

**CST2340**

**Database Systems: - Design and Implementation 2020-21**

**Coursework 1: - Part A**

**Database Design Case Study**

**Module Tutor: - Ms. Engie Bashir**

**Done By: -**

**Alester D’Costa (M00734829)**

**Table of Contents**

[A1. Case Study Description 3](#_Toc150339227)

[A2. Entity-Relationship Diagram 4](#_Toc150339228)

[i. Entity-Relationship Model 4](#_Toc150339229)

[ii. Relational Model 5](#_Toc150339230)

[A3. Validation of the Model 6](#_Toc150339231)

[i. Connection Traps 6](#_Toc150339232)

[ii. Functional Requirements 9](#_Toc150339233)

[A4. Professional, legal and ethical issues 10](#_Toc150339234)

[A5. Limitations of the relational model 11](#_Toc150339235)

[i. Relational model and its limitations 11](#_Toc150339236)

[ii. NoSQL and ways it overcomes the limitations of the relational model 11](#_Toc150339237)

[A6. References 13](#_Toc150339238)

# **Case Study Description**

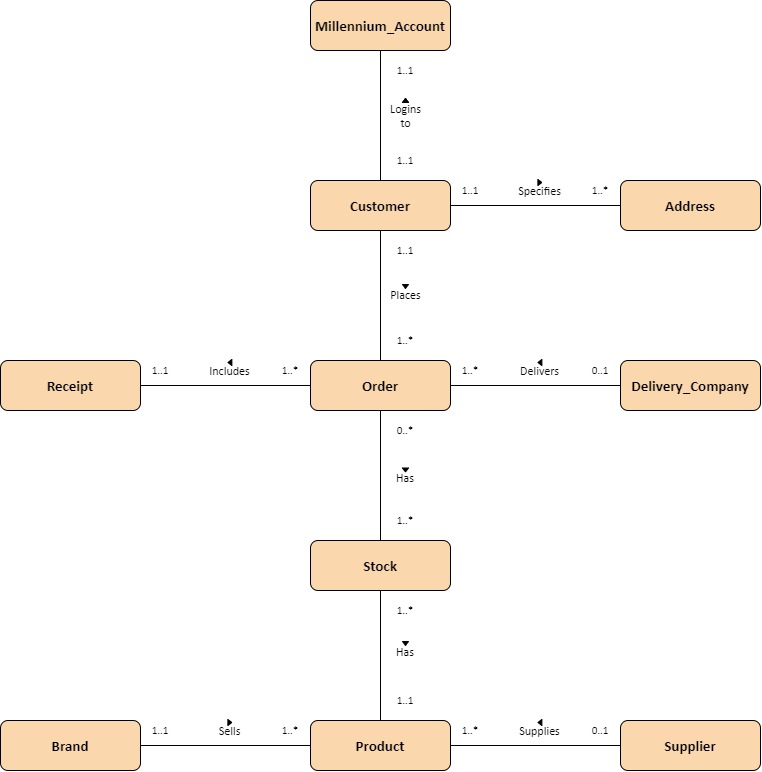
**Millennium Dealers** is a newly scaled e-commerce company based in Dublin, Republic of Ireland. It has a presence in all areas of the country. It deals in the sale of t-shirts, shirts and hoodies of various well-known brands such as Adidas, Puma, Levi’s and many more. A website has been created by the web designer and is linked to the Millennium database developed by the database designer that helps the company in managing orders placed by various customers and the delivery of products to the customers. This would help the company in taking better business decisions and to improve its performance based on the stored data.

Before the customers can place an order, the customers need to login to their account with an email address and password. The company has multiple customers with their various details such as a unique ID, customer name, and a contact number. A customer places an order with details that include a unique ID and order Date and each order has multiple stock with details that include a unique ID, color, size and quantity. Due to the company being a start-up, the customer can facilitate the payment of the order only through cash on delivery. Each stock has a single product and details of a product include a unique ID, product name, product type (i.e. t-shirt, shirt and/or hoodie), gender and price of the product. A customer may be interested in a particular brand and therefore, one brand would sell multiple products. Some details of a brand include a unique ID, brand name, and country of origin. The suppliers have the responsibility of supplying the products and its details include a unique ID, supplier name, contact number, address and city. During the ordering process, the customer needs to specify his/her address. The customer may have multiple addresses, but for a single order, only one address is taken into account based on the customers preference. The details that come under address includes a unique ID, description, city, P.O box and address type (i.e. either a residential, company or other address).

