# CS2208 INFORMATION STORAGE AND MANAGEMENT I

*2023/2024*

**Assignment**

**Due: 2359hrs 17th November 2023**

**The assignment is worth 10% of the final module score**

**All work must be your own**

You are supposed to be creative with your thinking and make your own assumptions if needed. Grading is based on the demonstration of domain understanding, technical correctness, design comprehensiveness, and innovativeness. Your database design should involve at least TEN (10) entities and at most SIXTY (60) entities with well-documented relationships.

**Project:** **2**. A Chatbot System for E-commerce Customers

**Question 1**. Develop your system description using fact finding technique(s).

**Part A)** Describe how you conduct your fact finding. You MUST declare any external sources if you reference them in your fact findings (e.g., Website URLs, Books or Articles).

Group Member 1:

***(Word Limit: 300 words)***

|  |  |  |
| --- | --- | --- |
|  | As we embarked on our assignment centred around eCommerce chatbots, I recognised the importance of grounding our work in real-world examples. Thus, I decided to study leading eCommerce businesses, in particular a deep dive into Amazon and eBay as my main points of examination. I chose Amazon and eBay as they both offer straightforward chatbot systems that would be easier for me to replicate, but both possess a wealth of knowledge online that would be readily available to me for a comprehensive study. |  |
|  |
|  | I started off with interacting with these company’s chatbots, to help me craft the foundation of our chatbot, how questions were asked, and how answers were processed. To deepen my understanding, I then studied technical blogs, alongside academic papers, and industry analyses, all of which gave me a greater understanding of how chatbots run.  Some examples of resources I look at include: |  |
| |  | | --- | | * <https://www.amazon.science/blog/amazon-com-tests-customer-service-chatbots> * Miklosik, N. Evans and A. M. A. Qureshi, "The Use of Chatbots in Digital Business Transformation: A Systematic Literature Review," in IEEE Access, vol. 9, pp. 106530-106539, 2021, doi: 10.1109/ACCESS.2021.3100885. |   More detailed database architectures were hard to come by due to the private and proprietary nature companies, but the information I did uncover was incredibly enlightening. After doing all this I decided what about these chatbots I liked, and what needed to be added to them so I could make a new and improved chatbot.  The focus of my analysis zeroed in on finding out how to design our database schema— how is everything organised, and what sort of data relationships we were looking at? I needed to understand not just how they store their data—customer interactions, product specs, and the logs of every little chat—but how they sift through this data to fetch what's needed when a customer pops a question.  Through this fact-finding exercise I was able to map out a database schema that I believe is both strong and flexible, ensuring it can handle the intricate nature of queries that modern eCommerce demands. |

**Part B)** Discuss and develop a system description based on your fact findings.

***(Word Limit: 1500 words)***

Regarding the significance of our fact-finding methodology towards the development of our system, the team’s efforts were well-rounded; including insights from industry leaders, stakeholder interviews, documentation reviews, competitor analysis and feasibility for implementation. We considered the past industry practices as well as what our users would need and the feasibility of certain aspects we imagined for the chatbot.

We first considered these factors in creating this design, such as:

Initial Database Size – We must assess how much data will initially be stored. It is critical to analyse initial storage requirements.

DB Rate of Growth – It is vital to observe how quickly the database may grow. This can be

