Northeastern University
College of Engineering

Multidisciplinary Graduate Engineering Course Syllabus

Course Information

Course Title	**Data Warheouseing and Business Intelligence**
Course Number	**DAMG 7290**
Term and Year	**Spring 2022**
Credit Hour	4
Location	**Online - https://northeastern.blackboard.com**

Instructor Information

Full Name	**Vincent Lattuada**
Email Address	**vlattuada@neu.edu**

Technical/Course Materials Requirements

- Business Intelligence Guidebook: From Data Integration to Analytics
  - Author: Richard Sherman
  - ISBN-13: 978-0124114616
  - ISBN-10: 012411461X
  - Publisher: Morgan Kaufmann
- Windows (or Mac running virtualization software with Windows)
- SQL Server Visual Studio Data Tools
- Python

Course Description/Prerequisite

Examines the technical and management aspects of building a data warehouse. Explores the architecture, infrastructure, processes, data quality, database design, and data analysis involved in building the data warehouse for business analysis. Management issues include business goals, tool selection, project management, personnel skills, training, and user requirements. Topics include dimensional data modeling, extraction/transformation/load processes, data quality problems, datamarts, operational data stores (ODS), staging databases, and online analytic processing (OLAP).

Student Learning/Course Outcomes (SLOs)

The goal of this course is to look at the application of a variety of data solutions to the healthcare industry. It will cover traditional relational databases in addition to new noSQL tooling. Solutions will also leverage the cloud and look at advancements in tooling with Python and a variety of plugins.

The course will begin with a quick review of relational database systems. Then, newer database technologies from the noSQL family will be explored. These systems will then be used to implement data warehouse and data lake techniques to develop a modern data and analytics platform. Key concepts will be covered from traditional warehouse design in addition to new tooling for the creation of a next generation system. Traditional ETL will be used as well as new data movement techniques. The data will be visualized with both traditional and new tools.

The class will be online / video streamed and we will make use of collaboration tools to facilitate project reviews. There will also be weekly live sessions that will also be recorded.

Objectives:

- Develop data stores with structured and unstructured data focused on providing solutions for analytics.
- Systems will be developed with a polyglot persistence leveraging the best storage for the data.
- Utilize tools such as Python, Mongo, SQL Server (SSIS, SSAS, SSRS), Talend both locally and in the cloud
- Gather requirements and develop specifications for designing a system
- Design database models for reporting and analytics
- Develop multidimensional OLAP models
- Use tooling to process and analyze data

Project:

Projects will focus on gathering requirements and developing a database to support an application. Initially we will run through a variety of exercises and then work to integrate individual learnings into a full analytics system.

Attendance Policy

Students are expected to complete course readings, participate in class discussions or other learning activities during the unit, and complete written assignments for each unit during the time of that unit. It is understood that there might be one week when active participation in ongoing class conversations and learning activities might be delayed. Beyond one week time, if there is an absence or lateness in participation (1) faculty must be notified in advance; (2) grades will be adjusted accordingly.

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.
Typically all assignments are due at 11:59 PM EST. If there is an online review session that will provide solutions the time will be 1 hour before the session.

Grading/Evaluation Standards

Grade Scale:

95-100%	**A**	**87-89.9%**	**B+**	**77-79.9%**	**C+**		
		84-86.9%	**B**	**74-76.9%**	**C**	**69.9% or below**	**F**
90-94.9%	**A-**	**80-83.9%**	**B-**	**70-73.9%**	**C-**		

Grade Breakdown:

Every item during the semester will have a point value based on the weight of the item in the semester. For examplea simple quiz or homework might be worth 40 points. A more complicated homework or quiz could be worh 100 points. Projects and tests will be worth more for example 300 points. This allows you to simply calculate your grade at any time. Note: participation in discussion boards / online session etc will count towards a participation grade. A high level summary is below. Note this is subject to change as course topics change and is only meant as a guideline.

Homework	20
Quizes	20
Tests	25
Projects	25
Participation	10

Course Schedule

- Overview and expectation setting
- SQL Review
  - Set based processing
  - Advanced SQL
  - Creating Databases and Tables
  - Design a Database
  - Foreign and Primary Keys
  - Indexes and Views
  - Surrogate and Natural Keys
- ETL
  - Traditional SSIS - Talend
  - Bulk Copy
- NoSQL
  - Mongo
  - Redis
  - Hadoop
- Other SQL Topics
  - Normal Forms
  - Programming in SQL
  - Cursors
- Data Warheouse Design
  - Staging
  - Error Handling
  - Aggregations
  - Star Schema
  - OLAP
- Business Intelligence Tools

        - OLAP / Multidimensional
        - Vizualization
    - Data Issues
        - Data Quality
        - Master Data Management
        - Data Cleansing
    - Next Generation Warehouses
        - NoSQL
        - Data Lakes
        - Lake Houses
        - Data Vault
        - Schema on read
    - Project
        - Analytics dashboard
        - Database modeling
        - Loading and maintaining data
        - Vizualizations
        - Requirements gathering and specifications
    - Additional Topics
        - Cloud (Azure / AWS / GCP)
        - Hardware implications on design

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.

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.

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.