# Project Title: "E-Commerce Database Design and Analysis for a Grocery Store"

## Project Description:

Your task is to design and implement a database for an imaginary grocery store. The database should capture and organize various aspects of the store's operations, such as product inventory, customer information, orders, and sales data. Additionally, you will analyze the data and provide insights to support business decision-making.

## Goals

Provide the business insights of the E-Commerce performance from 2021–2022.

## Result

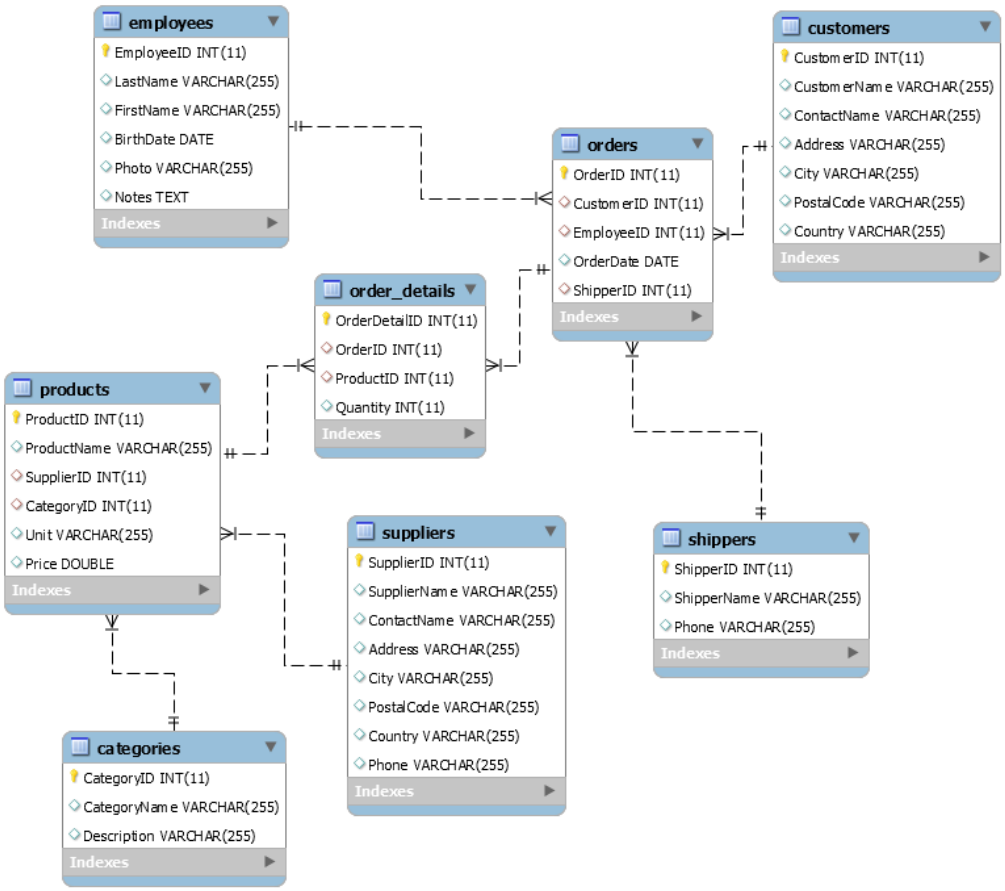
The insight into the E-Commerce performance is shown in the dashboard report.

## Project Steps:

1. **Database Design:**
    - Identify the entities and attributes relevant to the fashion store (e.g., products, customers, orders, shippers, etc.).
    - Design an appropriate database schema using a relational model, considering relationships, primary and foreign keys, and normalization principles.
2. **Data Collection and Loading:**
    - Generate or obtain sample data for the various entities in the database schema.
    - Load the data into the database using a suitable database management system (e.g., MySQL, PostgreSQL).
3. **Data Analysis:**
    - Explore the data to gain insights into the store's operations. Consider performing queries, aggregations, and statistical analysis to uncover trends, patterns, and key metrics.
    - Analyze sales data to identify popular products, customer purchasing behaviors, seasonal trends, etc.
    - Utilize SQL or other appropriate programming languages/tools for data analysis.
4. **Business Insights:**

- Based on the data analysis, extract meaningful insights and create visualizations to present the findings.
- Identify key performance indicators (KPIs) that are relevant to the fashion store (e.g., sales revenue, customer retention rate, inventory turnover) and track them over time.
- Provide recommendations to improve business operations and decision-making based on the insights derived from the data.
- Presentation and Documentation:
- Prepare a report summarizing the database design, data analysis process, and the insights gained.
- Create visualizations, charts, and graphs to present the findings effectively.
- Document the project's methodology, challenges faced, and any additional improvements or extensions that could be made.

This project will provide you students with hands-on experience in designing and implementing a database for an e-commerce store. It will also allow you to analyze real-world data, extract insights, and present their findings to demonstrate their understanding of data science principles.




**ER – Diagram**

## Analysis Questions:

The Analysis Questions is divided into 2 sections. Section A is only meant to test your use of SQL commands while Section B is the Business Case to be Study. **You are only meant to visualize Section B and not A.**

### A. Test Your SQL Skills

1. Select customer name together with each order the customer made.
2. Select order id together with name of employee who handled the order.
3. Select customers who did not placed any order yet.
4. Select order id together with the name of products.
5. Select products that no one bought.
6. Select customer together with the products that he bought.
7. Select product names together with the name of corresponding category.
8. Select orders together with the name of the shipping company.
9. Select customers with id greater than 50 together with each order they made.
10. Select employees together with orders with order id greater than 10400.
11. Select the most expensive product.
12. Select the second most expensive product.
13. Select name and price of each product, sort the result by price in decreasing order.
14. Select 5 most expensive products.
15. Select 5 most expensive products without the most expensive (in final 4 products).
16. Select name of the cheapest product (only name) without using LIMIT and OFFSET.
17. Select name of the cheapest product (only name) using subquery.
18. Select number of employees with LastName that starts with 'D'.
19. Select customer name together with the number of orders made by the corresponding customer, sort the result by number of orders in decreasing order.
20. Add up the price of all products.
21. Select orderID together with the total price of that Order, order the result by total price of order in increasing order.
22. Select customer who spend the most money.
23. Select customer who spend the most money and lives in Canada.
24. Select customer who spend the second most money.
25. Select shipper together with the total price of proceed orders.

### B. Exploratory Data Analysis

1. Total number of products sold so far.
2. Total Revenue So far.
3. Total Unique Products sold based on category.
4. Total Number of Purchase Transactions from customers.

5. Compare Orders made between 2021 – 2022
6. What is total number of customers? Compare those that have made transaction and those that haven't at all.
7. Who are the Top 5 customers with the highest purchase value?
8. Top 5 best-selling products.
9. What is the Transaction value per month?
10. Best Selling Product Category?
11. Buyers who have Transacted more than two times.
12. Most Successful Employee.
13. Most use Shipper.
14. Most use Supplier.

## Submission:

You're to submit this to your GitHub account. Due date will be communicated in class