Arjun Mehta

K036

B. Tech CSE(Cybersecurity)

DWM LAB 01

	Data Base	Data Warehouse	Data Analytics	Data Science	Machine Learning	Artificial Intelligence
Definition	Organized collection of structured data	Central repository for integrated data from multiple sources	Techniques to analyse data and extract insights	Field that uses methods, processes, and systems to extract knowledge from data	Field of study that gives computers the ability to learn without being explicitly programmed	Simulation of human intelligence processes by machines, especially computer systems
Purpose	Supports day-to-day operations	Supports business intelligence and analytics	Provides actionable insights from data	Extracts knowledge and insights from structured and unstructured data	Enables computers to learn from data and improve their performance over time	Enables machines to perform tasks that would normally require human intelligence
Data Type	Structured data (e.g., tables, rows, columns)	Structured and sometimes semi-structured data	Structured, semi-structured, and unstructured data	Structured, semi-structured, and unstructured data	Structured and unstructured data	Structured and unstructured data
Users	Database administrators, application developers	Data analysts, business analysts, data scientists	Data analysts, data scientists, business analysts	Data scientists, statisticians, data analysts	Data scientists, machine learning engineers, developers	Data scientists, AI engineers, developers
Scale	Typically, smaller scale	Large scale, handles	Can handle	Can handle	Can handle	Can handle

		vast amounts of data	large datasets	large datasets	large datasets	large datasets
Querying	SQL for querying	SQL for querying, optimized for read-heavy operations	SQL, NoSQL, data visualization tools	SQL, NoSQL, programming languages like Python and R	Programming languages like Python, R, frameworks like TensorFlow, PyTorch	Programming languages like Python, R, frameworks like TensorFlow, PyTorch, OpenAI
Examples	MySQL, PostgreSQL, Oracle Database	Amazon Redshift, Google BigQuery, Snowflake	Tableau, Power BI, Google Analytics	Jupyter, RStudio, Apache Spark	Scikit-learn, TensorFlow, Keras	IBM Watson, OpenAI GPT-3, Microsoft Azure AI

A notable case study of data warehousing and mining is **Amazon**, which effectively utilizes these technologies to enhance its e-commerce and cloud computing services.

E-commerce Optimization: Amazon collects and analyses vast amounts of customer data, including browsing history and purchase behaviour, through its data warehouse. This information powers its recommendation engine, significantly improving product suggestions and increasing sales through personalized shopping experiences.

Cloud Services: Amazon Web Services (AWS) employs data warehousing solutions like Amazon Redshift to manage and analyse performance data generated from its cloud infrastructure. This enables real-time analytics, helping businesses optimize their operations based on current usage patterns.

Impact on Decision-Making: The integration of real-time data allows Amazon to make informed decisions quickly, adapting strategies to meet customer demands and market trends effectively.

This case exemplifies how data warehousing and mining can drive growth, enhance customer satisfaction, and streamline operations in today's competitive landscape.