|  |
| --- |
| done via analysing past information or predicting the expected interactions by the user. We must ensure the system can grow without slowing it down, to make management easier as well as providing a seamless experience for users.    Number and complexity of queries – It is important to enable the chatbot to handle high volumes of queries if we expect many users to be on our platform at once. Additionally, the expected complexity of these queries that the chatbot can handle must be considered. We can optimise the schema to hasten and simplify the chatbot’s responses.    Security issues – Ties into access requirements. User data must be kept confidential and protected. Potential vulnerabilities must be assessed and the chatbot should only be able to access data that is relevant to the specific user to protect other users’ sensitive information.    Integration with existing systems – The chatbot should be able to seamlessly integrate with existing systems.    The team took a collaborative approach to designing the entities, relationships and attributes needed to ensure all bases were covered and that any potential problems were more likely to be spotted.  That said, I will now describe the system design my team and I have imagined.  Entities:  User – This entity describes the user which interacts with the chatbot. All attributes which may be significant to it are detailed. Such attributes include the user ID(key), the user’s name, password and email address. This entity is the integral piece which most other entities revolve around, ensuring that the chatbot can access important information relating to the user when necessary. This will be further elucidated through the entities listed below.  User Reviews – This entity is important as we intend for the chatbot to interact with these reviews to reduce the load of real-time operators who may otherwise be responsible for responding to these reviews. If the chatbot has access, then it will be able to collect information as to the positives and negatives provided by users about certain products, automating data collection. Such attributes for this entity would include review id(key), rating, review text and review date. All are important to understanding the context of why a certain review may be given to the product in question.  Order – As we are an online clothing store a customer may have problems with a past order of theirs, perhaps regarding a payment issue or the status of the order. Attributes: orderid(key), date, amountspent, status.  Shopping Cart – If a customer is having difficulty with a purchase or has had items in their cart for an extended period, we intend for the chatbot to pop up and assist the customer, or to incentivize the customer to continue with their purchase. Additionally, it may suggest similar or related items to the ones that are in the cart to sway the customer to purchase more than they originally intended. Attributes for this entity may include the cart id(key), created date and last updated.  Cart Items – This relates to the shopping cart as the chatbot would need to access the products that are in the cart to find related products to suggest. Additionally, it would be important if the customer is having difficulty purchasing an item, as the chatbot would then be able to recognize it is out of stock, for example. Attributes tied to this entity would be cart item id(key), user id and quantity.  Returns – If a user intends to return an item but is having difficulty doing so, they may request service for clarification or information as to the company’s return policy. This is the chatbot's |
| primary function, as it will be able to easily provide the user with the information they would need.  Order Items –  This would detail the specific items in an order. It’s important for the chatbot to be able to access the order items to see if a certain product in the order may have an issue. Attributes: order item id(key), quantity, price per unit  Product – This is a useful entity for the chatbot to access if a customer is having difficulty with a certain product on our site. It can quickly access all important details about the product and swiftly determine where the issue may lie. This could include the product’s name, description, quantity, price, etc. Therefore, attributes tied to this entity would be product id(key), its name, description, quantity, price and image url.  Clothing Size – As we decided for our website to exclusively sell clothing, we found it important to have an entity which focuses on the size of the article of clothing. This is a simple entity which would just have the attributes of size id(key) and size description.  Product Category – This is another entity which will simplify the process for the chatbot in finding certain products, as well as categorising them for users if they wish to browse our selection of broad types of clothing. Attributes for this would be category id(key), the name and description.  Material – Another simple entity which describes the material of the product. Attributes would include material id(key) and description.  Product Colours – Same as above, just describing the colour of the product. Attributes would be colour id(key) and name.  Shipping Status – This entity will describe the current shipping status of any pending order the customer has. The chatbot can provide customer support to provide details on the whereabouts of the user’s order. Attributes: status id(key), tracking number, current status, estimated delivery date  User Feedback – The chatbot will be able to collect valuable information from users by prompting them for their feedback. It will then be stored in this entity. Attributes:  feedback id(key), review string, rating  Shipping Information – This will contain all information about the location which the user’s order will be shipped to. It is important for the chatbot to know in case there are limitations for shipping to the input destination. Attributes:  shipping id(key), postal code, city, address, country  User Session – The central entity for the chatbot to know how long a user’s query has been outstanding for, so the chatbot can know when to close the session so that memory isn’t wasted. Attributes: session id(key), start time, end time  Query – This entity will save any query users have so the chatbot can store this information for the company to learn about any consistent issues or questions users may have. It also helps the chatbot learn how to respond to recurring queries. Attributes: query id(key), query string, response  FAQ - This entity logs all questions that have been asked prior and the frequency at which they are asked. It will then be accessed by the chatbot to display the most frequently asked questions to the user. Attributes: category id(key), frequency, question, answer  Query category – This includes broad categories the user’s query will pertain to, such as returns, products, orders, etc. It helps the chatbot narrow down the area the customer requires assistance in. Attributes:  category id(key), description |

