## **Part 1**

### **a)**

As B2B portal’s website works with a very huge amount of data, NoSQL queries will be very beneficial because it can handle a large number of datasets at a very high speed, can store and retrieve structured, unstructured, and semi-structured data, is more flexible and scalable, is low cost which provides a better B2B back-end data storage solution. Graph database will be used for the B2B portal’s website.

### **b)**

As there will be lots of activities of searching, buying, selling by the users from the different parts of the world on the website, it is really necessary to get the results back very fast as possible. Therefore, the Hadoop framework will be needed which works in distributed and parallel systems. This will use distributed file storage system and allows parallel and distributed processing which will give us results back at a very high speed for a better user experience.

### **c)**

There will be two different nodes for both products and customers. The order node will link the customers to the products they had bought. The customers will have a relationship with the order node and the order node will have relationships with the products they had bought. The order node will contain the total price, status, order date, etc. and the relationship between the order node and the product node will give the quantity of the product that they had bought.

### **d)**

The columns related to the activities of the customers will be required for the comparison while recommending products. The reviews text column, rating column, description of the wish-listed, and ordered product column will be used to recommend similar products to the customer. These are the columns that can identify similarities and differences between the products that the customer had bought or reviewed or wish-listed with the other products. The columns might also verify according to the algorithms used for recommending products.

### **e)**

Both nodes and relationships will be very important for better data design solutions. User, Seller, Buyer, Product, Orders, Payment, Address are the nodes that will be required. The relationships that will be required between these nodes are shown below in the table.

Table 1: Relationships between nodes

|  |  |
| --- | --- |
| **Relationship** | **Nodes** |
| Business | User --> Seller |
| Business | User --> Buyer |
| Sells | Seller --> Product |
| Wish List | Buyer --> Product |
| Shopping Cart | Buyer --> Product |
| Review | Buyer --> Product |
| Recommended | Product --> Buyer |
| Ordered | Buyer --> Orders |
| Ordered Product | Orders --> Product |
| Pay | Orders --> Payment |
| Delivery | Orders --> Address |

### **f)**

Some nodes are already split into other smaller nodes. For example, users are divided into sellers and buyers, and order detail is divided into the Ordered, payment, address, and “Ordered Product” relationship. The splitting into nodes and relationships is done for more flexible writing and accessing of data. Each node used for designing the database can be uniquely identified, contains unique information, and cannot be split furthermore.

### **g)**

Instead of having the whole database available for real-time querying, it is better to have the nodes that are frequently being used for real-time querying because all the nodes will not be used at the same time. Every node has its timing of use. Some are very rarely used and some are used mostly. Therefore, real-time querying in separate nodes when they are required will be a more effective solution, and the database connection will be made when needed not every time.

## **Part 2**

### **a)**

A diagram for explaining the data storage solution is given below.

