MSDS 7330

File Organization and Database Management Final Exam Due 23 April 2019

Name: Javier Saldana

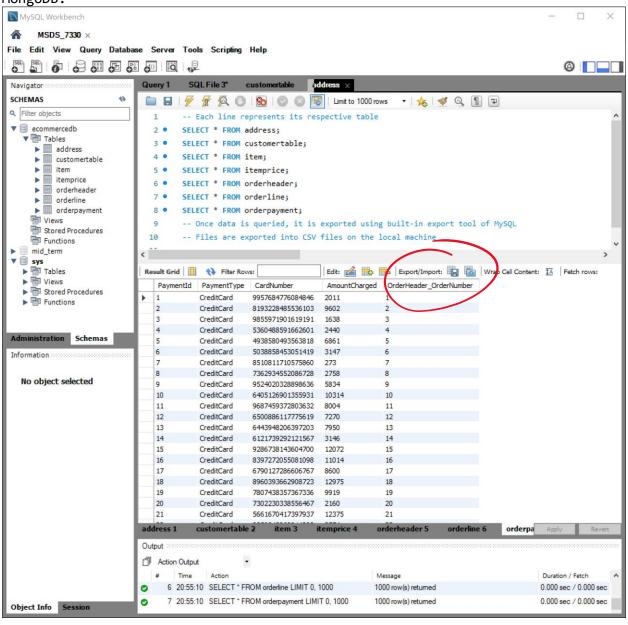
Final Exam: Directions

This is a final exam for MSDS 7330, File Organization and Database Management. This document contains the questions for the exam. Follow the question directions and submit the requested information by 23 April 2019.

Final Exam: Questions

Use MongoDB as a database to store the eCommerce Database that we used in class. Use Python to connect to the MongoDB database and insert, update, and query data from MongoDB. Submit the code and screenshots of your execution.

Run the following SQL to view the individual tables. Once the queries are run, the data is then exported into CSV files. Each table is exported into its own respective CSV file in order to facilitate the data import to MongoDB.



Initiate shell and create database in MongoDB

```
Microsoft Windows [Version_10.0.17134.706]
(c) 2018 Microsoft Corporation. All rights reserved.
C:\Users\javie>"C:\Program Files\MongoDB\Server\4.0\bin\mongo.exe"
MongoDB shell version v4.0.6
connecting to: mongodb://127.0.0.1:27017/?gssapiServiceName=mongodb
Implicit session: session { "id" : UUID("4b9c6baf-09e8-4505-ac00-01b55cd4ad6a") }
MongoDB server version: 4.0.6
Server has startup warnings:
2019-04-11T20:18:43.652-0500 I CONTROL [initandlisten]
2019-04-11T20:18:43.652-0500 I CONTROL [initandlisten] ** WARNING: Access control is not enabled
for the database.
Read and write access to data
and configuration is unrestricted.
2019-04-11T20:18:43.652-0500 I CONTROL [initandlisten]
Enable MongoDB's free cloud-based monitoring service, which will then receive and display
metrics about your deployment (disk utilization, CPU, operation statistics, etc).
The monitoring data will be available on a MongoDB website with a unique URL accessible to you
and anyone you share the URL with. MongoDB may use this information to make product
improvements and to suggest MongoDB products and deployment options to you.
To enable free monitoring, run the following command: db.enableFreeMonitoring()
To permanently disable this reminder, run the following command: db.disableFreeMonitoring()
> use ecommercedb
switched to db ecommercedb
```

Create the collections for each table export from MySQL

```
> db.createCollection("Address")
{ "ok" : 1 }
> 
> db.createCollection("CustomerTable")
{ "ok" : 1 }
> 
> db.createCollection("Item")
{ "ok" : 1 }
> 
> db.createCollection("ItemPrice")
{ "ok" : 1 }
> 
> db.createCollection("OrderHeader")
{ "ok" : 1 }
> 
> db.createCollection("OrderLine")
{ "ok" : 1 }
> 
> db.createCollection("OrderPayment")
{ "ok" : 1 }
```