When a customer has placed the order, the order is then delivered by a delivery company within the next 3-4 business days. A delivery company handles multiple delivery of orders and its details include a unique ID, delivery company name (FedEx, DHL or Aramex) and date of delivery. Each order is delivered only by one delivery company and the delivery charges are free of cost. When the order has been delivered to the customer, a receipt is included along with the delivery package that gives a proof of the purchase of goods by the customer. The details of the receipt include a unique ID and date.

# **Entity-Relationship Diagram**

## **Entity-Relationship Model**

The entities that have been identified from the case study are: -

|  |  |
| --- | --- |
| * Millennium \_Account | * Customer |
| * Order | * Stock |
| * Product | * Brand |
| * Supplier | * Address |
| * Delivery\_Company | * Receipt |

## **Relational Model**

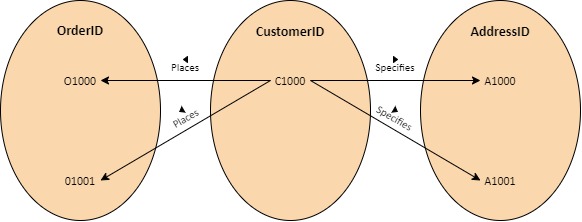
# **Validation of the Model**

## **Connection Traps**

There are two types of connection traps: - **Fan traps** and **Chasm Traps.** Based on the relational model of the case study prepared, a **fan trap** has been identified, as illustrated by the following: -

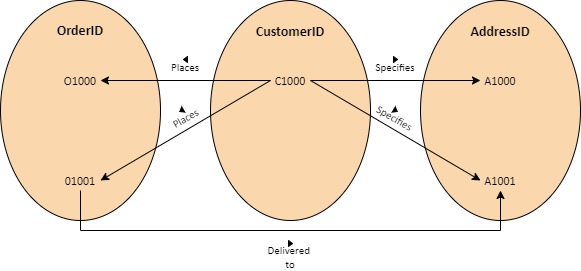
* + **Identification of trap: -** The fan trap that has been identified as per the case study is the relationship between order and address entities, where there is a relationship, but there is no path between them.

In the following individual example of the fan trap, it becomes difficult to identify which order of the customer has to be delivered to a particular address specified by the customer.



