



FIRST SEMESTER 2022-2023
Course Handout

Date: 30-08-2022

In addition to part-I (General Handout for all courses appended to the time table) this portion gives further specific details regarding the course.

Course No. : SS G515
Course Title : **Data Warehousing**
Instructor-in-Charge : **Narasimha Bolloju**

Description: Introduction, evolution of data warehousing; decision support systems; goals, benefit, and challenges of data warehousing; architecture; data warehouse information flows; software and hardware requirements; approaches to data warehouse design; creating and maintaining a data warehouse; Online Analytical Processing (OLAP) and multi-dimensional data, multidimensional modeling; view materialization; data marts; data warehouse metadata; data mining.

Scope and Objectives of the Course:

Decision makers require access to all the organization's data, wherever it is located, in appropriate structure and detail that helps in their decision making. This course will involve an in-depth study of various concepts and techniques needed to design, develop, and maintain a data warehouse. The course also offers OLAP techniques and business intelligence tools for accessing data warehouses and data marts by the decision makers. This course has the following learning objectives:

- to understand concepts of data warehousing, role of data warehouses in provisioning strategic information to decision makers at various organizational levels,
- to gather and analyze requirements for building data warehouses by selecting a suitable data warehouse architecture and developing necessary dimensional models, and
- to apply suitable techniques for online analytical processing using business intelligence and tools.

Textbooks:

- T1. Ponniah Paulraj, *"Data Warehousing Fundamentals for IT Professionals"*, WSE, 2ed., 2010.
T2. Kimball Ralph & M Ross, *"The Data Warehouse Toolkit"*, WSE, 3rd ed., 2013.

Reference books

- R1. Anahory S, & Dennis M, *"Data Warehousing in the Real World"*, Addison-Wesley, 2000.
R2. Kimball R, Reeves L, Ross M, & Thornthwaite, W, *"The Data Warehouse Lifecycle Toolkit"*, John Wiley, 1998.
R3. Adamson C, & Venerable M, *"Data Warehouse Design Solutions"*, John Wiley, 1998.
R4. Inmon, WH, *"Building the Data Warehouse"*, John Wiley, 2002.



Course Plan:

Lecture No.	Learning Objectives	Topics to be covered	Chapters in the textbook
1-12	<ul style="list-style-type: none"> To explain the need for data warehouses in organizations To design a simple data warehouse using the dimensional modeling technique To develop reports and dashboards on a populated data warehouse 	<ul style="list-style-type: none"> Introduction to data warehousing Components of data warehouses Processes for building data warehouses Introduction to dimensional modeling Introduction to reports and dashboards for business intelligence (BI) 	T1: 1, 2 T2: 1, 2
13-18	<ul style="list-style-type: none"> To differentiate popular data warehouse architectures To specify business requirements for data warehouse To explain data warehouse development processes To apply dimensional modeling technique for data warehouse design 	<ul style="list-style-type: none"> Popular architectures for data warehouses and data marts Gathering and defining business requirements Kimball's DW/BI lifecycle Principles of dimensional modeling (star schemas and data cubes) 	T1: 4,6,7,10 T2: 2,18
19-24	<ul style="list-style-type: none"> To explain the role of metadata in data warehousing To elaborate necessary processes and tasks associated with ETL To explain mechanisms for ensuring data quality in data warehouses 	<ul style="list-style-type: none"> Metadata – need, common types and providing metadata Extract-transform-load (ETL) – concepts, operational source systems, requirements, processes and tasks, data integration Data quality challenges and techniques/tools for ensuring data quality 	T1:9, 12,13 T2: 19, 20
25-30	<ul style="list-style-type: none"> To explain online analytical processing To apply dimensional analysis techniques To elaborate differences among OLAP models To differentiate among information delivery mechanisms 	<ul style="list-style-type: none"> Online analytical processing (OLAP) – need, major functions and feature Dimensional analysis OLAP models (MOLAP, ROLAP, HOLAP) Types of information delivery mechanisms and supporting tools 	T1: 14, 15 T2: 17
31-40	<ul style="list-style-type: none"> To explain advanced dimensional modeling and ETL techniques To elaborate data warehouse performance enhancement methods To explain the trends in data warehousing and BI 	<ul style="list-style-type: none"> Advanced dimensional modeling Real time data warehouses Advanced ETL techniques Data warehousing trends and BI including Web-enabled data warehouses Data warehouse physical design and performance enhancement techniques 	T1: 3, 11, 16, 18, 20 T2: 1, 17, 19, 20

Practical	Experiment Title
-----------	------------------



No.	
1	Identifying strategic information needs for data warehousing
2	Designing and preparing analytical reports with a sample data mart
3	Developing enterprise bus matrix with relevant dimension tables for a data warehouse
4	Dimensional modeling with different dimension table types
5	Dimensional modeling with different fact table types
6	Designing and validating business intelligence reports
7	Selecting appropriate visualizations for the designed business intelligence reports
8	Connecting various elements of reports to corresponding elements in dimensional models
9	Populating data warehouse fact tables with data and validating the generating reports
10	Planning and verification of data staging for a given data warehouse
11	Developing dashboards for access via desktop and mobile devices
12	Practicing with advanced dimensional modeling techniques

Note: During lab sessions (#5 to #11 above), the students are required to present assigned data warehousing case studies (from Textbook T2) and to present & demonstrate their team project work.

Evaluation Scheme:

Component	Duration	Weightage(%)	Date & Time	Nature of component
Mid-semester test	90Minutes	20	02/11 1.30 - 3.00PM	Closed Book
Class activities		5+5	TBA	Open Book
Project work		10+15	TBA	Open Book
Lab activities		5+5	TBA	Open Book
Presentations		5	TBA	Open Book
Comprehensive Examination	180 Minutes	30	23/12 FN	Closed Book

Note: 40% of the evaluation will be completed by mid-sem grading (mid-sem 20%; class activities 5%; project work 10% and lab activities 5%).

Chamber Consultation Hour: Wednesday/Friday 10 AM to 11 AM (or by appointment)

Notices: All the notices and announcements concerning this course will be posted on **the course page on Google Classroom.**



Make-up Policy:For tests and comprehensive exam components, make-up policy is according to the institute regulations. For the other component (fourth in the table above), there will be no make-up.

Academic Honesty and Integrity Policy: Academic honesty and integrity are to be maintained by all the students throughout the semester and no type of academic dishonesty is acceptable.

INSTRUCTOR-IN-CHARGE