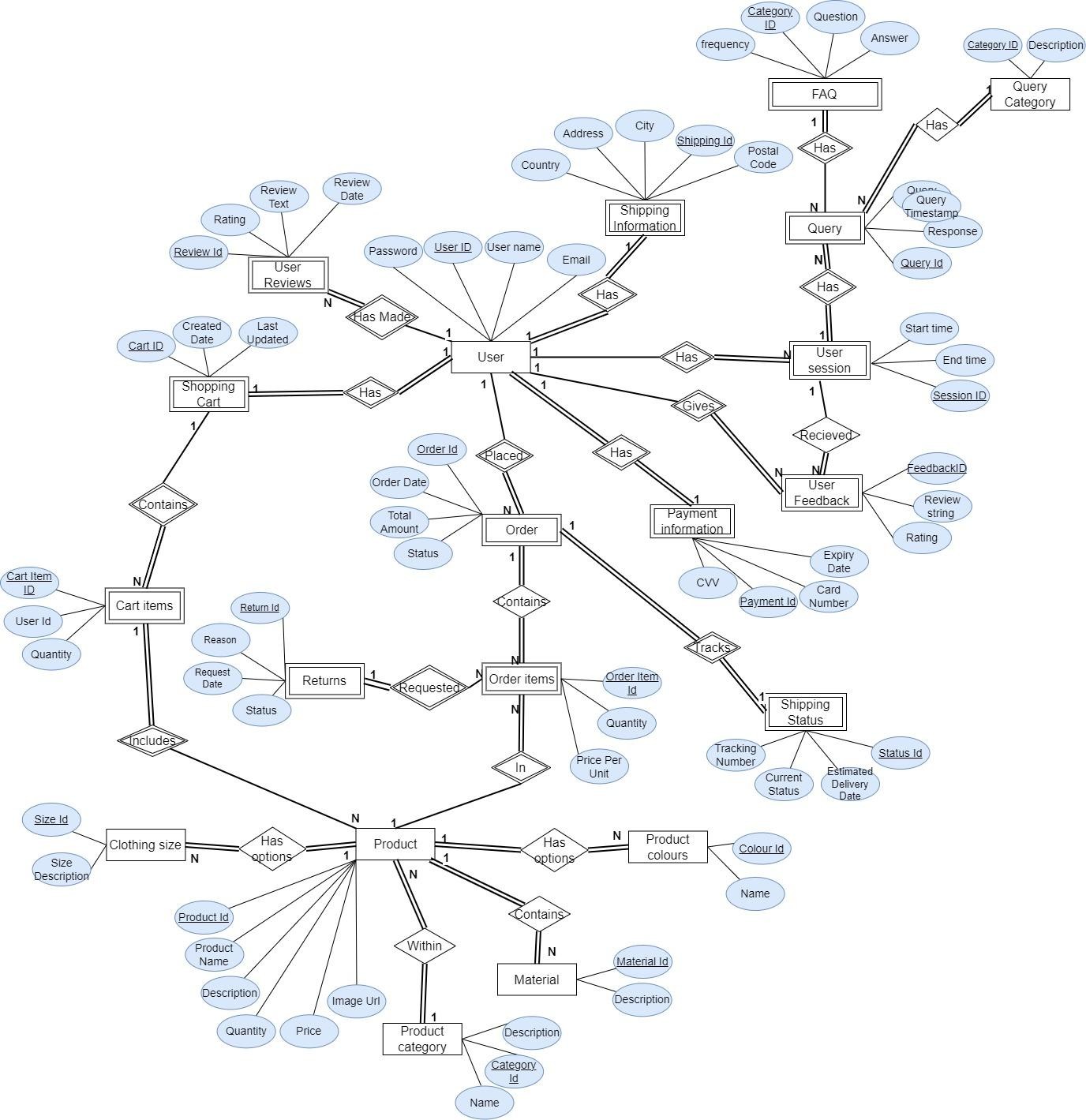
|  |
| --- |
| Relationships:  User <placed> Order - 1...1 - 1...N  User <has> Shopping Cart – 1...1 - 0...1  User <has made> User Reviews – 1...1 - 0...N  User <has> Shipping Information – 1...1 - 1...N  User <has> Payment Information – 1...1 - 1...N  User <placed> Order – 1...1 - 1...N  User <has> User Session – 1...1 - 0...1  User <gives> User Feedback - 1...1 - 0...N  User Session <received> User Feedback - 1...1 - 0...N  User Session <has> Query – 1...1 - 0...N  Query <has> FAQ – 0...N - 0...1  User Session <has> Query – 1...1 - 0...N  Order <contains> Order Items – 1...1 - 1...N  Product <in> Order Items – 1...1 - 1...N  Shopping Cart <contains> Cart Items – 1...1 - 0...N  Cart Items <includes> Product – 1...N - 1...1  Product <has options> Clothing Size – 1...1 - 1...N  Product <within> Product Category – 0...N - 1...1  Product <contains> Material – 1...1 - 1...1  Product <has options> Product Colours – 1...1 - 1...N |

