Northeastern University
College of Engineering

DAMG 7370 Data Management and Database Design

Course Information
Course Title: Data Management and Database Design
Course Number: DAMG 7370
Term and Year: Spring 2024
Credit Hour: 4
Course Format: Online

Instructor Information
Full Name: Naveen Kuragayala
Email Address: n.kuragayala@northeastern.edu

Course Prerequisites
DAMG 6210 (Minimum Grade of B)

Course Description
Focuses on designing advanced data architectures supporting structured and semistructured data sources on cloud (**Azure** and **AWS**) and on-premise. Data preparation and data integration tools are used to gather and integrate data. Business Intelligence (BI) used for data visualization and data analysis.

Topics include data architcture; designing data models (dimension and hydrid dimensional); developing data integration and data preparation worflows to gather, integrate and load "curated" data; designing BI data visualizations and reports for data analysis. Technologies include databases; hybrid data integrations and cloud integration; data preparation; data virtualization; and on-premise and on-cloud deployments.

By end of this course you will complete tool-based tutorials along with targeted hands-on labs, assignments and projects.

Standard Learning Outcomes
Learning outcomes common to all College of Engineering Graduate programs:

1. *An ability to identify, formulate, and solve complex engineering problems.*
2. *An ability to explain and apply engineering design principles, as appropriate to the program's educational objectives.*
3. *An ability to produce solutions that meet specified end-user needs with consideration of public health, safety, and welfare, as well as global, cultural, social, environmental, and economic factors.*

The Information Systems Program accepts students of different engineering backgrounds with minimum programming skills and produces first class Information Systems engineers that operate at the intersection of real-world complexity, software development, and IT management. Graduating students will be able to construct end-to-end advanced software applications that meet business needs.

Specific Learning Outcomes for the Information Systems program:

1. *Create a strong technical foundation through diverse, high-level courses*
2. *Built crucial interpersonal skills needed to succeed in any industry*
3. *Foster a deep level of applied learning through project based case studies*

Reference Books

Business Intelligence Guidebook: From Data Integration to Analytics
Author: Rick Sherman, Published by: Morgan Kaufmann

The Kimball Group Reader
Author: Ralph Kimball, Margy Ross, Warren Thornthwaite, Joy Mundy, Bob Becker, Published by: Wiley

Resources for help

Slack for communication and help TA's during their office hours

Course Grading

All assignments / mid-term and final examinations, presentations will receive a numerical value. Midterm and Final exam must be taken at the time and location determined by the schedule and announced in class. Grading is based on absolute grading system. You will have about 5 days to get your assignment graded by the TA. Make up exams permitted only for exceptional circumstances in accordance with NEU policy.

Assessment – Tentative	Allocation %
Lectures, Readings, Workshops, Assignments & Quizzes	40%
Team Projects	15%
Midterm	20%
Final Team Project	25%

Grades scored %	Grade
94% and 100%	A
90% and < 94%	A-
87% and < 90%	B+
84% and < 87%	B
80% and < 84%	B-
77% and < 80%	C+
74% and < 77%	C
70% and < 74%	C-
0% and < 70%	F

Topics –

Note: This course uses Cloud (**Azure** and **AWS**) and Installation based softwares and tools on personal laptop/notebook which student has to install, configure and if required student has to **purchase the credits** for learning and projects submissions. Atleast 16GB RAM laptop is required. Windows is highly preferred however MAC also works with few work additional tools installations (Mac Students may need to purchase virtualization. In most scenarios we use free software however dependeing of chipset student might endup buying a virtualization software. For students there is a discount).

