Coursera RDBMS

(By Vanshika Malviya)

Following is the summary of what was taught in the course modules followed by the assignment summary:

Module 1

A well-designed relational database ensures that the users and applications that depend on the data will know that it is:

- Accurate. Can you rely on the accuracy of the data as new information is added or it is modified?
- Easy to access. Is the data organized to make it fast, easy, and predictable to query and maintain?
- Reliable. Can your database design ensure data integrity and maintain consistent and reliable data?
- Flexible. Can you update or expand on the design to meet future data requirements?

Data: Unprocessed information that goes under processing. It includes facts, observations, perceptions, numbers, characters, symbols, images, or a combination of these elements.
Data structure: Structure, unstructured and semi- structured
Types of models: Information model and
Data model: A data model operates at a tangible level, serving as the blueprint for translating conceptual information models into practical database structures.
Difference between data and information model
In relationships, an information model conceptualizes without specifying storage, while the hierarchical model physically structures data in a tree-like format.
Types of data model: The most common types of data models include the relational model, and entity relationship data model.
Concepts in dbms:

- Logical data independency
- Physical data independency
- Physical storage independence.

Fundamental concepts of ERD
Best examples:

- Primary key designation
- Data validation
- Default values
- Use of views
- Concurrency control

Data types: These include date, time, float, or decimal.
Advantages of data types(like varchar and char)
Properties of a relation:

- Reflexitivity
- Symmetry
- Transitivity
- Asymmetry

Deployment topology: Arrangement or configuration of hardware, software and network components.
Common deployment topologies:

- Single-tier architecture

- Client server or two-tier
- Three tier
- Cloud based

DBMS layers:

- Data access layer
- Database engineer layer
- Database storage layer

Types of database architecture:

- shared disk architecture
- shared nothing architecture
- combination and specialized architectures

Key users of databases:

- DBA
- Datascientists and business analysts(DS and BI tools)
- Application developers

RDB timeline

Open source relational databases:

- Oracle
- MySQL
- Microsoft SQL Server
- PostgreSQL
- MangoDB
- Redis
- Elasticsearch
- IBM Db2
- SQLite
- Microsoft access

Db2 is a versatile family of products deployable across various platforms, providing high availability, disaster recovery, and scalability.

MySQL supports various programming languages and is a reliable, scalable, and widely adopted database system compatible with UNIX, Windows, and Linux.

PostgreSQL is an open-source, object-relational database supporting a range of languages for client application development, including features for handling relational, structured, and non-structured data.

Advanced relational model concepts:

- Functional dependency
- Multi-valued dependency
- Candidate keys

Module-2

Common DDL terms:

- Create
- Alter
- Truncate
- Drop

Common DML terms

- Insert
- Select
- Update
- Delete

Loading data is quicker, more efficient, and more scalable than using multiple INSERT statements.

How to load data into a table from a CSV file in DB2?

Four steps of loading data:

- Source
- Target
- Define
- Finalise

Database hierarchy:

- Instance(Not all rdbms uses this)
- Databases
- Schemas(internal and external)
- Database partition
- Database objects(Tables, constraints, indexes, views, aliases)

Primary key and foreign key
Index allows you to find specific rows based on criteria
Examples: Online shopping, ATM, airline reservation system, etc.
Advantages and disadvantages of indexes
Normalisation: 1NF, 2NF, 3NF and BCNF
Normalisation in OLTP and OLAP
Constraints, 6 types of constraints:

- Entity integrity constraint
- Referential integrity constraint
- Semantic integrity constraint
- Domain constraint
- Null constraint
- Check constraint

Module-3

MySQL
Practical applications
Popular SQL tools: MySQL CLI, MySQL Workbench, phpMyAdmin
Creating databases and tables: using the command line, graphical user interface and API
Command line backup and restore
Import function to load large amount of data
Types of keys and constraints: Primary key, foreign key, unique constraint and null constraint
PostgreSQL, how to use, deployment, etc.
Creating database, tables, import function, backup in psql
View, materialised view

Module-4

Importance of database design
Database design process:

