and Workbench, Data types</td>
<td>Chapter 7, 8, &amp; Online materials</td>
<td>P1 Out</td>
</tr>
<tr>
<td>4</td>
<td>09/17/2020</td>
<td>MySQL: Functions, Operators and Statement Syntax</td>
<td>Chapter 7, 8, &amp; Online materials</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>09/24/2020</td>
<td>MySQL: SQL</td>
<td>Chapter 7, 8, &amp; Online materials</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>10/01/2020</td>
<td>MySQL: Advanced SQL</td>
<td>Chapter 7, 8, &amp; Online materials</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>10/08/2020</td>
<td>MySQL: Procedure &amp; Trigger</td>
<td>Online materials</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>10/15/2020</td>
<td>Mid-term Exam</td>
<td></td>
<td>P1 Due</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P2 Out</td>
</tr>
<tr>
<td>9</td>
<td>10/22/2020</td>
<td>Feedback on Mid-term Exam Introduction to Unstructured data analysis And MongoDB</td>
<td>Online materials</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>10/29/2020</td>
<td>MongoDB: CRUD &amp; Aggregation</td>
<td>Online materials</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>11/05/2020</td>
<td>Web Scraping</td>
<td>Online materials</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>11/12/2020</td>
<td>Database on AWS</td>
<td>Online materials</td>
<td>P2 Due</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P3 Out</td>
</tr>
<tr>
<td>13</td>
<td>11/19/2020</td>
<td>Data Mining and Machine Learning: Sentiment analysis and Topic Modeling I</td>
<td>Online materials</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>11/26/2020</td>
<td>Thanksgiving Break</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>12/03/2020</td>
<td>Data Mining and Machine Learning: Sentiment analysis and Topic Modeling II</td>
<td>Online materials</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>12/10/2020</td>
<td>Final Presentation and Report Due</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
I have developed several policies that seem to work well in this course. Please review these very closely. You’ll have an opportunity to voice your opinion on these policies and other aspects of the course when we reach evaluation points during the semester.

**Attendance Policy**

This course will be offered synchronously in person and on Zoom or Webex.

**Blended Section:** Students are expected to attend their assigned in-person class sessions. Attendance and participation will count towards 15% of the final course grade. There would be no penalty for joining lecture via Zoom or Webex, but students need to notify their instructor in advance. For the remaining students, they can still **virtually** attend a class session they are not assigned to attend in-person via Zoom or Webex. Students If a student develops a fever, cough or shortness of breath they should not go to class and should notify their instructor immediately. The student will work with the instructor to develop a plan to complete the necessary course content, activities, and assessments in order to meet the course student learning outcomes. The attendance policy for this class is in alignment with the “Policy on Class Attendance” in the Student Code of Conduct.

**Online Section:** Students are expected to attend class sessions virtually via Zoom or Webex. Attendance and participation will count towards 15% of the final course grade.

**Make-up Exam Policy**

There is one mid-term exam in this course. Date for the exam is already set on the Tentative Course Schedule below. If there is an excusable reason for being unable to be present during the exam dates, please let me know as soon as possible to schedule a make-up exam. The make-up exam if at all possible, will take place before the scheduled exam date. Students with unexcused absences for an exam will earn a 0 on the exam.

**Project Submission Policy**

Late submission policy: Each day an assignment is late, it is marked down 20%. All submissions have to be submitted through iCollege.

**Course Evaluation and Evolution**

The evaluation for the mid-term exam and 2 individual projects will be provided in the following week. And, online appointments with me are available if you want to discuss your performance.

Your constructive assessment of this course plays an indispensable role in shaping education at Georgia State. Upon completing the course, please take time to fill out the online course evaluation.
OTHER POLICIES

Academic Honesty
Being responsible for your own learning does not mean that you must always work in isolation. However, when working in groups we encourage you to be mindful of how much effort and learning you are experiencing. Below, we outline our expectations for work in this course.

For projects, I encourage students to work together to solve and understand the problems. Nevertheless, each student is responsible for demonstrating he or she has good grasp of the material. Ultimately, each student’s project solution should reflect his or her own learning and be written in the students’ own words. While students may work together to figure out how to solve the problems, each student must run his or her own analyses and turn in their own output. For the Exam, each should work independently, no discussion is allowed. Under no circumstance should a student email his or her project solutions, project reports, and codes to a classmate. Working together (for the project) is for the purpose of collaborating, not copying.

