Assignment as on Date – 9/3/2024

**Download SSMS –** Downloaded SQL Server From <https://www.Microsoft.com/en-us/sql-server-downloads>.

* Click on install server 2022 for developer and download it
* We will get the developer edition page with 3 options which has 1.Basic 2.Custom 3.Download media - Select basic version
* Next comes with developer edition – Click on installing SSMS.
* SSNS – Setup-ENU is downloaded on ur system along with SQL2022-SSEI-DEV which will successfully install sql server on ur pc
* Set up ur server details and it is all ready to use.

**Structured Query Language (SQL):** It is a programming language. It interacts with RDBMS It is widely used by data analytics, software engineers etc.,

Most common used are- ORACLE, MYSQL, MICROSOFT SQL SERVER, POSTGRE SQL.

**what is database** - it is like a container filled with information in an electronic method of storing data. Which can be

> easily accessed

> modified

> Protected

> Analysed

**what is data warehouse –** A SQL data warehouse is a centralized repository that integrates data from various sources for analytical querying and reporting. It uses structured schemas, like star or snowflake, and is optimized for complex queries and historical data analysis, supporting business intelligence and decision-making.

**what is DataMart - A** data mart is a specialized subset of a data warehouse which is used for a specific business unit or department, such as sales or finance, Accounting. It contains a focused segment of data relevant to that particular area, designed to support departmental reporting and analysis. Data marts simplify data access for users by filtering out unnecessary information, enhancing query performance, and improving the relevance of insights. They often draw from the central data warehouse but are optimized for targeted, departmental needs.

**what is data lake - A** data lake is a large, centralized repository that stores raw, unstructured, and structured data in its native format. Unlike a traditional data warehouse, a data lake accommodates diverse data types, including text, images, and logs, without pre-processing. It leverages scalable storage solutions to handle vast volumes of data, making it suitable for big data analytics and machine learning. In SQL contexts, data lakes can integrate with SQL engines for querying and analysis, though they are typically more flexible in handling varied data formats and structures.

**What are keys in sql -**

1. Primary Key: Uniquely identifies each record in a table. It must contain unique values and cannot be NULL. Each table can have only one primary key, which may consist of a single column or multiple columns (composite key).
2. Foreign Key: A field (or combination of fields) in one table that refers to the primary key of another table. It establishes and enforces a link between the data in the two tables, ensuring referential integrity.
3. Unique Key: Ensures that all values in a column (or a combination of columns) are unique across the table, except for NULL values. Unlike the primary key, a table can have multiple unique keys.
4. Candidate Key: A column or set of columns that can uniquely identify a record within a table. Each table can have multiple candidate keys, but one is selected as the primary key.

Ex : In A employee table we can use Emp ID and Emp Email address as candidate key

1. Alternate Key: Any of the candidate key which is not a part of primary key.so when we talk about alternative key the column may not be a primary key but still it is a unique key in the column.

EX: In student table Were ROLL NO is a primary key and rest of All the columns like Name, Subjects etc., are alternate key

1. Composite Key: A primary or unique key that consists of two or more columns used together to uniquely identify a record in a table.

EX: CREATE TABLE Enrollment (

StudentID INT,

CourseID VARCHAR(10),

EnrollmentDate DATE,

PRIMARY KEY (**StudentID, CourseID**)

**what are joins in sql -** In SQL, joins combine rows from two or more tables based on a related column. Key types include:

* INNER JOIN: Returns matching rows from both tables.
* LEFT JOIN: Returns all rows from the left table, with matched rows from the right table or NULL.
* RIGHT JOIN: Returns all rows from the right table, with matched rows from the left table or NULL.
* FULL JOIN: Returns all rows from both tables, with NULLs where no match exists.

Week 2 – 9/12/2024

1. What is SQL? what is No-SQL?

**SQL (Structured Query Language)** is used for managing and querying relational databases, where data is stored in structured tables with predefined schemas. It supports complex queries and transactions, ensuring data integrity and consistency.

