Course Number	EBB 6103
Course Title	**Data Engineering**
Course Outline	**Module 1: Introduction to Data Engineering** ● Overview of Data Engineering ● Key Concepts in Data Modeling ● Relational Database Design Principles ● Data Warehousing Concepts **Module 2: Data Processing and Storage** ● Data Pipelines and ETL (Extract, Transform, Load) ● Distributed Systems and Parallel Computing ● Data Storage Technologies, including NoSQL databases ● Data Quality and Validation **Module 3: Managing and Optimizing Data Systems** ● Performance Tuning and Optimization ● Data Security and Privacy ● Scalability and Availability ● Disaster Recovery and Backup **Module 4: Data Engineering Tools and Technologies** ● SQL and Relational Database Management Systems ● Big Data Frameworks, including Hadoop and Spark ● Cloud-Based Data Warehousing, including Amazon Redshift and Google BigQuery ● Data Visualization and Reporting Tools
Learning Outcome	● Demonstrate an understanding of data engineering concepts and principles. ● Design and implement efficient and scalable data pipelines for data processing and storage. ● Manage and optimize data systems for performance and reliability. ● Apply data engineering tools and technologies to real-world data problems.
Assessment Method	Quiz / Assignment / ESE

TEXTBOOKS:

● Designing Data-Intensive Applications by Martin Kleppmann (O'Reilly Media)
● Data Warehousing in the Age of Big Data by Krish Krishnan (Morgan Kaufmann)
● The Data Warehouse Toolkit by Ralph Kimball and Margy Ross (Wiley)