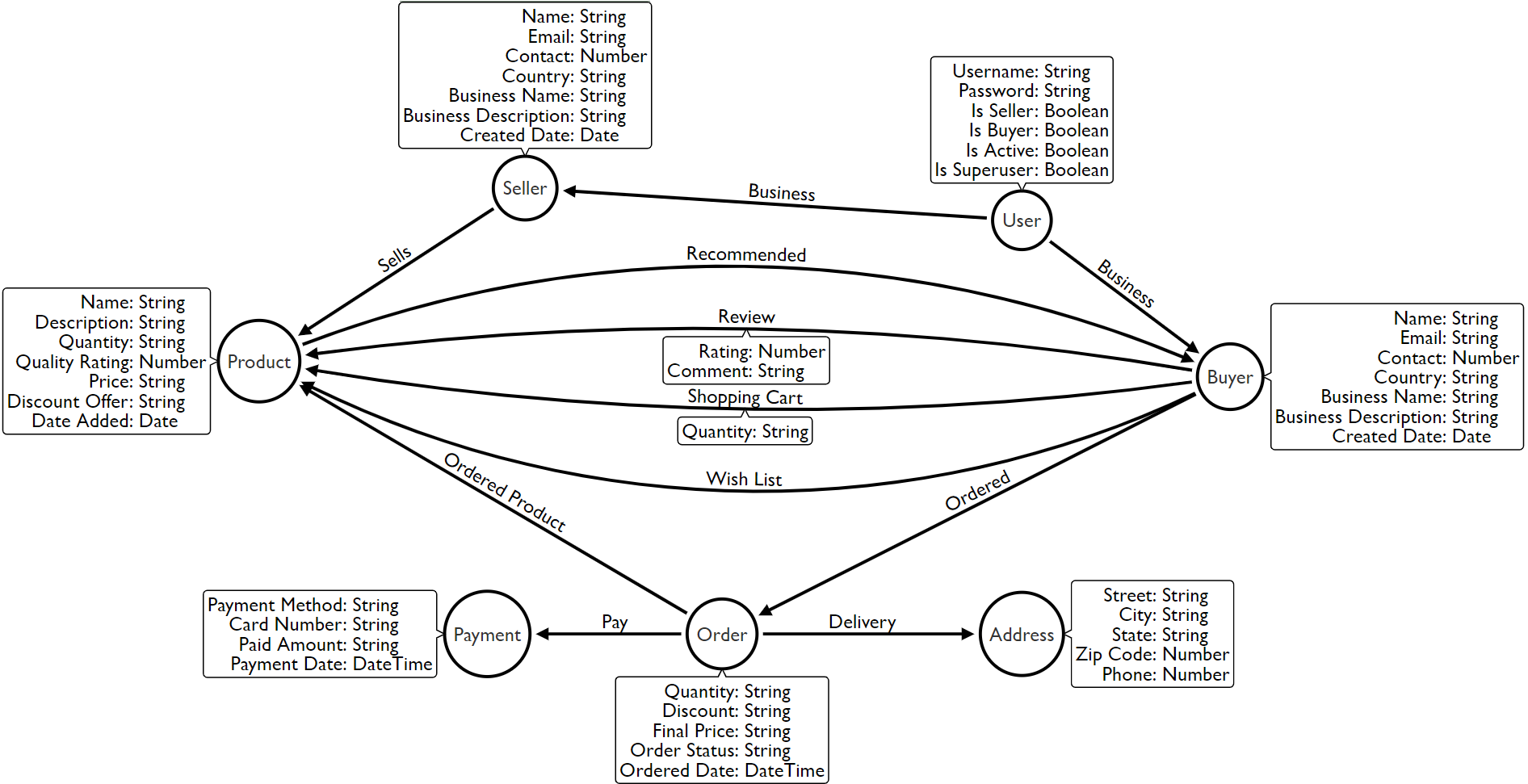


Figure 1: Data Model for Graph Database Design

The users are divided into sellers and buyers, and if the “Is Seller” property of the user is true then the user will be the seller but if “Is Buyer” is true then it will be the buyer. Sellers can add products and the “Sells” relationship will be created between them. The “Quality Rating” property of the products is calculated by the average rating given by the buyers to the product. The review, shopping cart, wish list activities of the buyer will be recorded with the respective relationship with the product. The product will have a relationship with the buyer if it is recommended to the buyer. When a buyer will order a product then an “Order” node will be created and the “Ordered Product” relationship will link the node with the ordered product. The “Order Status” property will tell whether the order was delivered or not. The user will also have to pay and give the delivery address to complete the order. After that, the “Payment” and “Delivery” nodes will be also created and linked with the “Ordering” node with relationships.

### **b)**

#### **Inserting Data**

The queries used to create the nodes and relationships are shown below.

