Assignment 4

The assignment uses the **UCI Online Retail Dataset** (https://archive.ics.uci.edu/dataset/352/online+retail) , a comprehensive transactional dataset containing over 540,000 real e-commerce transactions from a UK-based online retailer between December 2010 and December 2011. The dataset structure is as follows:

InvoiceNo: 6-digit transaction identifier (prefix 'C' indicates cancellations)

StockCode: 5-digit product identifier

Description: Product name/description

Quantity: Number of items per transaction

InvoiceDate: Transaction timestamp

UnitPrice: Product price in Sterling

CustomerID: 5-digit customer identifier

Country: Customer's country of residence

1. Design and implement a 2nd normalized relational database and insert the dataset into database using SQL (atleast 1000 records). (5 Marks)
2. Design and implement MongoDB's document-oriented approach (implement both transaction and customer centric approaches) and use PyMongo with connection pooling and error handling. (4+4+2 = 10 Marks)
3. Perform CRUD operations on both approaches and **write the differences you have observed regarding the performance and provide proof for your claims.** (5 Marks)
4. Use MongoDB Atlas local deployment with specific cluster configuration (either transaction or customer centric). (5 Marks)

Upload the Q3 answer as a PDF and the rest of the code needs to be committed on your Git repo.