NATIONAL UNIVERSITY
of Computer & Emerging Sciences, Lahore

F A S T   S c h o o l   o f   C o m p u t i n g

DS3003 – Data Warehousing & Business Intelligence

FALL 2024

Instructor Name: Muhammad Ishaq Raza   **TA Name:** TBA
Email address: ishaq.raza@nu.edu.pk   **Email address:**
Office Location/Number: NB
Office Hours: Mon, Wed 1:30 - 3:00 PM

Course Information

Program: DS   **Credit Hours: 3**   **Type:** Core
Pre-requisites (if any): DS2001 – Introduction to Data Science
Course Website (if any):

Course Description/Objectives/Goals

This course covers the concepts and techniques in the design and construction of high-performance data warehouses. The software, hardware and design factors influencing performance characteristics of the data warehouse will be emphasized. A special focus will be given to features and functions in RDBMS implementations that are appropriate in a data warehouse environment. Distinction between DSS (Decision Support System) and OLTP workloads will be made with an emphasis on performance characteristics and functionality required.

Course Learning Outcomes (CLOs):		
At the end of the course students will be able to:	**Domain**	**BT* Level**
Introduction to the concepts and techniques in data warehousing and business intelligence.	C	2
Design of high performing data warehouses.	C	3
Construction of high performing data warehouses.	C	3
* BT= Bloom's Taxonomy, C=Cognitive domain, P=Psychomotor domain, A= Affective domain. **Bloom's taxonomy Levels:** 1. Knowledge, 2. Comprehension, 3. Application, 4. Analysis, 5. Synthesis, 6. Evaluation		

Course Textbook

1. Paulraj Ponniah, _Data Warehousing Fundamentals_, John Wiley & Sons, 2010.
2. Handouts

Additional references and books related to the course

1. Ralph Kimball, The Data Warehouse Lifecycle Toolkit: Expert Methods for Designing, Developing and Deploying Data Warehouses, John Wiley & Sons, 1998.
2. Ralph Kimball, _The Data Warehouse Toolkit._ John Wiley & Sons, June 1996.
3. W. H. Inmon, _Building the Data Warehouse_ (3rd Edition), John Wiley & Sons, 2002.
4. Articles

Tentative Weekly Schedule

Week	Topics to be covered	Readings (Textbook)	No of Lectures	Asst.
1-2	1. **Overview and Concepts:** DW Fundamentals, need for a DW and BI, decision support versus transaction processing, evolution of a DW	Ch. 1,2,3 Handout	3	A1
2-3	2. **Logical and Physical Data Modeling:** Normalization vs. denormalization, pre-join denormalization, column replication/ movement, pre-aggregation denormalization	Handout	2	
3-4	3. **OLAP Implementation Techniques:** OLAP framework for decision support, Physical implementation techniques: MOLAP, ROLAP, HOLAP, and DOLAP, Star schema design	Ch. 15 Handout	2	
4-5	4. **Dimensional Modeling:** Principles of dimensional modeling, Physical database design for ROLAP deployment, Natural versus surrogate key design	Ch. 10,11 Handout	3	A2
6	5. **Extract, Transform, Load (ETL) Processing**	Ch. 12 Handout	2	
7	6. **Join Techniques and Performance Evaluation for Data Warehousing:** DSS vs. OLTP queries, nested loop join, sort merge join, merge join, hash join, pointer-based join, query optimization	Handout	2	A3
8-9	7. **Indexing Techniques for Data Warehousing:** Traditional B-tree indexing, hash indexing, primary vs. secondary indexing, single index access vs. scanning, combining multiple indexes, dynamic bitmap indexing, static bitmap indexing, composite indexing, covered indexing, cluster indexing, partial indexing	Handout	4	A4
10-11	8. **Advanced Physical Database Design:** Horizontal and vertical partitioning, materialized views framework, materialized views for geography manipulation, advanced aggregation functions	Ch. 18 Handout	3	
11-12	9. **Data Mining and Data Visualization**	Handout	3	
13-14	10. **Advanced Data Warehousing & BI Concepts**	Handout	4	

(Tentative) Grading Criteria

1. Assignments (10%)
2. Quizzes/ Class Participation (10%)
3. 2 Midterm Exams (30%)
4. Final Exam (50%)

Grading Scheme: Absolute

Course Policies

1. Quizzes may be un-announced.
2. No makeup for missed quizzes or assignments.
3. Minimum eligibility to pass this course is to get 50% marks.