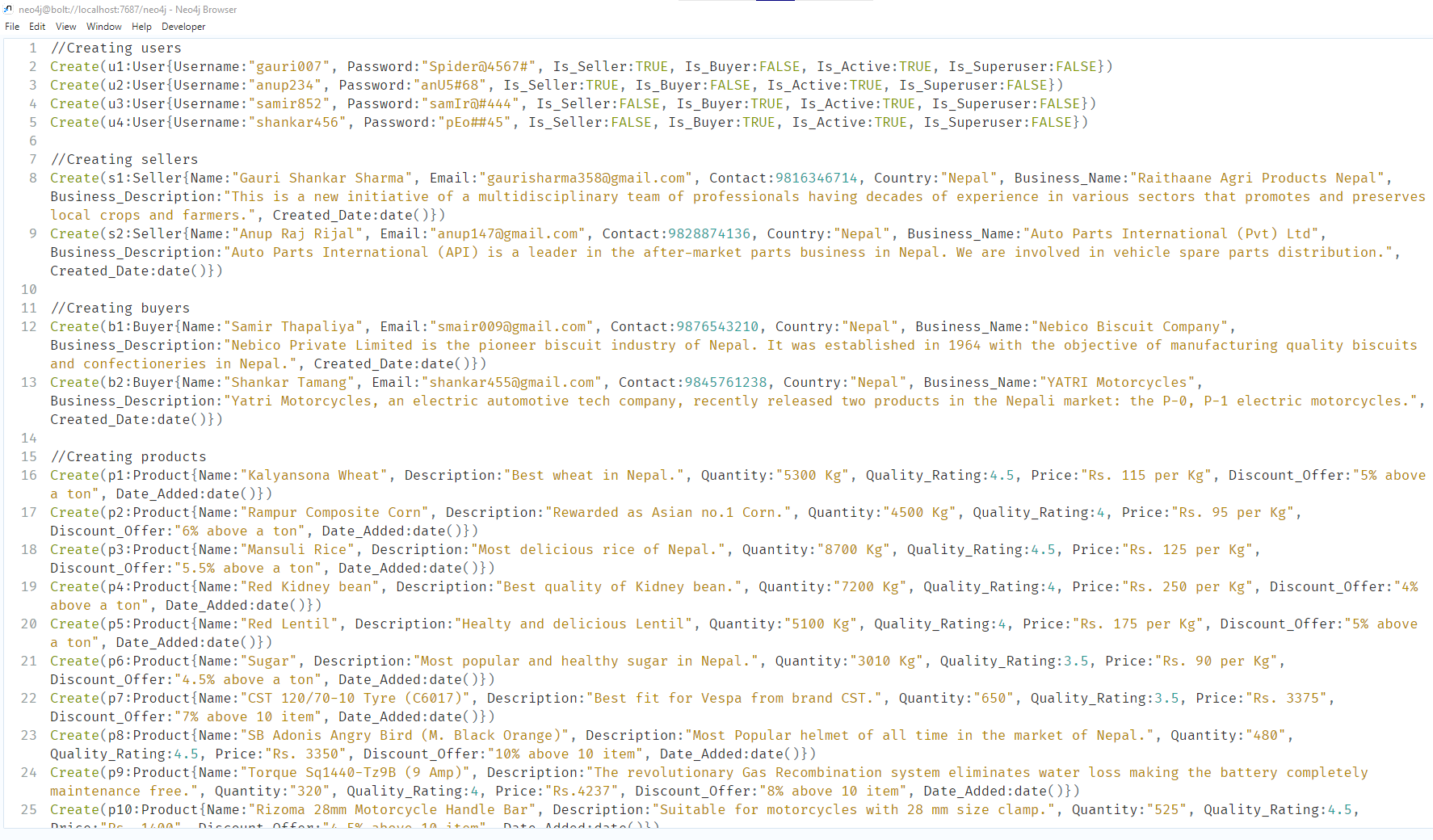


Figure 2: Queries for Inserting data



Figure 3: Queries for Inserting data



Figure 4: Queries for Inserting data