**Question 2**. Draw a logical ER diagram using your system description. Use the Chen’s notations and symbols.

***(***

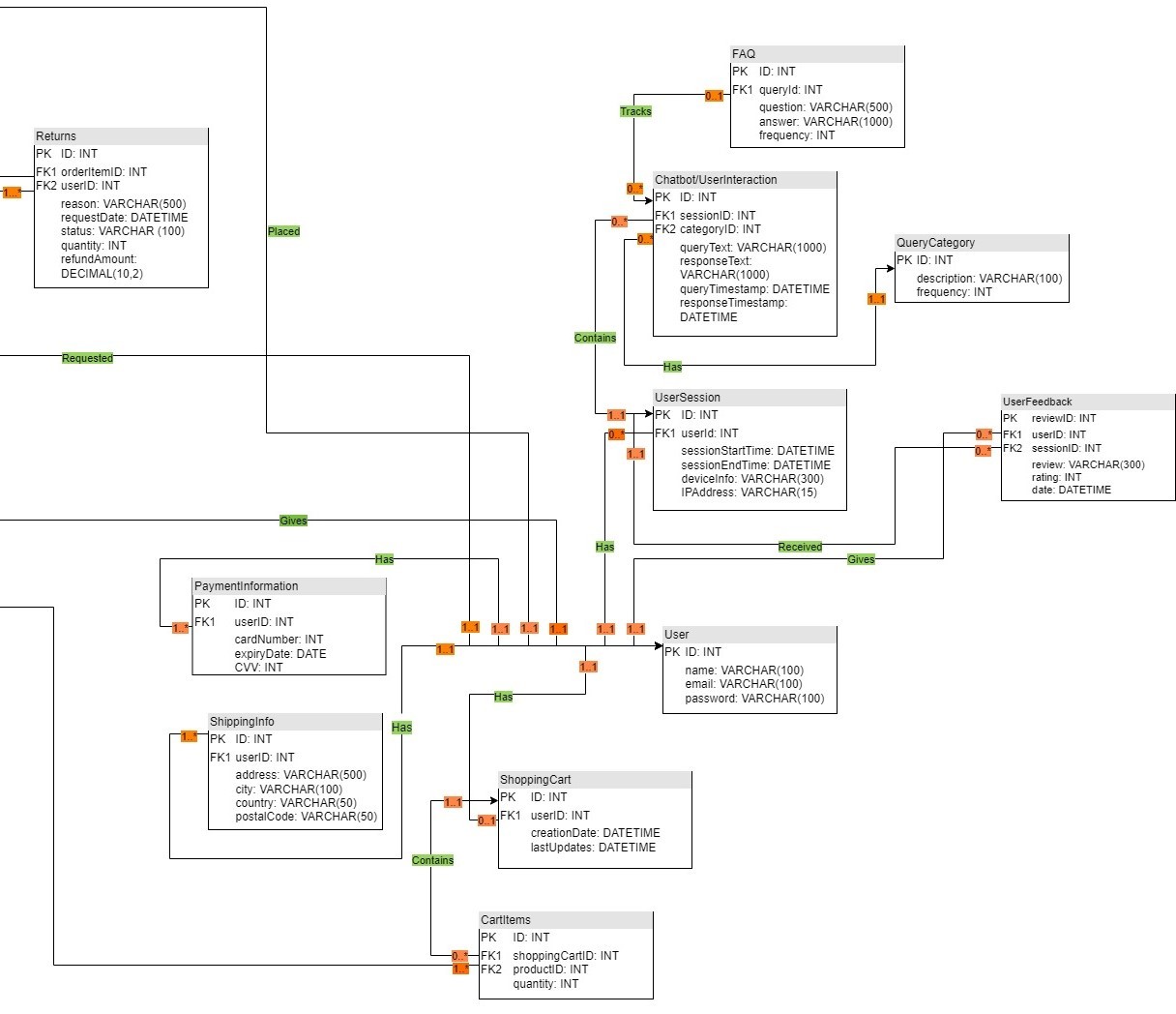
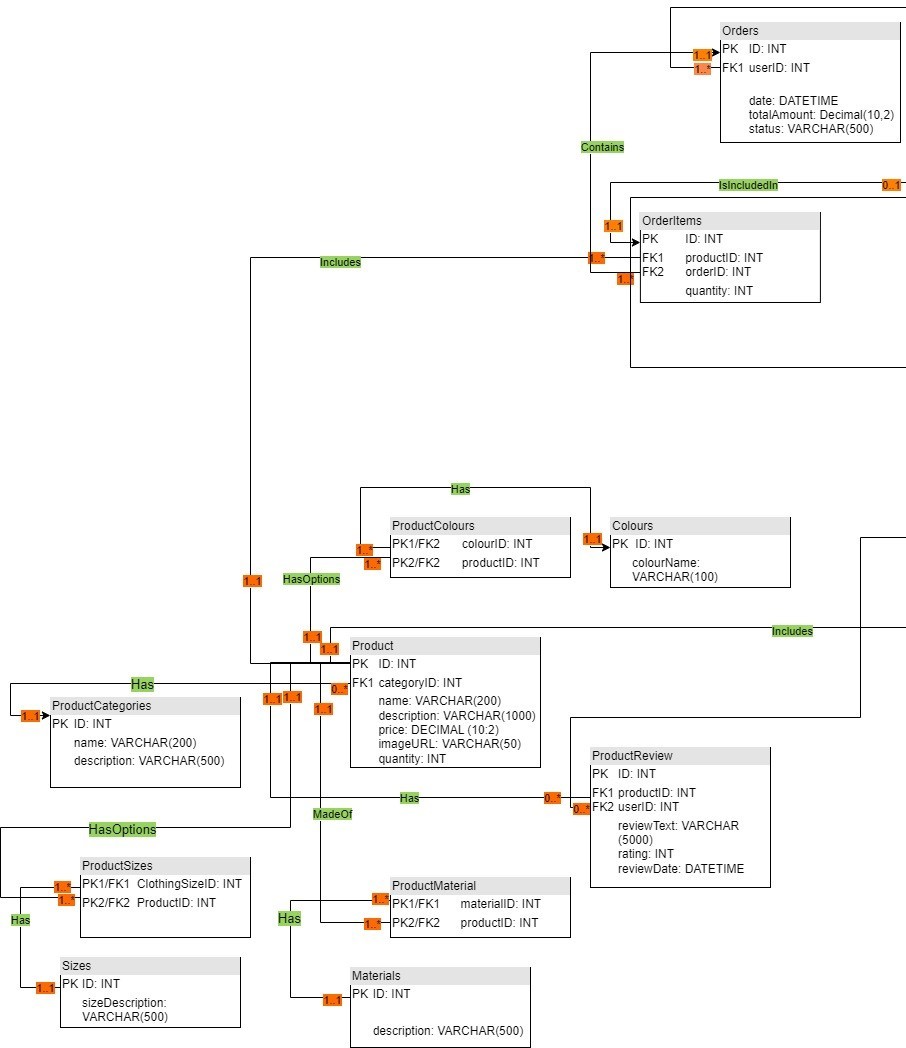
***Page Limit: 3 pages***

