

# College of San Mateo

## Course Outline

- New Course  
 Update/No change  
 Course Revision (Minor)  
 Course Revision (Major)

Date: November 10, 2008

Department: CIS

Number: 364

Course Title: Enterprise Data Warehousing Units: 4.0

Total Semester Hours: Lecture: 48 Lab: 48 Homework: 96 By Arrangement: 0

### Length of Course

- Semester-long  
 Short course (Number of weeks \_\_\_)  
 Open entry/Open exit

### Grading

- Letter  
 Pass/No Pass  
 Grade Option (letter or Pass/No Pass)

1. **Prerequisite** (Attach Enrollment Limitation Validation Form.)

CIS 362 or CIS 363 or equivalent

2. **Corequisite** (Attach Enrollment Limitation Validation Form.)

None

3. **Recommended Preparation** (Attach Enrollment Validation Form.)

ENGL 838/848

4. **Catalog Description** (Include prerequisites/corequisites/recommended preparation.)

CIS 364 Enterprise Data Warehousing (4) (Pass/No Pass or letter grade option.) 48 lecture hours and 48 lab hours per semester. Prerequisite: CIS 362 or CIS 363 or equivalent. Recommended Preparation: ENGL 838/848. Introduction to data warehousing architecture, data extraction, management, and load. Also covered are metadata management, logical and physical models, dimensional modeling, data aggregation, and project management. Hands-on design and development of a data warehouse using Oracle or MySQL. Advanced topics such as OLAP query processing, security management, and data mining techniques will be introduced. A \$2.00 materials fee is payable upon registration. (CSU)

5. **Class Schedule Description** (Include prerequisites/corequisites/recommended preparation.)

CIS 364 Enterprise Data Warehousing (4) Introduction to data warehousing architecture, data extraction, management, and load. Also covered are metadata management, logical and physical models, dimensional modeling, data aggregation, and project management. Hands-on design and development of a data warehouse using Oracle or MySQL. Advanced topics such as OLAP query processing, security management, and data mining techniques will be introduced. A \$2.00 materials fee is payable upon registration. Prerequisite: CIS 254 or equivalent. Recommended Preparation: ENGL 838/848. (Pass/No Pass or letter grade option.) (CSU)

6. **Student Learning Outcomes** (Identify 1-6 expected learner outcomes using active verbs.)

Upon successful completion of the course, the student will be able to:

1. Determine the best data warehouse architecture using proven analytical modeling concepts
2. Perform data extraction, transformation, and load
3. Model dimensions for the data warehouse
4. Manage metadata
5. Explain how and when to use aggregates
6. Design and develop a data warehouse using Oracle or MySQL
7. Query and manage the data warehouse

7. **Course Objectives** (Identify specific teaching objectives detailing course content and activities. *For some courses, the course objectives will be the same as the student learning outcomes. If this is the case, please simply indicate this in this section).*

See Student Learning Outcomes

8. **Course Content** (Brief but complete topical outline of the course that includes major subject areas [1-2 pages]. Should reflect all course objectives listed above. In addition, you may attach a sample course syllabus with a timeline.)

See attached topical outline

9. **Representative Instructional Methods** (Describe instructor-initiated teaching strategies that will assist students in meeting course objectives. Include examples of out-of-class assignments, required reading and writing assignments, and methods for teaching critical thinking skills.) **If hours by arrangement are required by this course, indicate the additional instructional activity which will be provided during this time.**

The course will include the following instructional methods as determined appropriate by the instructor:

Lecture will be used to introduce new topics;  
Teacher will model problem-solving techniques;  
Class will solve a problem together, each person contributing a potential "next step";  
Students will participate in short in-class projects (in teacher-organized small groups) to ensure that students experiment with the new topics in realistic problem settings;  
Teacher will invite questions AND ANSWERS from students, generating discussion about areas of misunderstanding;  
Teacher will create and manage an Internet conference for discussion of course topics; and  
Students will work in small groups to solve significant programming assignments.

10. **Representative Methods of Evaluation** (Describe measurement of student progress toward course objectives. Courses with required writing component and/or problem-solving emphasis must reflect critical thinking component. If skills class, then applied skills.)