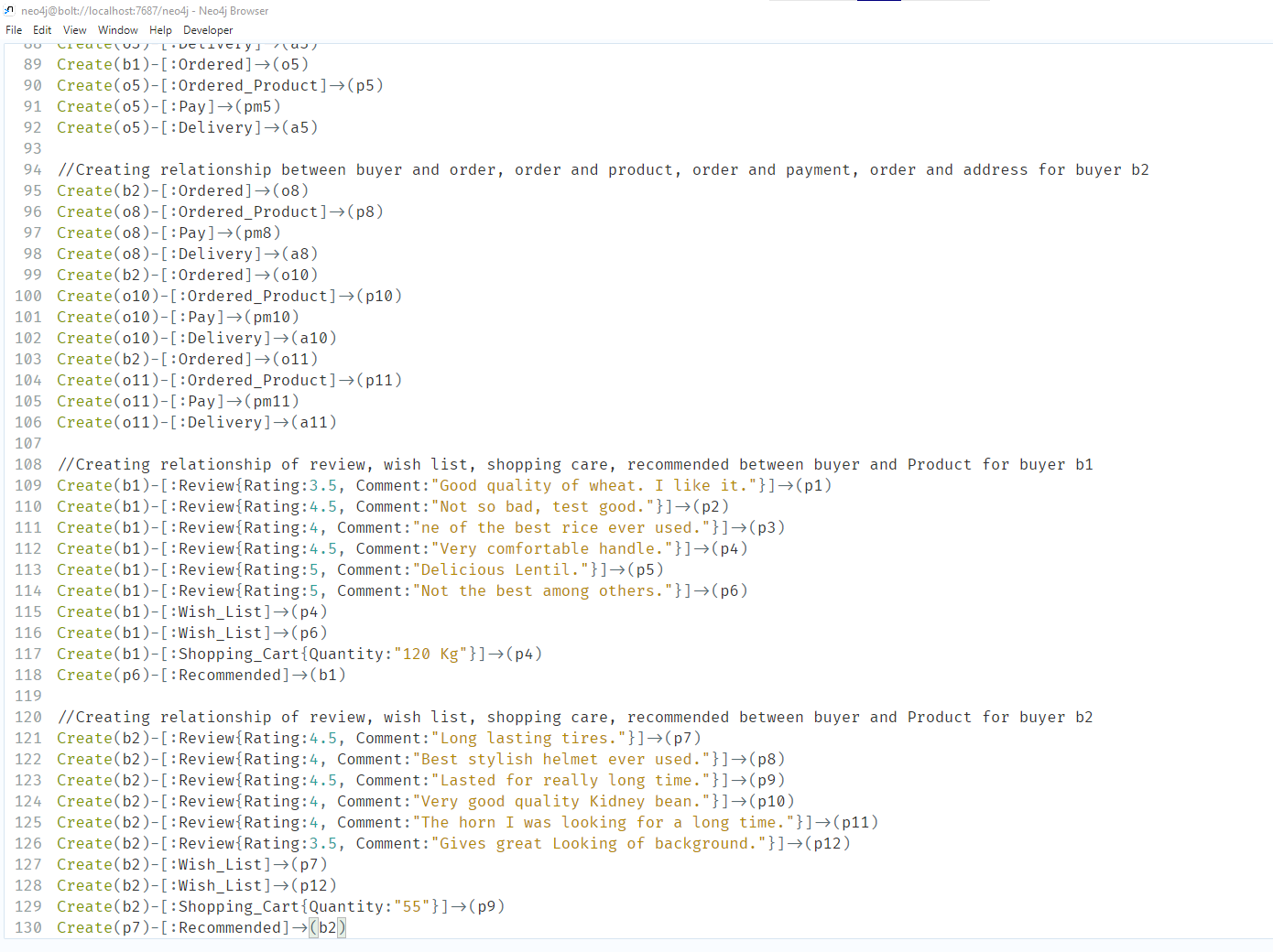


Figure 5: Queries for Inserting data



Figure 6: Inserting Data

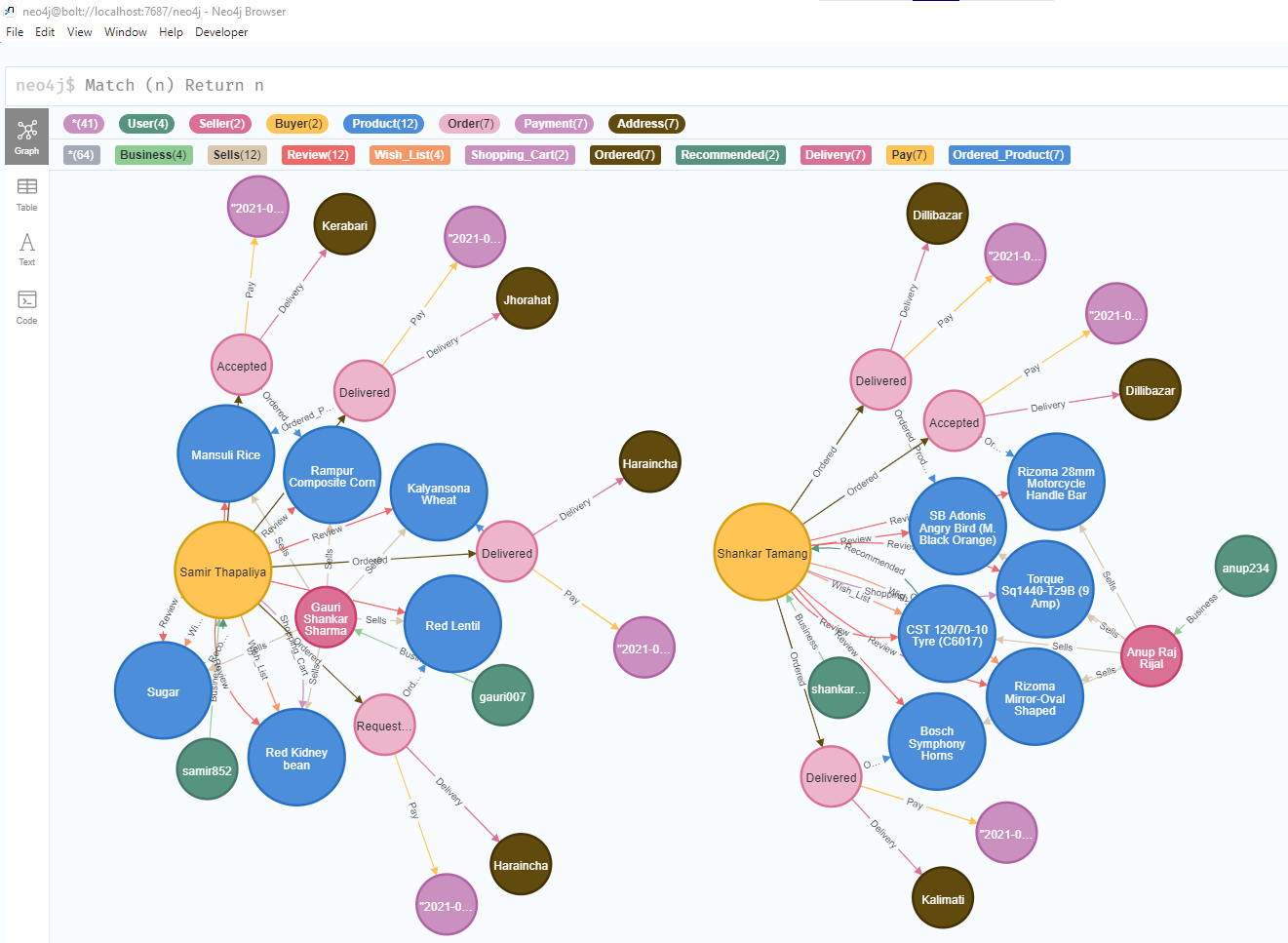


Figure 7: Final Output after inserting the data

#### **Queries that might happen**

1) Getting all the products with the sellers and order status ordered by a buyer.