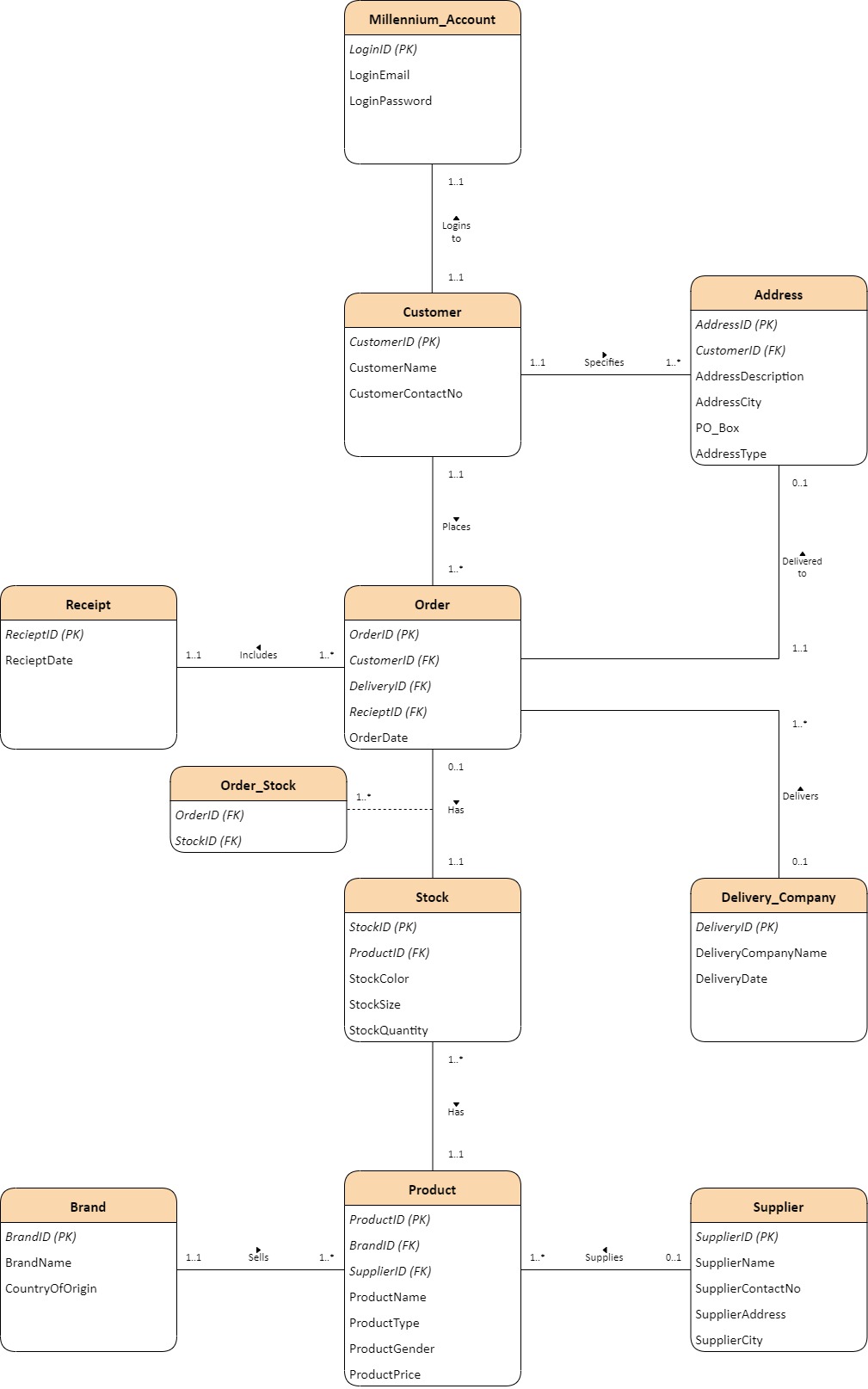
* + **Resolving the trap: -** The fan trap is resolved by establishing a relationship between order and address.

With the trap being resolved, it becomes easier to understand which order of the customer has to be delivered to the address specified by the customer.

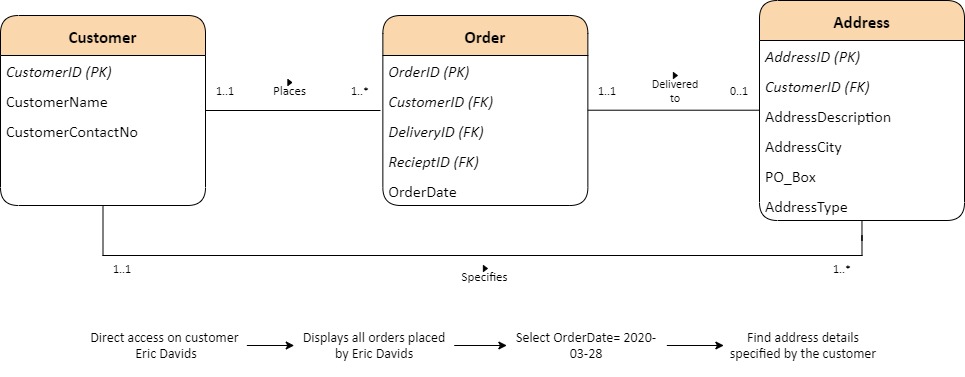


* + **Reordering of entities: -** The entities have been reordered to establish a relationship between the entities.

After resolving the fan trap, the relational model is revised and represented.



## **Functional Requirements**

Finding out address details of an order such address description and city specified by a customer named Eric Davids for an order placed on the date 28th March 2020.

# **Professional, legal and ethical issues**

According to the case study, Millennium Dealers is an e-commerce company based in the Republic of Ireland. As a member country of the European Union, any organization based in the Republic of Ireland is subject to the **General Data Protection Regulation (GDPR)** laws of the European Union law. As part of the GDPR of the European Union, the **data controller** and **data processors** have also been stated in the case study.

According to (Dubrovskaya, no date), a data controller refers to any company, individual or body which specifies the conditions and measures of processing personal data, controlling of data and is accountable, alone or jointly. This could be a non-profit organization or any other form of organization that stores personal information about individuals employed in the company. According to (Dubrovskaya, no date), a data processor is an individual or organization who on behalf of the data controller, processes the data. This could be a financial company like VISA and Mastercard.

