

LSC Use Only

Number: _____

Submission Date: _____

Action Date: _____

99-28d
UWUCC USE Only 98-411d age 25

Number: _____

Submission Date: _____

Action Date: UWUCC App 1/25/00
Senate App 2/29/00

CURRICULUM PROPOSAL COVER SHEET
University-Wide Undergraduate Curriculum Committee

I. CONTACT

Contact Person Elizabeth M. Pierce Phone 5713

Department MIS & Decision Sciences

II. PROPOSAL TYPE (Check All Appropriate Lines)

X **COURSE** Data Warehousing & Mining
Suggested 20 character title

X **New Course*** IM 455 Data Warehousing & Mining
Course Number and Full Title

 Course Revision _____
Course Number and Full Title

 Liberal Studies Approval+ _____
for new or existing course Course Number and Full Title

 Course Deletion _____
Course Number and Full Title

 Number and/or Title Change _____
Old Number and/or Full Old Title

_____ **New Number and/or Full New Title**

 Course or Catalog Description Change _____
Course Number and Full Title

PROGRAM: Major Minor Track

 New Program* _____
Program Name

 Program Revision* _____
Program Name

 Program Deletion* _____
Program Name

 Title Change _____
Old Program Name

_____ **New Program Name**

III. Approvals (signatures and date)

Laurie B. Busby
Department Curriculum Committee

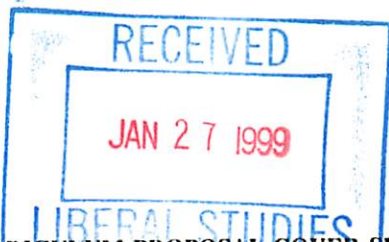
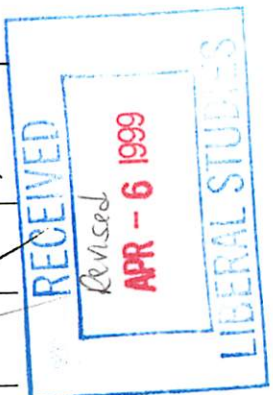
Wm. Sebala
College Curriculum Committee

Kenneth L. Sweet
Department Chair

Kate C. Carr
College Dean

NS
*Provost (where applicable)

+Director of Liberal Studies (where applicable)



Syllabus of Record Format: IM 455 Data Warehousing & Mining

I. Catalog Description

IM 455 Data Warehousing & Mining

3 credits
3 lecture hours
0 lab hours
3c-01-3sh

Prerequisite: IM 450 or IM 261 or CO441, and MA 214 or MA 216

This course introduces students to the strategies, technologies, and techniques associated with this growing MIS specialty area. This course is composed of two main parts. In the first part of the course, students will learn the basic methodology for planning, designing, building, using, and managing a data warehouse. In the second part of the course, students will learn how to use different data mining techniques to derive information from the data warehouse for strategic and long term business decision making.

II. Course Objectives

By the end of this course, students will be able to:

- Describe what is meant by a data architecture and be able to explain how legacy systems, data warehouses, data marts, and operational data stores fit into that data architecture.
- Understand the process by which data warehouses are designed, built and managed. Students also need to understand the many issues that complicate data warehouse construction and maintenance. For example, performance is a problem as data warehouses grow beyond gigabytes to terabyte, petabyte or even exabyte size. By the end of the course, students should be able to explain the steps they would follow to create a data warehouse as well as the different types of problems and issues that they will need to address.
- Discuss the variety of end user tools for analyzing the data contained in the data warehouse. These tools range from simple spreadsheets to sophisticated neural net software packages. Students should understand the selection process for matching the right set of tools to the various types of users in an organization.

III. Detailed Course Outline

- A. Introduction to Data Warehouses (3 hours)
 - 1. Review of Databases & Legacy Systems
 - 2. Evolution of Decision Support Systems
 - 3. What is a Data Warehouse?
- B. Planning the Data Warehouse (3 hours)
 - 1. The Data Warehouse Environment
 - 2. Developing a Data Warehouse Strategy

- C. Designing the Data Warehouse (8 hours)
 - 1. Data Warehousing Architectures
 - 2. Granularity in the Data Warehouse
 - 3. Requirements for a Data Warehouse
 - 4. Designing the Metadata Repository
- D. Building the Data Warehouse (4.5 hours)
 - 1. Transforming Data from Legacy Systems
 - 2. Technology (Hardware & Software Choices) for the Data Warehouse
 - 3. Data Quality Issues in the Data Warehouse
 - 4. Migration to the Architected Environment
 - 5. Connecting to the Data Warehouse
- E. Using the Data Warehouse (3 hours)
 - 1. The Manager's Perspective
 - 2. The End User's Perspective
 - 3. Creating the Analysis & Using Information for Competitive Advantage
- F. Managing the Data Warehouse (4.5 hours)
 - 1. Monitoring Data Warehouse Activity
 - 2. Monitoring Data Warehouse Data
 - 3. Managing Security in the Data Warehouse
 - 4. Administering the Data Warehouse Data Model
 - 5. Managing Metadata in the Data Warehouse
 - 6. Selecting End-User Access Tools
 - 7. Managing the Integration and Transformation Interface
 - 8. On-Line Application Processing and the Data Warehouse Environment
 - 9. Managed Redundancy in the Data Warehouse Environment
 - 10. Data Warehouse Team Roles and Responsibilities
- G. Variations on the Data Warehouse (1.5 hours)
 - 1. Introduction to Data Marts
 - 2. Introduction to Operational Data Stores
- H. Data Mining (10 hours)
 - 1. Introduction to Data Mining
 - 2. The Data Mining Process
 - 3. Statistical Analysis of Data
 - 4. Market Basket Analysis
 - 5. Memory Based Reasoning
 - 6. Cluster Detection
 - 7. Link Analysis
 - 8. Decision Trees and Rule Induction
 - 9. Artificial Neural Networks
 - 10. Genetic Algorithms
 - 11. On-Line Analytic Processing
 - 12. Visualization Techniques
- I. Case Studies (1.5 hours)
- J. Testing and Evaluation (3 hours)