- Requirements analysis
- Logical analysis
- Physical design

Entity relationship management
Physical design
ERD designer

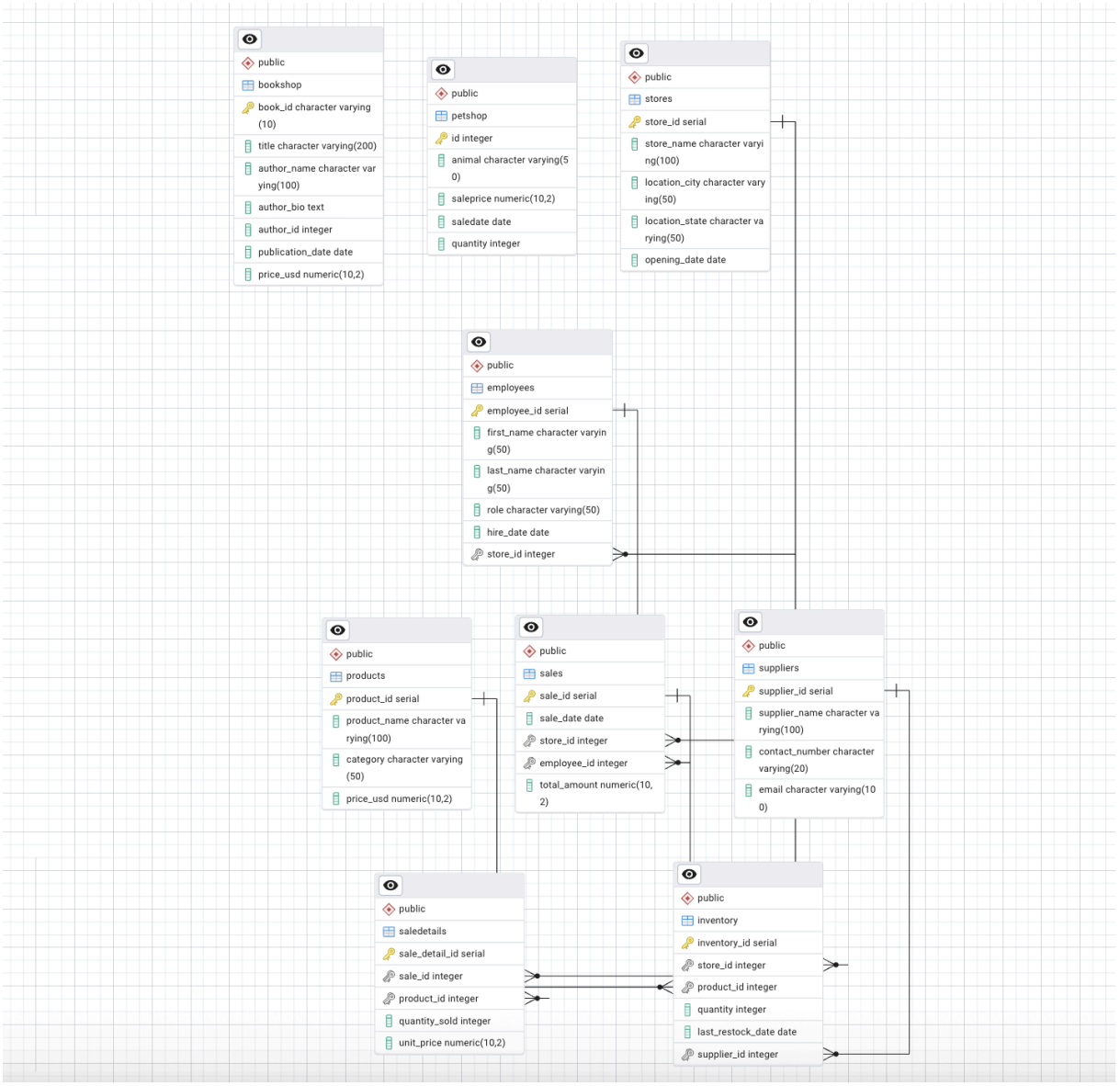
Assignment part:

Sample Coffee Shop Chain Database Design

Entities & Attributes

1. **Stores**
    - store_id (PK)
    - store_name
    - location_city
    - location_state
    - opening_date
2. **Employees**
    - employee_id (PK)
    - first_name
    - last_name
    - role
    - hire_date
    - store_id (FK → Stores.store_id)
3. **Products**
    - product_id (PK)
    - product_name
    - category
    - price_usd
4. **Suppliers**
    - supplier_id (PK)
    - supplier_name
    - contact_number
    - email
5. **Inventory**
    - inventory_id (PK)
    - store_id (FK → Stores.store_id)
    - product_id (FK → Products.product_id)
    - quantity
    - last_restock_date
    - supplier_id (FK → Suppliers.supplier_id)
6. **Sales**
    - sale_id (PK)
    - sale_date
    - store_id (FK → Stores.store_id)
    - employee_id (FK → Employees.employee_id)
    - total_amount
7. **SaleDetails**
    - sale_detail_id (PK)
    - sale_id (FK → Sales.sale_id)
    - product_id (FK → Products.product_id)
    - quantity_sold
    - unit_price

ER diagram:


public
bookshop
book_id character varying (10)
title character varying(200)
author_name character varying(100)
author_bio text
author_id integer
publication_date date
price_usd numeric(10,2)
public
petshop
id integer
animal character varying(50)
saleprice numeric(10,2)
saledate date
quantity integer
public
stores
store_id serial
store_name character varying(100)
location_city character varying(50)
location_state character varying(50)
opening_date date
public
employees
employee_id serial
first_name character varying(50)
last_name character varying(50)
role character varying(50)
hire_date date
store_id integer
public
products
product_id serial
product_name character varying(100)
category character varying (50)
price_usd numeric(10,2)
public
sales
sale_id serial
sale_date date
store_id integer
employee_id integer
total_amount numeric(10, 2)
public
suppliers
supplier_id serial
supplier_name character varying(100)
contact_number character varying(20)
email character varying(100)
public
saledetails
sale_detail_id serial
sale_id integer
product_id integer
quantity_sold integer
unit_price numeric(10,2)
public
inventory
inventory_id serial
store_id integer
product_id integer
quantity integer
last_restock_date date
supplier_id integer