DAMG 7370: Designing Advanced Data Architectures for Business Intelligence

Course Information
Course Title: **Designing Advanced Data Architectures for Business Intelligence**
Course Number:DAMG 7370
Term and Year: SPRING2025
Credit Hour: 4
CRN:
Course Format: On-Ground

Instructor Information
Full Name: Naveen Kuragayala
Email Address: n.kuragayala@northeastern.edu
Office Hours: By Appointment

Instructor Biography
A renowned data professional with over 20+ years of experience in the field. expertise in areas such as data migrations, data processing, data analysis, visualization and integration. worked with organizations across various industries, such as healthcare, e-commerce and finance helping them design and implement robust data management solutions. In my courses I facilitate real-time use cases. I focus more on core areas such as database management systems, data modeling methodologies, SQL programming languages, Data warehousing Architectures,Data Engineering and Data Analytics.

Teaching Assistant Information
Refer to canvas for TA details

Course Prerequisites
Data Management and Database Design

Course Description
Focuses on designing advanced data architectures supporting structured and semistructured data sources on cloud (Azure mandatory and AWS optional) 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.

Note: This course uses Cloud (Azure / 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 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).

Course Learning Outcomes
1) **Advanced Data Architecture Design**: Students will be able to design advanced data architectures that support both structured and semi-structured data sources, leveraging cloud platforms as well as on-premise solutions to create efficient and scalable data models.

2) **Proficiency in Data Preparation and Integration**: Students will gain proficiency in using data preparation and integration tools to gather, integrate, and load curated data, developing comprehensive data workflows for business intelligence applications. Will introduce Python notebooks to use on Databricks. Complete hands to get started with Databricks and Engineering workspace, Databricks notebook operations, Integrating with GIT, Transform data with Spark and Delta live tables. Additionally will introduce the data access using unity catalog.

3) **BI Data Visualization and Reporting Skills**: Students will obtain skills to create effective BI data visualizations and reports for data analysis, utilizing technologies such as databases, hybrid data integrations, cloud integration, data preparation, and data virtualization.

Required Tools and Course Textbooks.
Business Intelligence Guidebook: From Data Integration to Analytics (Must buy this Book)
Author: Rick Sherman, Published by: Morgan Kaufmann
The Kimball Group Reader
Author: Ralph Kimball, Margy Ross, Warren Thornthwaite, Joy Mundy, Bob Becker

Software & Development Environments:
- Business Intelligence (Data discovery & Data visualization)
  - Tableau
  - Microsoft PowerBI
- Data Integration
  - Databricks (PySpark preview) and Unity catalog
  - Azure Data Factory (Data flow scripting)
  - Talend Cloud Data Integration - Big Data & Cloud Integration, ETL/ELT
  - Alteryx - Data Preparation
- Data Modeling
  - ER/Studio Data Architect
  - Navicat
- Cloud Databases (used as source / target)
  - Snowflake
  - Databricks SQL endpoint using Deltalake tables
  - Azure SQL Database
- Databricks Notebook

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

Course Schedule/Topics Covered

#	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 • Datalake implementations	• Partitioning • Datalake / Delta lake / Iceberg	• Ch 16
13	• BI Architecture (Data Lakes, Sandboxes) • Data Integration Workflow	• Ch 15 Advanced Analytics (Architecture)	• Review of Final Project	• Ch 15
14 15	Final Presentation Demos			

Assignment Grading

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

Grading Scale

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

Attendance/Late Work Policy

Attendance Policy

Students registered in MGEN courses (INFO, CSYE, and DAMG) are allowed **a maximum of 2 absences per course, with 3 or more absences resulting in an automatic 'F' for that course.** Students are expected to inform their instructors of any absences in advance of the class; if a student is sick long-term or experiences a medical issue that prevents class attendance, it is strongly encouraged that they speak with their Academic Advisor (coe-mgen-gradadvising@northeastern.edu) to learn more about the Medical Leave of Absence. Should a student anticipate being unable to attend 3 or more classes, they should discuss their situation with their Academic Advisor to explore other types of leave in accordance with the University's academic and global entry expectations. International students should review the Office of Global Services webpage to understand their visa compliance requirements.

Teaching Assistants (TAs) or Instructional Assistants (IAs) will be present at each class to collect student attendance.

Late Work Policy

Students must submit assignments by the deadline in the time zone noted in the syllabus. Students must communicate with the faculty prior to the deadline if they anticipate work will be submitted late. Work submitted late without prior communication with faculty will not be graded.

End-of-Course Evaluation Surveys
Your feedback regarding your educational experience in this class is particularly important to the College of Engineering. 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 Northeastern University Mail account notifying you when surveys are available.

Academic Integrity
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.

MGEN Student Feedback
Students who would like to provide the MGEN unit with anonymous feedback on this particular course, Teaching Assistants, Instructional Assistants, professors, or to provide general feedback regarding their program, may do so using this survey: https://neu.co1.qualtrics.com/jfe/form/SV_cTIAbH7ZRaaw0Ki

University Health and Counseling Services
As a student enrolled in this course, you are fully responsible for assignments, work, and course materials as outlined in this syllabus and in the classroom. Over the course of the semester if you experience any health issues, please contact UHCS.

For more information, visit https://www.northeastern.edu/uhcs.

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 https://drc.sites.northeastern.edu.

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 https://library.northeastern.edu
Network Campus Library Services: Northeastern University Library Global Campus Portals

24/7 Canvas Technical Help

For immediate technical support for Canvas, call 617-373-4357 or email help@northeastern.edu

Canvas Student Resources: https://canvas.northeastern.edu/student-resources/

For assistance with my Northeastern e-mail, and basic technical support:
Visit ITS at https://its.northeastern.edu
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, members 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 **https://www.northeastern.edu/ouec** for a complete list of reporting options and resources both on- and off-campus.*