Project Report

Database Design:

 This is a database structure of an e-commerce website (or a system including mobile or desktop applications) of a company selling only computer-related products.

  At first, I started creating this project with no specification on the product types, like a website selling almost everything (e.g. hepsiburada, trendyol etc.). However, after I started writing SQL code for the database tables, I realized that it would be imposible to develop this kind of system alone (at least like in 2 months). So, with the help of the instructor of the lecture (you) too, I decided to create a database for a website that sells products of only 3 categories and their 11 sub-categories:

Parent categories: Computers, Computer Parts, Peripherals
Sub-categories: Desktop, Laptop, Motherboard, Processor, Graphic Card, RAM, Computer Case, Power Supply, Monitor, Keyboard, Mouse

  It keeps parent categories and their subcategories as a value in "categories" table. Category-specific attributes are kept in the table "attributes" with the column "category_id" to specify category having this attribute. Products of these categories have these attributes through the table "attribute_values". This table keeps "product_id", "attribute_id", and the value of that attribute for the product having "product_id".

  Addition to attribute values, products have images too. Since one product can have (and generally has) multiple images, best solution to this was creating a table just for "product_images". This table has columns like "product_id", "image_url", "image_order", and "is_primary".

  "users" table is one of the main tables used in data analysis, has columns like "job", and "gender" which are specifically added for data analysis. Addition to these, it has columns "cart_id", and "wishlist_id" to bind it to the tables "carts", and "wishlists".

  "carts" and "wishlists" are similar tables identifying: "cart_items", and "wishlist_items". Although these tables are crucial for any e-commerce system, they do not provide us with much analysis.

  "orders" and "order_items" are other main tables for the data analysis. These provide us the information "who bought what, and how much?", which simply means money.

Triggers and Constraints:

1) **Update the total price of an order when its items are changed:**
   We have an order having id=1. Whenever a new item is inserted with order_id=1, "total_price" column in order table must be updated.
   It's same with the "update" operation. Whenever an order item having order_id=1 is being updated, total_price of the order should also be updated.

2) **Users must purchase the product to review it:**
   People should not be able to review products they have not bought, so it is being checked before review inserts.

3) **Products with no stock cannot be bought:**
   People should not be able to buy products having quantity=0, so it is being checked before cart_item inserts.

4) **Update product quantity after a new order:**
   Reduce the quantity of a product after a new order is set.

5) **Warn us when the quantity of a product is lower than a threshold:**
   Show a warning when a quantity of a product is lower than a threshold like 3 or 5, so we can do something about it.

6) **A user can review a specific product max once:**
   A user should not be able to review one product multiple times.

7) **Username's cannot be changed:**
   People should not be able to change their usernames for better user management and sustainability.

8) **Rating values must be between 1-5:**
   "rating" column in "reviews" table can have only values in range 1-5.

Data Analysis:

1) **Data Analysis for Products and Categories:**
   1.1) retrieving parent categories
   1.2) retrieving sub-categories for a category
   1.3) retrieving all attributes of a category
   1.4) retrieving all attributes and their values for a product
   1.5) filtering products using multiple attribute-value pairs
   1.6) retrieving all products with all of their attribute values
   1.7) retrieving products under a specific category
   1.8) retrieving average price per category
   1.9) retrieving most reviewed products
   1.10) retrieving products with their average review rating information

1.11) retrieving all products with their stock and pricing information

1.12) retrieving products with low stock

1.13) retrieving the total value of inventory: sum of (price * quantity) for each product

2) **Data Analysis for Orders, Order Items, and Users:**

2.1) retrieving "number of orders", "total money spent", "average money spent per order" information for each user

2.2) retrieving products each user bought: showing "who bought what, and how many times?"

2.3) retrieving users who spent the most, top 10

2.4) retrieving number of orders for each order status ("CREATED", "SHIPPED", "DELIVERED", "CANCELED")

2.5) retrieving products most frequently bought, top 10

2.6) retrieving average order total prices and total spent for each gender: which gender spends more?

2.7) retrieving order history of users

2.8) retrieving orders per month

2.9) retrieving orders per year and month

2.10) retrieving orders per day

2.11) retrieving top categories among customers

2.12) retrieving top brands among customers

2.13) classifying users based on their spending and purchase frequency

2.14) filtering orders by date range

2.15) retrieving the number of products bought in each category by each job

2.16) retrieving how many each job spent

2.17) retrieving products bought together

2.18) retrieving the earning by category and by month: how many earning did we get from which category and which month