“As members of the academic community, students are expected to recognize and uphold standards of intellectual and academic integrity.” As listed on https://deanofstudents.gsu.edu/student-conductpolicy-on-academic-honesty/.

Special Needs
Students who wish to request accommodation for a disability may do so by registering with the Office of Disability Services. Students may only be accommodated upon issuance by the Office of Disability Services of a signed Accommodation Plan and are responsible for providing a copy of that plan to instructors of all classes in which accommodations are sought.

Students with special needs should then make an appointment with me during the first week of the class to discuss any accommodations that need to be made.

FERPA
In keeping with USG and university policy, this course website will make every effort to maintain the privacy and accuracy of your personal information. Specifically, unless otherwise noted, it will not actively share personal information gathered from the site with anyone except university employees whose responsibilities require access to said records. However, some information collected from the site may be subject to the Georgia Open Records Act. This means that while we do not actively share information, in some cases we may be compelled by law to release information gathered from the site. Also, the site will be managed in compliance with the Family Educational Rights and Privacy Act (FERPA), which prohibits the release of education records without student permission.

Sexual Harassment
In instances of sexual misconduct, the present instructor(s) and teaching assistants, are designated as Responsible Employees who are required to share with administrative officials all reports of sexual misconduct for university review. If you wish to disclose an incident of sexual misconduct confidentially, there are options on campus for you do so. For more information on this policy, please refer to the Sexual Misconduct Policy which is included in the Georgia State University Student Code of Conduct.
Basic Needs Statement
Any student who faces challenges securing their food or housing and believes this may affect their performance in the course is urged to contact the Dean of Students for support. Furthermore, please notify the professor if you are comfortable in doing so. This will enable us to provide resources that we may possess. The [Embark program at GSU](#) provides resources for students facing homelessness.

POLICIES IN TERMS OF COVID 19

Face coverings
Georgia State University and the University System of Georgia have mandated a face covering policy. Students are required to wear an appropriate face covering while inside campus facilities (classrooms, hallways, elevators, labs and in all other public spaces) because six feet of social distancing may not always be possible. Face coverings will be worn in addition to and not as a substitute for social distancing. Face coverings aren’t required in one’s own dorm room or suite, when alone in an enclosed office or study room, or in campus outdoor settings where social distancing requirements are met.

Anyone not using a face covering when required will be asked to wear one or leave the area. Repeated refusal to comply with the requirement may result in discipline through the applicable conduct code for students. Reasonable accommodations may be made for those who are unable to wear a face covering for documented health reasons. To request an accommodation, start with the Access & Accommodations Center Welcome Form found at How to Connect.

When you wear a face covering, you’re protecting yourself, but equally important, you’re protecting other members of the university community. You’re showing you’re taking responsibility for the welfare of those around you.

The [university’s policy on face coverings](#) is available in the university policy library.

Learner-centered language from Senate FAC you may choose to include:
I will be wearing my face mask in class this semester, and you are required to wear yours. I know that face masks may make some aspects of class more difficult. It will be harder for us all to project our voices and read each other’s facial expressions. However, I am willing to sacrifice these elements since wearing a mask is one thing I can control to support the health and safety of our community. This is a new situation for all of us and if we start the semester off from a place of mutual understanding it will be easier for us all to get through this together. Be aware that wearing face mask is required by GSU and there is a penalty if you choose to not wear a mask. Our university community has a strong tradition of upholding the value of mutual respect; we therefore ask students to not engage in behavior that would be disruptive to class regarding this policy.

Assigned seats:
On the first day of in-person class, students should sit only in seats indicated by a blue GSU “sit here” label. Instructors will then create a seating chart. Students should sit in their
assigned seats for each in-person class session. If a student becomes ill, the seating chart will help identify those who may have been in close contact.

**Student illness**
If a student develops a fever, cough or shortness of breath they should stay at home, not go to class or work, and stay away from other people. If a student becomes sick or is required to quarantine during the semester, they should notify their instructor as soon as possible. The student will work with the instructor to develop a plan to complete the necessary course content, activities, and assessments in order to meet the course student learning outcomes.

Students in residence halls who report an infection will be asked to return home. For those unable to do so, GSU will provide a sequestered area that will allow residential students to isolate while infected.