IV. Evaluation Methods

50% Exams (1 mid-terms & final). Exams will emphasize concepts and techniques.

40% Lab exercises, Assignments and Projects using a Database and Data Mining Software

10% Term paper on a topic related to Data Warehousing and Data Mining.

Note: Scores will be recorded in points and letter grade equivalents will be computed at the end of the semester. Assignment of letter grades will be based on the percentage of correct points accumulated where: A = 90% and above, B = 80% - 89%, C = 70% - 79%, D = 60% - 69%, and F = Less than 60%.

V. Required Textbook(s)

Barquin, R. C. and Edelstein, H. A. *Building, Using, and Managing the Data Warehouse*, Prentice Hall, 1997.

Barquin, R. C. and Edelstein, H. A. *Planning and Designing the Data Warehouse*, Prentice Hall, 1997.

Groth, R. *Data Mining: A Hands-On Approach for Business Professionals*, Prentice Hall, 1998.

VI. Special Resource Requirements

Students can utilize existing PC labs in the Eberly College of Business for completion of assignments.

VII. Bibliography

Barquin, R. C. and Edelstein, H. A., (1997). *Building, Using, and Managing the Data Warehouse*, Prentice Hall.

Barquin, R. C. and Edelstein, H. A., (1997). *Planning and Designing the Data Warehouse*, Prentice Hall.

Berry, M. J. A. and Linoff, G., (1997). *Data Mining Techniques*, John Wiley & Sons.

Date, C. J., (1987). *An Introduction to Database Systems*, Addison-Wesley.

Groth, R., (1998). *Data Mining: A Hands-On Approach for Business Professionals*, Prentice Hall.

Inmon, W. H., (1996). *Building the Data Warehouse*, John Wiley & Sons.

Inmon, W. H., Imhoff, C., and Battas, G., (1996). *Building the Operational Data Store*, John Wiley & Sons.

Inmon, W. H. and Hackathorn, R. D., (1994). *Using the Data Warehouse*, John Wiley & Sons.

Inmon, W. H., Welch, J. D., and Glassey, K. L., (1997). *Managing the Data Warehouse*, John Wiley & Sons.

Levine, Berenson and Stephan, (1997). *Statistics for Managers using Microsoft Excel*, Prentice Hall.

Mattison, Rob (1996). *Data Warehousing: Strategies, Technologies and Techniques*, McGraw-Hill.

Redman, Thomas, (1992). *Data Quality: Management and Technology*, Bantam Books.

Course Analysis Questionnaire

Section A: Details of the Course

- A1 This course will be a controlled elective for MIS majors. Business majors who have had MA 214 and a course in databases may also elect to take this course. Acceptable database courses will include IM 450, IM 261 or equivalents like CO 441. Computer Science majors who have completed their math requirement of MA 216 and their database requirement will also be allowed to take this course.
- A2 This course does not require changes in the content of existing courses.
- A3 This course is currently being offered as special topic, IM 481/581, for Fall 1998.
- A4 This course is intended to be dual level.
- A5 This course is not intended to be taken for variable credit.
- A6 This course will give interested students an opportunity to study a fast growing area in the MIS field that combines both a knowledge of databases and statistical decision making techniques.
- A7 While this course is not explicitly required by accrediting bodies, such bodies emphasize the need to be current in the business technology area particularly newly emerging technologies destined to impact the business world. Students as well as the MIS - Eberly Business Advisory Council have expressed a need for a course of this type.

Section B: Interdisciplinary Implications

- B1 This course may be team-taught.
- B2 This course does not overlap with any other courses at the University.
- B3 Seats will be made available to Continuing Education students meeting the prerequisite.

Section C: Implementation

- C1 Faculty resources are adequate for teaching this course.
- C2 Other Resources
 - 1. Current space allocations are adequate for this course.
 - 2. The course will utilize existing ECOB PC labs.
 - 3. Laboratory supplies/consumable goods are not required.
 - 4. Library holdings are adequate.
 - 5. Travel funds are not required.
- C3 No resources are funded by grant sources.
- C4 This course will be offered in the Fall semester each year.

- C5 One section of this course will be offered each Fall depending on demand for the course.
- C6 Suggested maximum enrollment is 35.
- C7 No professional society recommends enrollment limits or parameters for a course of this nature.

Section D: Miscellaneous

No additional information.