***)***



**Question 3**. Normalise your database to the Boyce–Codd normal form (BCNF). Draw a physical ER diagram of the normalised database using UML notations and symbols.

***(Page Limit: 2 pages)***



**Question 4**. Write SQL Queries to create database tables and insert dummy data (3-10 rows). Each group member is responsible for working on a sub-set of the database.

Group Member 1:

***(Word Limit: 1000 words)***

**Tables**

CREATE TABLE `Orders` (

`ID` INT AUTO\_INCREMENT PRIMARY KEY,

`userID` INT,

`date` DATETIME,

`totalAmount` DECIMAL(10,2),

`status` VARCHAR(500),

FOREIGN KEY (`userID`) REFERENCES `users`(`ID`)

);

CREATE TABLE `ProductCategories` (

`ID` INT AUTO\_INCREMENT PRIMARY KEY,

`name` VARCHAR(200),

`description` VARCHAR(500)

);

CREATE TABLE `OrderItems` (

`ID` INT AUTO\_INCREMENT PRIMARY KEY,

`productID` INT,

`orderID` INT,

`quantity` INT,

FOREIGN KEY (`productID`) REFERENCES `products`(`ID`),

FOREIGN KEY (`orderID`) REFERENCES `Orders`(`ID`)

);

CREATE TABLE `Returns` (

|  |
| --- |
| `ID` INT AUTO\_INCREMENT PRIMARY KEY,  `orderItemsID` INT,  `userID` INT,  `reason` VARCHAR(500),  `requestDATE` DATETIME,  `status` VARCHAR(100),  `quantity` INT,  `refundAmount` DECIMAL(10,2),  FOREIGN KEY (`orderItemsID`) REFERENCES `OrderItems`(`ID`), FOREIGN KEY (`userId`) REFERENCES `users`(`ID`)  );  CREATE TABLE `Colours` (  `coloursID` INT AUTO\_INCREMENT PRIMARY KEY,  `colourName` VARCHAR(100)  );  CREATE TABLE ProductColours (  `coloursID` INT,  `productID` INT,  PRIMARY KEY (`coloursID`, `productID`),  FOREIGN KEY (`coloursID`) REFERENCES `Colours`(`ID`),  FOREIGN KEY (`productID`) REFERENCES `products`(`ID`)  );  **Insertion of Dummy Data**  INSERT INTO `Orders` (`userID`, `date`, `totalAmount`, `status`)  VALUES  (1, '2023-11-11 10:00:00', 110.00, 'Processing'),  (2, '2023-11-14 11:30:00', 85.00, 'Shipped'),  (3, '2023-11-16 09:45:00', 95.00, 'Delivered'),  (4, '2023-11-12 14:00:00', 95.00, 'Processing'),  (5, '2023-11-11 16:00:00', 40.00, 'Cancelled'), |
| (5, '2023-11-10 17:00:00', 75.00, 'Delivered');  INSERT INTO `ProductCategories` (`name`, `description`)  VALUES  ('T-Shirts', "Just your regular T-shirts that go with pretty much everything. Perfect for when you want to keep it simple."),  ('Jeans', "Everyone's got a favorite pair of jeans, right? We've got lots - from those comfy ones you wear on a lazy day to ones that dress you up a bit."),  ('Dresses', "Whether it's a sunny day out or a small party, we've got a dress for that. Easy to throw on and always looks good."),  ('Jackets', "Need something for chilly days or just to add some style? Our jackets range from light ones to those warm, puffy ones."),  ('Shoes', "Casual kicks, fancy heels, comfy flats - shoes for days when you're out and about, and even for those stay-at-home times."),  ('Accessories', "Belts that hold it together, hats for sunny days, and scarves for the chilly ones. Sometimes, it’s all about the accessories.");  INSERT INTO `OrderItems` (`productID`, `orderID`, `quantity`)  VALUES  (1, 1, 2),  (2, 1, 1),  (3, 2, 3),  (4, 2, 1),  (5, 3, 2),  (1, 3, 1),  (2, 4, 1),  (3, 4, 2),  (4, 5, 1),  (5, 6, 3);  INSERT INTO `Returns` (`orderItemsID`, `userID`, `reason`, `requestDATE`, `status`, `quantity`, `refundAmount`)  VALUES  (1, 1, 'Too large in length.', '2023-11-17 10:00:00', 'Pending', 1, 45.00),  (8, 4, 'Incorrect item received.', '2023-11-18 12:30:00', 'Approved', 2, 80.00), |

(10, 5, 'Did not match the photo.', '2023-11-20 14:45:00', 'Completed', 1, 25.00);

INSERT INTO `Colours` (`colourName`)

VALUES

('Black'),

('White'),

('Red'),

('Blue'),

('Green'),

('Grey');

INSERT INTO `ProductColours` (`coloursID`, `productID`)

VALUES

(1, 1),

(1, 5),

(2, 5),

(3, 3),

(3, 5),

(4, 1),

(5, 2),

(5, 4),

(6, 4);

**Question 5**. Write your reflections on contributions and lessons learnt.

Group Member 1:

***(Word Limit: 300 words)***

|  |  |
| --- | --- |
| Now that we’ve finally finished our assignment, I’d like to quickly mention my personal contributions and lessons I’ve learned from this experience.  We started off the process with completing different fact-finding approaches, where I focused on researching the industry eCommerce leaders and their respective chatbots. This involved identifying their strengths and finding the areas of improvement in their chatbot implementations.  Building on this, I then took the initiative to draft up a preliminary list of entities and their attributes needed for an efficient chatbot. This was a collaborative effort, where some entities were discarded, and others added after group discussions.  Another significant portion of my effort went into collaborating with my group members to come up with the two ER diagrams: one in Chens’s notation, and the other in UML, fully normalised and in BCNF form.  Reflecting on my journey through this project, I’ve realised just how little experience I’ve had with making ER diagrams, and now I am more than confident I would be able to make one in a professional setting. It was a steep learning curve, with aspects such as multiplicity and normalisation starting off as a very confusing topic for me. I’ve also greatly improved my SQL capabilities during this process, which will be a great asset when I go into the world of working.    I believe my personal rating of 9/10 is more than justified for the effort that went into the project. With this being my first college group project, I had to learn to work as a team, and I believe I was able to do so, putting all help and effort into the assignment. I also have my team members to thank for having such a successful project.   |  | | --- | | **Self-rating: 9 / 10** | |