BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI
INSTRUCTION DIVISION
SECOND SEMESTER 2015-2016
COURSE HANDOUT (PART II)

Date: 13/01/2016

In addition to Part-I (general handout for all courses appended to this time table) this portion gives further details pertaining to the course.

Course No.: **SS G515**
Course Title: **Data Warehousing**
Instructor-in-charge: **NAVNEET GOYAL (goel@bits-pilani.ac.in)**

1) Objective and Scope

Corporate decision makers require access to all the organization's data, wherever it is located. To provide comprehensive analysis of the organization, its business, its requirements and any trends, require access to not only the current data in the database but also to historical data. To facilitate this type of analysis, data warehouses have been created to contain data drawn from several sources, maintained by different departments of the organization. This course will involve an in-depth study of various concepts needed to design, develop, and maintain a data warehouse. It also provides an introduction to end user access tools like OLAP and reporting.

2) Text Books:

T1. Ponniah P, ***"Data Warehousing Fundamentals"***, 2e, John Wiley, 2012.
T2. Kimball R, ***"The Data Warehouse Toolkit",*** 3e, John Wiley, 2013

3) Reference Books

R1. Anahory S, & Dennis M, "***Data Warehousing in the Real World"***, Addison-Wesley, 2000.
R2. Kimball R, et. al. ***"The Data Warehouse Lifecycle Toolkit",*** 2e, John Wiley, 2008.
R3. Kimball R, & Caserta J, The Data Warehouse ETL Toolkit, John Wiley, 2004.
R4. Inmon, WH, ***"Building the Data Warehouse"***, 4e,John Wiley, 2005.

4) Course Plan

Lecture No.	Learning Objective	Topic(s)	Chapter Reference
1-2	To understand the need, definition, & applications of a Data Warehouse	Introduction to Data Warehousing • Present Business Scenario • Operational and Informational Systems • What is a Data Warehouse? • Applications of Data Warehouse • Problems with Data Warehousing	T1: 1
3-4	To understand the components, & processes of a Data Warehouse	Data Warehouse Components, & Processes • Source Systems • Data Staging Area • Presentation Server • Data Marts • Operational Data Store (ODS) • Metadata • Information Delivery • Basic Processes of a Data Warehouse	T1: 2 R2: 2
5-6	To understand the Data Warehouse Architecture	Data Warehouse Architecture	T1: 7 R1: 3 & 4
7-8	To learn how to collect business requirements for a Data Warehouse	Collecting Business Requirements	T1: 5 R2: 4 Self Study
9-10	To learn dimensional modeling for designing database schemas for a Data Warehouse	Data Warehouse Data Design • Dimensional Modeling Basics • Facts, Dimensions, & Star Schemas • Snowflake & Starflake Schemas • Design Steps • ER modeling vs. Dimensional modeling	T1: 10 T2: 1 R1: 5 R2: 5, 6, & 7
11-12	To understand the role of Data Marts & ODS in Data Warehousing	Data Marts & ODS • Architecture • Design • Cost	T1: 19 R1: 8 R4: 5
13-15	To understand advanced Dimensional Modeling concepts	Advanced Dimensional Modeling Concepts • Surrogate Keys • Changing Dimensions • Conformed Dimensions	T1: 11 + Class Notes

		• Factless Fact Tables • Minidimensions & Outriggers • Role-playing Dimensions • Multi-valued Dimensions	
16-17	To understand the ETL Process	Extraction, Transformation, & Loading (ETL) • Data Extraction • Data Transformation • Data Loading • ETL Data Structures • ETL Tools: Build or Buy?	T1: 12 R3
18-20	To understand OLAP, its features, functions, & variations	Online Analytical Processing (OLAP) • Need for OLAP • Features & Functions • ROLAP, MOLAP, HOLAP, & DOLAP • OLAP Implementation • OLAP Tools	T1: 15
21-22	To understand role of Multidimensional Databases in Data Warehousing	Multidimensional Databases (MDDBs)	Class Notes
23-24	To understand the new data warehousing related features of SQL	SQL Features for DW • CUBE Operator • Roll-up Operator • Top-N Queries • Window Queries	Class Notes
25-27	To understand efficient cube computation techniques	Cube Computation • Complexity • Optimization Techniques (ROLAP & MOLAP)	Class Notes
28	Case Study	Financial Services-Banks	T2: 9
29-32	To understand and implement various techniques used to reduce the query response time	Performance Enhancing Techniques • Partitioning • Aggregation • Materialization of Views • Bitmap Indexes	T1: 11, 18 T2: 16 R1: 6 & 7 R2: 14 + Class Notes
33-34	Case Study	Academic Data Warehouse: BITS Pilani	T2: 12
35	To understand the role of Metadata	Metadata • Role • Design	T1: 9 R1: 9 R2: 11
36-37	To understand the need for Real Time Data Warehousing	Real-Time Data Warehousing	Class Notes
38-39	To understand the role of HPC in Data Warehousing	HPC Solutions for Data Warehousing • MapReduce/Hadoop • HIVE	Class Notes
40	To expose students to the research trends in Data Warehousing	Data Warehousing Research Trends • NoSQL Databases • Data Lakes	Class Notes

5). Evaluation Schedule

Component	Duration	Weightage(%)	Date & Time	Remarks
Midsem Test	90 Mins.	30	15/3 11:00 - 12:30 PM	Closed Book
Assignments	Take Home	15+15	To be announced	
Comprehensive	3 Hours	40	5/5 AN	Partly open

6). Assignments & Labs.

A series of study, design, and implementation assignments will be given to the students on a regular basis. These assignments will immensely help the students in gaining a better understanding of the subject. Students will also get hands on experience on popular DW software during the weekly 2-hour practical sessions.

7). Chamber-Consultation Hours: T, Th - 10 (in addition to this, students can seek appointment through email)

8). Notices: All the notices concerning this course will be displayed on **NALANDA o**nly.

Instructor-in-charge
SS G515