

# FIRST SEMESTER 2019-2020

# COURSE HANDOUT (PART II)

**Date: 24/07/2019**

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*: **Narasimha Bolloju (narsi.bolloju@hyderabad.bits-pilani.ac.in)**

# Scope and Objectives

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.

# Text Books:

**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** | **References** |
| 1-12 | * 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 | * 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 | * 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 | * 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 | * 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 | * 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 | * To explain online analytical processing * To apply dimensional analysis techniques * To elaborate differences among OLAP models * To differentiate among information delivery mechanisms | * 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 | * To explain advanced dimensional modeling and ETL techniques * To elaborate data warehouse performance enhancement methods * To explain the trends in data warehousing and BI | * 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 |

# Evaluation Schedule

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Component** | **Duration** | **Weightage(%)** | **Date & Time** | **Remarks** |
| Mid Sem Test | 90 Mins | 20 | 03/10 09:00 – 10:30 am | Closed Book |
| Case study presentation + discussion | NA | 10 | Details will be provided in the first week | Open Book/ During the lab sessions |
| Project work  + Lab Sessions | NA | 30 | Details will be provided in the first week | Open Book/Take  Home |
| Comprehensive  Examination | 3 Hours | 40 | 09/12 FN | Closed Book |

* + - 1. **Chamber-Consultation Hours** T.B.A. in the class.

# Notices

All the notices concerning this course will be posted on the course page on Google Classroom.

**8) 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

**SS G515**