Week	Weekly Topics	Online Lecture Reading Chapters	Workshops & Tutorials	Quizzes
1	• Course Overview • Architecture High-Level Overview (Ch 4) • Defining requirements		• BI Introduction -- quick review of data sources & tools used -- tool basics & querying different source types • Data Profiling	
2	• BI Applications • Defining requirements	• Ch 13 BI Applications • Ch 03 Defining requirements	• Data Viz Basics & Discussion • Workshop: Tableau & Power BI	• Ch 13 • Ch 03
3	• Dimensional Modeling • Databases & Data Structures	• Ch 09 Dimensional Modeling • Ch 08 Foundational Modeling (ER Modeling)	• Review concepts & hands-on: workshop - Data Modeling	• Ch 09 • SQL
4	• BI Dimensional Modeling • Data Preparation	• Ch 10 BI Dimensional Modeling	• Review concepts & hands-on: workshop - Data Modeling - Data Preparation	• Ch 10
5	• Data Architectures • Data Engineering & Data Integration • BI, Analytics • Deployment (Cloud, On-Premise & Hybrid)	• Ch 06 Data Architecture • Data Architecture: Cloud, BigData • Ch 07 Technology Architecture	• Review concepts & hands-on: workshop - Data Integration	• Ch 06
6	• Data Integration • Data Preparation	• Ch 11 Data Integration Design & Development • Ch 12 Data Integration Processes • Data Integration Tutorials	• Review concepts & hands-on: workshop - Data Integration	• Ch 11 • Ch 12
7	• Cloud Data Architectures & Cloud DW • Data Integration	• Cloud Data Architecturesm & Cloud DW • Data Integration Tutorials	• Review concepts & hands-on: workshop - Data Integration	• Data Integration
8	Midterm Exam (4 Parts) P1: Quiz MCQs, P2: Theory, P3: Data Modeling, P4: Team Project			
9	• Data Integration	• Data Integration Tutorials	• 1st Team Project Review • Data Integration	• Data Integration
10	• BI Data Models & Data Visualizations	• Ch 14 BI Design & Development	• BI Best Practices Recap & Discussion • Workshop: Tableau, Power BI	• Ch 14
11	• BI Design & Development Expanded	• Self-Service BI (SSBI)	• Workshop: Tableau, Power BI • Workshop: Data Integration	• BI Design
12	• Self-Service BI (SSBI) • Data Shadow Systems	• Ch 16 Data Shadow Systems	Catch up	• Ch 16
13	• BI Architecture (Data Lakes, Sandboxes) • Data Integration Workflow	• Ch 15 Advanced Analytics (Architecture)	• Review of Final Project	• Ch 15
14	Final Presentation Demos			
15	Final Presentation Demos			

Software & Development Environments:

- Business Intelligence (data discovery & data visualization)
    - Tableau Cloud & Tableau Desktop
    - Microsoft PowerBI

- Data Integration
    - Azure Data Factory
    - Talend Cloud Data Integration - Big Data & Cloud Integration, ETL/ELT
    - Alteryx - Data Preparation

- Data Modeling
    - ER/Studio Data Architect

- Databases (used as source or target systems)
    - Cloud:
        - Microsoft Azure SQL (Relational)
        - Synapse & Databricks (Preview)
        - Snowflake (Preview)

    - Relational (on your notebook:)
        - Microsoft SQL Server 2019 Developers' Edition (Relational)
        - MySQL & PostgreSQL (Relational)
    - Files: csv, tsv, json, Parquet (data lake) (S3/ADLS2)

Notes:

- The above software will be licensed full functioning versions NOT trials with the exception of cloud platforms that use trial credits or academic credits: Microsoft, Google, Oracle
- Windows (or Mac running virtualization software with Windows) fyi many students have used Mac in this cousre
    - For Mac with Intel chip: Macs Boot camp is fastest option
    - For Mac with M1/M2 chip: Parallels Desktop for Mac Student Edition
        - Recommended to Buy virtualization software

Class Structure:

- Lectures & Tutorials (Weekly)
- Hands-on Workshops (Weekly)
- Class Collaborate Sessions
- Design reviews
- Q&A
- Supplemental Readings & Tutorials

Tests:

- Quizzes (based on lectures & readings)
- Midterm & Final Exams

Deliverables:

- Team Projects
- Workshops & homework Assignments

Workshops, assignments & team projects involve:

- Data Design, Modeling and Integration
- Data Visualization
- Data Profiling

Final Project:

Project deliverables include data design, data modeling data integration and BI/analytics dealing with multiple data sources.