Based on the case study, the **data controller** identified is Millennium Dealers. They interact with the data processors, i.e., the delivery companies for using customer data after being subject to a contract signed. The **data processors** are the delivery companies (DHL, FedEx, or Aramex). It protects and safeguards any customer's information, such as his/her name, address, emails, and more data. Being a third-party company to Millennium Dealers, the delivery companies would act under a contract signed between the two parties (any delivery company and Millennium Dealers), which stipulates that personal information would have to be shared on behalf of the customers after delivering an order. This is indicated during the ordering process, where the customer would have to agree to the terms and conditions set by any delivery company that will deliver the order.

# **Limitations of the relational model**

## **Relational model and its limitations**

A **relational model** is defined based on the mathematical theory of a relation, which is physically described as a table (Connolly and Begg, 2015).

The limitations of the relational model are: -

* **Cost: -** A drawback of relational databases is its cost of setting up and managing the database management system. A particular software needs to be acquired to set up a relational database.  Regardless of the company's size, any personal data of the customer, such as his/her credit card number, security numbers, and more, must be stored securely to fulfill the regulatory standards (Martin, no date).
* **Abundance of Information: -** Another drawback of relational databases is the wide variety of information stored, leading to data complexity. A relational database is developed to arrange data that have common characteristics. Simple classification defies complex images, figures, layouts, and digital items, paving the way for a new type of database called object-relational database management systems, which has been designed to accommodate applications that are more complex and can be extensible (Martin, no date).
* **Structured Limits:** - There are restrictions on field lengths in individual, relational databases. The quantity of data that can fit into a field needs to be stated during the database development. Several names or queries are lesser than the actual ones; this could result in data loss (Martin, no date).
* **Isolated Databases: -** A relational database system that is complex could result in complexity where it becomes difficult to share information from one system to another. A large organization may use distinct relational databases in different divisions. An example could be a hospital where a single database is used by the finance department, while the other staff may utilize a separate database (Martin, no date).

## **NoSQL and ways it overcomes the limitations of the relational model**

**NoSQL** refers to a class of database management systems.  These databases are not meant to replace relational databases but instead focuses on providing alternatives to relational models. (Rogue Wave, 2016).

The following points explain the classification of the NoSQL model: -

* **Key-Value: -** NoSQL databases are designed to hold key-value pairs. Key-value databases are developed for quick, unprocessed-speed trade durability. Although a string is used in typical key-value data structures, the value is not restricted to a raw string in a key-value database. It may still retain complex data structures, including lists, sets, hashes, and bit arrays. Redis is an instance of a key-value database (Rogue Wave, 2016).
* **Document-based: -** NoSQL databases are identical to databases of key values, but store documents in the value section. The documents stored are usually in self-described forms, such as XML, BSON, and JSON. MongoDB is an instance of a key-document database (Rogue Wave, 2016).
* **Column based: -** To store and process quite large quantities of data shared over several systems, NoSQL databases have been developed. Keys also exist, but they refer to many columns. The family of columns organizes the columns. Cassandra and HBase are instances of column-based databases (Rogue Wave, 2016).
* **Graph-based: -** References to several other rows and relations are implied in relational databases by referencing foreign-key columns to their primary-key attributes. Relationships are first class citizens of the graph data model and it is swift to traverse the joins or relationships. The relationship between nodes is not determined at the time of the query but is eventually maintained as a relationship. Neo4j is one instance of a graph-based database (Rogue Wave, 2016).

# **References**

Connolly, T. and Begg, C. (2015) *Database Systems: A Practical Approach to Design, Implementation, and Management*. Harlow: Pearson.

Dubrovskaya, S. (no date) *EU GDPR Knowledge Base*. Available at: https://advisera.com/eugdpracademy/knowledgebase/eu-gdpr-glossary (Accessed: 8 November 2020).

Martin, A. (no date)*Disadvantages of a Relational Database*. Available at: https://www.techwalla.com/articles/disadvantages-of-a-relational-database (Accessed: 10 November 2020).

Rogue Wave (2016) *Overcoming Relational Database Limitations with NoSQL*. Available at: https://www.roguewave.com/sites/rw/files/attachments/RW-Relational-database-and-NoSQL-WP-FNL-20160706.pdf (Accessed: 10 November 2020).