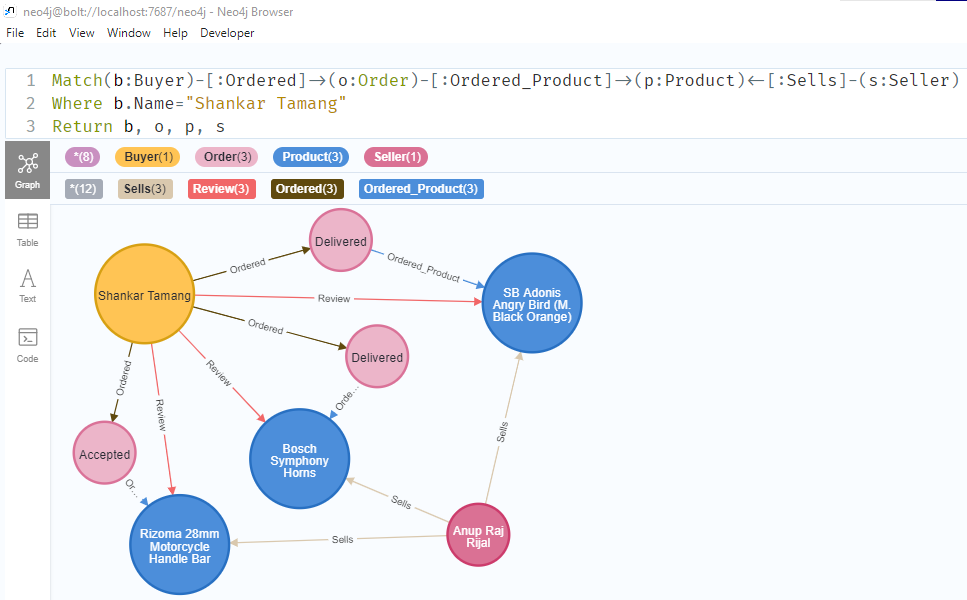
****

Figure 8: Products ordered by a buyer

2)Total number of orders come in the sellers

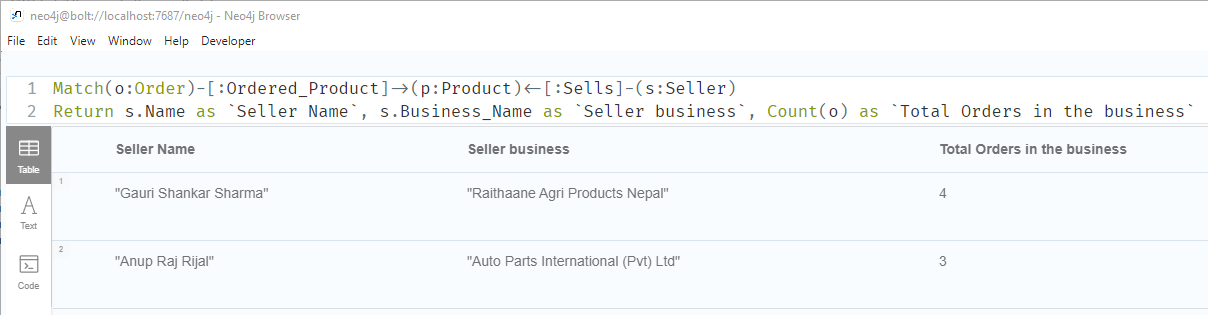


Figure 9: Total orders in each seller

3)Products recommended to the buyers

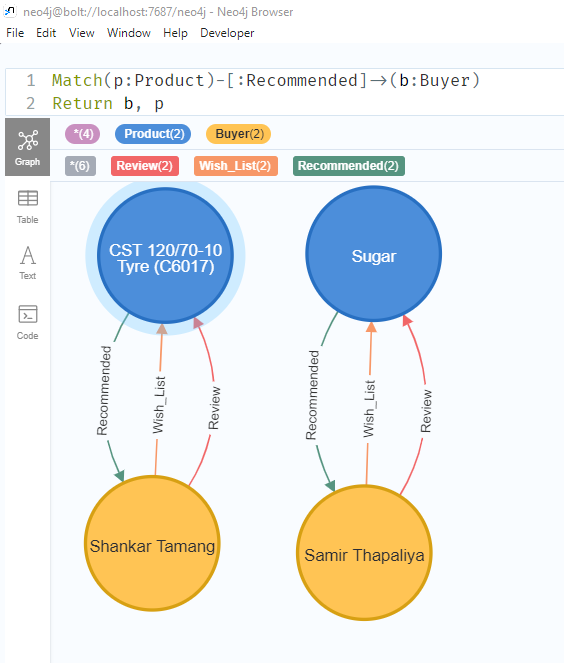


Figure 10: Products recommended to the buyers

4)Products that the buyers had wish-listed

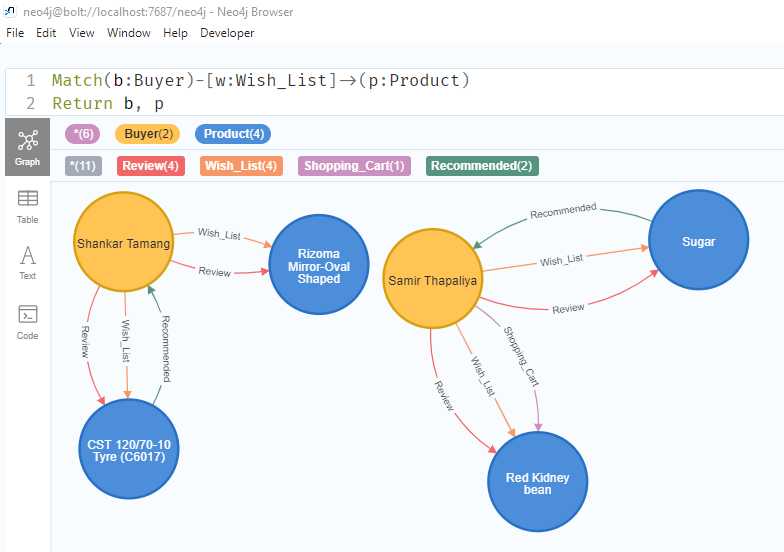
****

Figure 11: Products wish-listed by the buyers

5)The buyer who has ordered the highest number of products

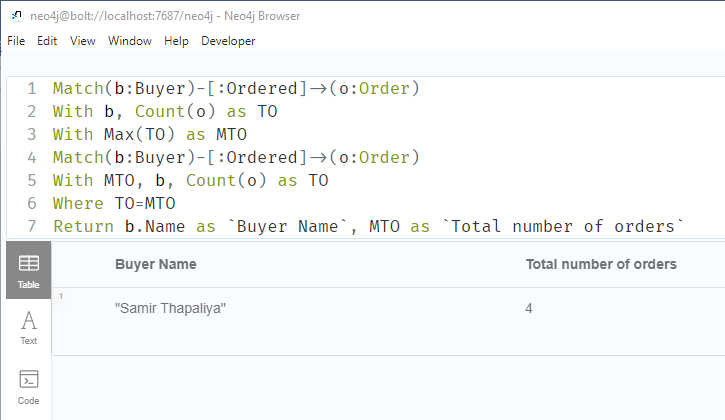