*Knowledge Prerequisites (**Must**):*

- A fundamental understanding of databases, relational dbms and SQL
- A basic understanding of Entity-Relationship (ER) Modeling

End-of-Course Evaluation Surveys

Your feedback regarding your educational experience in this class is very important to the College of Professional Studies. Your comments will make a difference in the future planning and presentation of our curriculum.

At the end of this course, please take the time to complete the evaluation survey at **https://neu.evaluationkit.com**. Your survey responses are **completely anonymous and confidential**. For courses 6 weeks in length or shorter, surveys will be open one week prior to the end of the courses; for courses greater than 6 weeks in length, surveys will be open for two weeks. An email will be sent to your HuskyMail account notifying you when surveys are available.

Academic Integrity

Midterm and Final exams are individual effort and collaboration of any kind would be considered violation. For group activities, you are strongly encouraged to work with your group and submissions must be done by each individual. Don't share your code to other teams.

A commitment to the principles of academic integrity is essential to the mission of Northeastern University. The promotion of independent and original scholarship ensures that students derive the most from their educational experience and their pursuit of knowledge. Academic dishonesty violates the most fundamental values of an intellectual community and undermines the achievements of the entire University.

As members of the academic community, students must become familiar with their rights and responsibilities. In each course, they are responsible for knowing the requirements and restrictions regarding research and writing, examinations of whatever kind, collaborative work, the use of study aids, the appropriateness of assistance, and other issues. Students are responsible for learning the conventions of documentation and acknowledgment of sources in their fields. Northeastern University expects students to complete all examinations, tests, papers, creative projects, and assignments of any kind according to the highest ethical standards, as set forth either explicitly or implicitly in this Code or by the direction of instructors.

Go to http://www.northeastern.edu/osccr/academic-integrity-policy/ to access the full academic integrity policy.

Student Accommodations

Northeastern University and the Disability Resource Center (DRC) are committed to providing disability services that enable students who qualify under Section 504 of the Rehabilitation Act and the Americans with Disabilities Act Amendments Act (ADAAA) to participate fully in the activities of the university. To receive accommodations through the DRC, students must provide appropriate documentation that demonstrates a current substantially limiting disability.

For more information, visit http://www.northeastern.edu/drc/getting-started-with-the-drc/.

Library Services

The Northeastern University Library is at the hub of campus intellectual life. Resources include over 900,000 print volumes, 206,500 e-books, and 70,225 electronic journals.

For more information and for Education specific resources, visit http://subjectguides.lib.neu.edu/edresearch.

24/7 Blackboard Technical Help

For immediate technical support for Blackboard, call 617-373-4357 or emailhelp@northeastern.edu

Within Blackboard, open a support case via the red support button on the right side of the screen, click Create Case

myNortheastern, e-mail, and basic technical support
Visit the Information Technology Services (ITS) Support Portal
Email: help@northeastern.edu
ITS Customer Service Desk: 617-373-4357

Diversity and Inclusion

Northeastern University is committed to equal opportunity, affirmative action, diversity and social justice while building a climate of inclusion on and beyond campus. In the classroom, member of the University community work to cultivate an inclusive environment that denounces discrimination through innovation, collaboration and an awareness of global perspectives on social justice.

Please visit http://www.northeastern.edu/oidi/ for complete information on Diversity and Inclusion

TITLE IX

Title IX of the Education Amendments of 1972 protects individuals from sex or gender-based discrimination, including discrimination based on gender-identity, in educational programs and activities that receive federal financial assistance.

Northeastern's Title IX Policy prohibits Prohibited Offenses, which are defined as sexual harassment, sexual assault, relationship or domestic violence, and stalking. The Title IX Policy applies to the entire community, including male, female, transgender students, faculty and staff.

In case of an emergency, please call 911.

Please visit www.northeastern.edu/titleix for a complete list of reporting options and resources both on- and off-campus.