**No-SQL (Not Only SQL)** databases handle unstructured or semi-structured data and offer flexible schemas. They are designed for scalability and performance across distributed systems, supporting various models like document, key-value, wide-column, and graph databases. No-SQL is ideal for dynamic, large-scale applications.

2. what is database? what is relational and non relational database ?

A **database** is an organized collection of data that is stored and managed electronically, enabling efficient access, modification, and management.

**Relational databases** use structured tables with rows and columns, adhering to a fixed schema. They support complex queries and transactions using SQL, ensuring data integrity and consistency.

**Non-relational databases** (No-SQL) handle unstructured or semi-structured data and offer flexible schemas. They include various models such as document, key-value, wide-column, and graph, and are designed for high scalability and performance in handling dynamic and large-scale datasets. Non-relational databases are often used for applications requiring rapid data retrieval and flexible data structures.

3. Types of SQL Statements ?

1. **Data Definition Language (DDL):** Defines and manages database structures. Examples include:
   * CREATE: Defines new tables, views, or databases.
   * ALTER: Modifies existing database objects.
   * DROP: Deletes tables, views, or databases.
2. **Data Manipulation Language (DML):** Manages data within tables. Examples include:
   * SELECT: Retrieves data.
   * INSERT: Adds new records.
   * UPDATE: Modifies existing records.
   * DELETE: Removes records.
3. **Data Control Language (DCL):** Controls access to data. Examples include:
   * GRANT: Provides user permissions.
   * REVOKE: Removes user permissions.
4. **Transaction Control Language (TCL):** Manages transactions. Examples include:
   * COMMIT: Saves changes.
   * ROLLBACK: Reverts changes.
   * SAVEPOINT: Sets a transaction savepoint.
5. What are Database Integrity Constraints :

These constraints are rules that ensure the accuracy and consistency of data within a database. They enforce data validity by restricting the types of data that can be entered and maintained. Common constraints include:

**Primary Key:** Ensures unique identification for each record.

**Foreign Key:** Maintains referential integrity between tables.

**Unique:** Guarantees that all values in a column are distinct.

**Not Null:** Requires that a column must contain a value.

**Check:** Enforces specific conditions on data values.

1. What is OLTP and OLAP?

**OLTP (Online Transaction Processing)** systems are designed for managing and processing real-time transactional data, such as order processing and inventory management. They focus on quick query responses and data integrity, supporting day-to-day operations.

**OLAP (Online Analytical Processing)** systems are used for complex data analysis and reporting, often involving large volumes of historical data. They support multidimensional queries and are optimized for read-heavy operations, enabling business intelligence and data warehousing tasks like trend analysis and decision-making.

1. What is Three-tier architecture?

Three-tier architecture is a software design model that separates an application into three distinct layers:

**Presentation Layer:** The user interface, which interacts with users and displays data.

**Application Layer (Business Logic Layer):** Handles the core functionality, processing user inputs, and applying business rules.

**Data Layer:** Manages data storage and retrieval, interacting with databases or other data sources.

This separation promotes modularity, scalability, and maintainability by isolating different concerns within the application.

1. .What is Database Design ?

Database design is the process of defining the structure, relationships, and constraints of a database to ensure efficient data storage, retrieval, and management. It involves creating a schema that includes tables, columns, data types, and keys, and establishing relationships between tables through foreign keys.

Proper database design aims to optimize performance, maintain data integrity, and support scalability, ensuring that the database meets the needs of the application and users effectively.

1. What is Normalization?

Normalization is a database design process that organizes data to reduce redundancy and improve data integrity. It involves dividing a database into two or more related tables and defining relationships between them to minimize duplicate data and ensure that each piece of information is stored in only one place. The process follows a series of normal forms (e.g., 1NF, 2NF, 3NF) that progressively eliminate data anomalies and dependencies, enhancing the efficiency and consistency of the database.

**SQL Queries to Practice: (Attached in SQL file)**

1. Create a Database?

2. Create a Customer table in it?

3. Perform DDL/DML/DQL commands on Customer table?

4. Create two different tables example: Customer and Order tables ? Insert data into it.

5. Perform all Join Operations?