****

Figure 12:The buyer ordering the highest number of products

6) The Products that have the highest number of ratings with their sellers

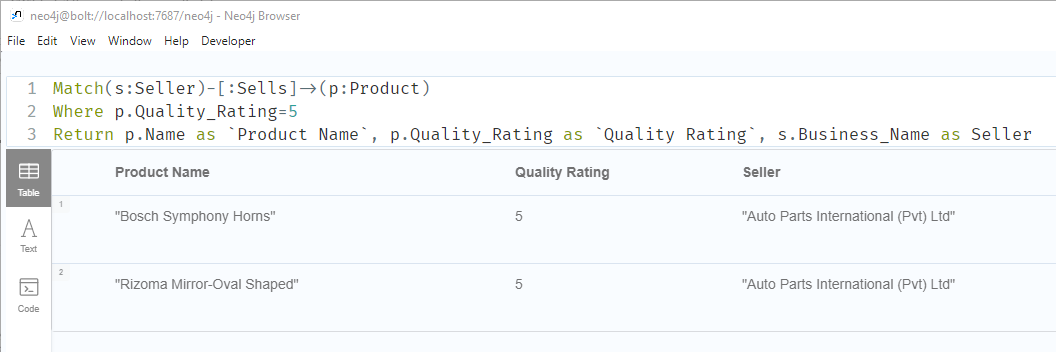
****

Figure 13: The products with the highest number of ratings

7)Total number of sellers and buyers on the B2B website

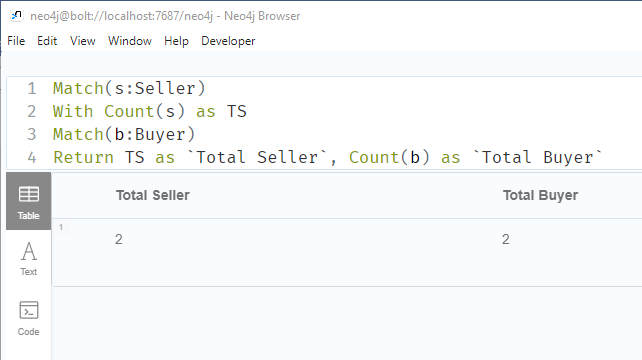
****

Figure 14: Total number of buyers and sellers

8)Payment methods used by a buyer while ordering products

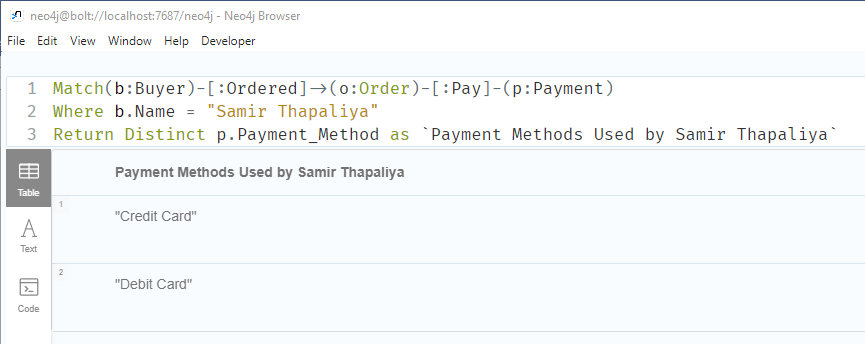


Figure 15: Payment method used by a buyer while ordering products