



ISTM 6211 – Data Warehousing

Fall 2011

Schedule: Tuesday 7:00 – 9:40 PM (Funger Hall, Duques 350)

Instructor:

Margherita Bruni, MS, PMP
Office: Funger Hall, Bull Pen 515-H Main Campus
Office Hours: by appointment only

Contact Information:
E-mail: mbruni@gwmail.gwu.edu
Phone: 202-492-5377

Course Description:

The course provides an introduction to data warehousing (DW), dimensional data modeling, Extract Transform, Load (ETL) and Business Intelligence (BI). The course will be based on case studies, industry software based labs, and a design project.

Prerequisites:

ISTM 6202 Relational Databases

Course Objectives:

The objective of this course is to provide students with knowledge of techniques, methodologies, and tools used in data warehousing. The first part of the course will cover the various phases of the data warehouse lifecycle and design. Current and emerging technologies will be introduced and discussed in the second part of the course.

Software:

Microsoft Visio and Access. Please note that Visio and Access are available in most computer labs on campus, and use of the software in this course is required. Other specific BI/DW software will be leveraged over the web.

References:

Modern Data Warehousing, Mining, and Visualization
George M. Marakas
ISBN 9780131014596
Prentice Hall
www.prenhall.com/marakas

Building the Data Warehouse, Fourth Edition
William H. Inmon
Wiley

The Data Warehouse ETL Toolkit
Ralph Kimball, Joe Caserta
Wiley

The Data Warehouse Lifecycle Toolkit, 2nd Edition
Ralph Kimball, Margy Ross, Warren Thornthwaite, Joy Mundy and Bob Becker
Publisher: Wiley

The Data Warehouse Toolkit

Ralph Kimball

Margy Ross

Wiley

Agile Data Warehousing

Ralph Hughes

iUniverse, Inc

Data Warehouse Design

Matteo Golfarelli and Stefano Rizzi

McGraw Hill

Lectures:

All lecture notes will be available on Blackboard. A few articles and cases may be distributed during the semester for class discussion and/or assignment. You are expected to attend all classes, and will be responsible for assignments due on the days you do not attend class.

Grading:

The course grade will be a weighted average of assignments, a project, and one mid-term exam. The relative importance of the different components is given below.

In-class Exercises and Labs	33 %
Assignments	33 %
Project	33 %

Project:

The course will have a group project. The objective of the project will be to design and build a data mart to provide analytics to business users.

The project deliverables will include project documents (data model, architectural diagram, etc.) and presentation. A list of deliverables for the project and their due dates will be distributed in the sixth week of the semester.

Assignments and In-Class Exercises/Labs:

Assignments will be based on material covered in lectures and the text, articles, and cases. Assignments may include article critiques or discussions, case analysis, quizzes, and problem solving using techniques learned in class. All assignments, unless otherwise stated, **must** be submitted electronically in MS Word format on Blackboard (no emails) **AND** hand delivered hard copy before class starts. **Homework assignments will be due at the start of class.** Late assignments will be penalized. If you are away and cannot attend class you are still required to submit your assignment on time.

Please note the following assignment and in-class exercises rules:

1. Name your homework (hard copy and files) in the following way:

Assignment_Number_Your_FIRST_AND_LAST_NAME

2. Assignments are due **BEFORE** class - **YOU MUST BRING IN A HARD COPY**
3. Assignments must also be uploaded in the appropriate location on blackboard
4. No cover page is necessary for assignment hard copies
5. No comments are necessary for uploaded assignments
6. Late Assignments will be penalized **(-1 for each day they are late)**
7. For **group assignment**, each group member must upload a soft copy though **only 1 hard copy is necessary** – include each member name in the hard copy.

Note:

All examinations, papers, and other graded work products and assignments are to be completed in conformance with **The George Washington University Code of Academic Integrity**.

Tentative Schedule: The schedule of lectures and assignments below is subject to change without notice. Updated information will be available on our Blackboard site.

Date	Topic	Readings and Assignments Due
9/6	Course Overview Introduction to DW, Components <i>Lab 1: register at Teradata Site</i>	Chapter 1 (1.1, 1.2, 1.3), Ch. 2 (excluding Table 2-6) Module 1 Article: The Operational Data Store, Inmon
9/13	OLTP, OLAP, BI DW Design: Star and Snowflake schemas <i>Lab 2: OLAP Operations</i>	Course Expectations* Chapter 3, paragraph 2 only Module 2 Article: Who in the World Needs a DW? – Information Management
9/20	DW Design and Dimensional Modeling <i>Lab: In-class Exercises</i>	Module 3 Article: The Disconnect of Intelligence and Analytics - Information Management Assignment 1*: Star schema example. Assignment 2*: Convert star in snowflake schema. Lab 1*: Post OLAP Operations Results
9/27	Terminology and Architecture DW Design: Strategy and Methodology <i>Lab: star schema exercise</i>	Chapter 6.1, 6.2 Module 4 Module 5
10/4	NO CLASS	
10/11	DW Project Management DW Design: Dimensional Modeling <i>Lab: aggregation exercise</i>	Chapter 6.3, 6.4, 6.5 Make-up Assignment*: Lab1 with Tableau (3 extra points) Assignment 3*: Turn-in In Class Exercise (5 points) Module 6
10/18	DW Design: Slow Changing Dimensions Project Introduction <i>Lab: Create a report</i>	Module 7 Assignment 4* and 5 (Lab 2)*: Sample receipt
10/25	ETL <i>Lab: ETL exercise</i>	Assignment 6*: Slow Changing Dimension Exercise. Module 8 Lab 3*: post report
11/1	DW Agile Methodology Review for Mid-Term Exam <i>Lab: Project Work</i>	Chapter 7, Module 9 Articles: Agile BI: Best Practices for Breaking through the BI Backlog, Forrester; Foundation for BI, Information Week; Cloud BI Perfectly Stormy Future, TDWI; BI in the Cloud: Yes and on the Ground too, Forrester; Not All Mobile Applications are Created Equal, Forrester; BI and the Cloud, CIO Magazine.
11/8	Mid-Term Exam	
11/15	In class discussion on Agile* <i>Lab: Project Work</i>	Assignment 7* (Optional): Agile: 3 Most Relevant Concepts (1 paragraph each, 1 page maximum)
11/22	Data Mining and Data Visualization <i>Lab: Project Work</i>	Chapter 3 (excluding 3.2, 3.7)
11/25	Fuzzy Logic and Neural Networks <i>Lab: Project Work</i>	Assignment 8*: In class quiz on Data Mining Module 10 Module 11, Chapter 4
12/6	Project Presentation	Assignment 9*: In class quiz on Machine that can Learn

*The asterisk indicates assignment due. Assignments are subject to change; please see Blackboard for updated information. Additional assignments may be given in class or posted on Blackboard.