Bi-weekly quizzes to provide feedback to students and teacher ;  
(Short answer--from textbook material)  
Assessment of student contributions during class discussion and project time;  
Individual database development assignments to assess objectives 4-5;  
Midterm and Final exams; and  
(Short answer (textbook material), general problem solving (similar to in-class work), short program segments (similar to database development assignments)

Assessment of group participation on course projects, including peer-assessment of participation and contribution to the group effort.

11. **Representative Text Materials** (With few exceptions, texts need to be current. Include publication dates.)

The Data Warehouse Lifecycle Toolkit, Second Edition, Kimball, Ross, et al, John Wiley  
ISBN 0470149779, 2008

Dimensional Data Warehousing with MySQL: A Tutorial, Darmawikarta, BrainySoftware  
ISBN 0975212826, 2007

Prepared by: \_\_\_\_\_  
(Signature)

Email address: greenm@smccd.edu

Submission Date: \_\_\_\_\_

## **CIS 364 – Enterprise Data Warehousing Topical Outline**

For Objective 1

1. Basic Elements of the Data Warehouse
  - a. Source System
  - b. Data Staging Area
  - c. Presentation Server
  - d. Dimensional Model
  - e. Business Process
  - f. Data Mart
  - g. Data Warehouse
  - h. Operational Data Store (ODS)
  - i. OLAP (On-Line Analytic Processing)
  - j. ROLAP (Relational OLAP)
  - k. MOLAP (Multidimensional OLAP)

*For Objectives 1 - 6*

2. Project Management and Requirements
  - a. The Business Dimensional Lifecycle
    - i. Lifecycle Evolution
    - ii. Lifecycle Approach
  - b. Project Planning and Management
    - i. Define and Plan the Project
    - ii. Collect the Requirements
  - c. Prepare and Publish the Requirements Deliverables

*For Objectives 3 - 6*

3. Data Design
  - a. Dimensional Modeling
  - b. The Data Warehouse Bus Architecture
  - c. Basic Dimensional Modeling Techniques
    - i. Fact Tables and Dimension Tables
    - ii. Foreign Keys, Primary Keys, and Surrogate Keys
    - iii. Additive, Semiadditive, and Nonadditive Facts
  - d. Extended Dimension Table Designs
    - i. Many-to-Many Dimensions
    - ii. Many-to-One-to-Many Traps
  - e. Extended Fact Table Designs
  - f. Build Dimensional Models

*For Objectives 1 - 6*

4. Data Warehouse Architecture
  - a. Architectural Framework
  - b. Logical Models and Physical Models
  - c. Back Room Technical Architecture
    - i. Back Room Data Stores
    - ii. Back Room Services
      - a. Extract Services
      - b. Data Transformation Services
      - c. Data Loading Services
    - iii. Backup and Archive Planning
  - d. Architecture for the Front Room
    - i. Front Room Data Stores
    - ii. Front Room Services for Data Access
      - a. Warehouse Browsing
      - b. Access and Security Services
      - c. Activity Monitoring Services
      - d. Query Management Services

*For Objective 4*

5. Infrastructure and Metadata
  - a. The Evolution of Infrastructure
    - i. Back Room Infrastructure Factors
    - ii. Front Room Infrastructure Factors
    - iii. Connectivity and Networking Factors
  - b. Metadata and the Metadata Catalog
    - i. Source System Metadata
    - ii. Data Staging Metadata
    - iii. DBMS Metadata
    - iv. Front Room Metadata

*For Objectives 5 - 7*

6. Implementation
  - a. Aggregation
    - i. Develop the Aggregate Table Plan
    - ii. Processing Aggregates
    - iii. Administering the Aggregates
  - b. Completion of the Physical Design
    - i. Develop Standards
    - ii. Develop the Physical Data Model

- iii. Develop the Initial Index Plan
- iv. Design and Build the Database Instance

7. Security Management in a Data Warehouse Environment

- a. Security: Vulnerabilities
  - i. Physical Assets
  - ii. Information Assets: Data, Financial Assets, and Reputation
  - iii. Software Assets
  - iv. Network Threats
- b. Security: Solutions
  - i. Routers and Firewalls
  - ii. The Directory Server
  - iii. Encryption