Import tables into collections

```
C:\Users\javie>"C:\Program Files\MongoDB\Server\4.0\bin\mongoimport.exe" --db ecommercedb --type
csv --collection Address --drop --headerline --file "C:\Users\javie\OneDrive\Documents\Academics\
SMU\MSDS 7330 - File Organization and Database Management\Final\Address.csv"
                                connected to: localhost
2019-04-15T21:36:43.113-0500
2019-04-15T21:36:43.120-0500
                                dropping: ecommercedb.Address
2019-04-15T21:36:44.206-0500
                                imported 100 documents
C:\Users\javie>"C:\Program Files\MongoDB\Server\4.0\bin\mongoimport.exe" --db ecommercedb --type
csv --collection CustomerTable --drop --headerline --file "C:\Users\javie\OneDrive\Documents\Acad
emics\SMU\MSDS 7330 - File Organization and Database Management\Final\CustomerTable.csv'
2019-04-15T21:58:04.731-0500
                                connected to: localhost
2019-04-15T21:58:04.743-0500
                                dropping: ecommercedb.CustomerTable
                                imported 100 documents
2019-04-15T21:58:05.810-0500
C:\Users\javie>"C:\Program Files\MongoDB\Server\4.0\bin\mongoimport.exe" --db ecommercedb --type
<u>csv --collection Item --drop --headerline --file "C:\Users\javie\OneDrive\Documents\Academics\SMU</u>
\MSDS 7330 - File Organization and Database Management\Final\Item.csv"
2019-04-15T21:58:47.629-0500
                                connected to: localhost
2019-04-15T21:58:47.633-0500
                                dropping: ecommercedb.Item
2019-04-15T21:58:48.711-0500
                                imported 100 documents
C:\Users\javie>"C:\Program Files\MongoDB\Server\4.0\bin\mongoimport.exe" --db ecommercedb --type
csv --collection ItemPrice --drop --headerline --file "C:\Users\javie\OneDrive\Documents\Academic
s\SMU\MSDS 7330 - File Organization and Database Management\Final\ItemPrice.csv"
2019-04-15T21:59:15.056-0500
                               connected to: localhost
2019-04-15T21:59:15.059-0500
                                dropping: ecommercedb.ItemPrice
2019-04-15T21:59:16.132-0500
                                imported 100 documents
C:\Users\javie>"C:\Program Files\MongoDB\Server\4.0\bin\mongoimport.exe" --db ecommercedb --type
csv --collection OrderHeader --drop --headerline --file "C:\Users\javie\OneDrive\Documents\Academ
ics\SMU\MSDS 7330 - File Organization and Database Management\Final\OrderHeader.csv"
2019-04-15T21:59:49.227-0500
                                connected to: localhost
                                dropping: ecommercedb.OrderHeader
2019-04-15T21:59:49.230-0500
2019-04-15T21:59:50.320-0500
                                imported 1000 documents
C:\Users\javie>"C:\Program Files\MongoDB\Server\4.0\bin\mongoimport.exe" --db ecommercedb --type
csv --collection OrderLine --drop --headerline --file "C:\Users\javie\OneDrive\Documents\Academic
s\SMU\MSDS 7330 - File Organization and Database Management\Final\OrderLine.csv"
                                connected to: localhost
2019-04-15T22:00:17.225-0500
2019-04-15T22:00:17.229-0500
                                dropping: ecommercedb.OrderLine
                                imported 1000 documents
2019-04-15T22:00:18.336-0500
C:\Users\javie>"C:\Program Files\MongoDB\Server\4.0\bin\mongoimport.exe" --db ecommercedb --type
csv --collection OrderPayment --drop --headerline --file "C:\Users\javie\OneDrive\Documents\Acade
mics\SMU\MSDS 7330 - File Organization and Database Management\Final\OrderPayment.csv"
2019-04-15T22:01:33.371-0500
                               connected to: localhost
2019-04-15T22:01:33.374-0500
                                dropping: ecommercedb.OrderPayment
                                imported 1000 documents
2019-04-15T22:01:34.483-0500
```

Query Data from ecommercedb Database in MongoDB for each document

```
"_id": ObjectId("Scb544bdd781fc84082e5bb1), "CustomerIs"
"Address_AddressID": 51 }
"ad': ObjectId("Scb544bdd781fc84082e5bb2"), "CustomerID": 21, "CustomerName"
"Address_AddressID": 64 }
"_id": ObjectId("Scb544bdd781fc84082e5bb3"), "CustomerID": 22, "CustomerName"
, "Address_AddressID": 40 }
"_id": ObjectId("Scb544bdd781fc84082e5bb3"), "CustomerID": 23, "CustomerName"
ddress_AddressID": 83 }
"_id": ObjectId("Scb544bdd781fc84082e5bb5"), "CustomerID": 24, "CustomerName"
"_id": ObjectId("Scb544bdd781fc84082e5bb5"), "CustomerID": 24, "CustomerName"
                                                                                                                                                                                                                      AddressID" : 48 }
jectId("Scb544bdd781fc84082e5bb7"), "CustomerID" : 12, "CustomerName" : "Bill Rang
    ss").
"Stadesdadesd781fc8d882e5c6f"), "ItemNumber" : 1, "ItemDescription" : "Descript
"ItemNeight" : "286 lbs", "CountryOfOrigin" : "US", "CostPrice" : 67, "ItemCatego
       .
Id("Scb544e8d781fc84082e5c10"), "ItemNumber" : 3, "ItemDescription" : "Descripti
ItemWeight" : "754 lbs", "CountryOfOrigin" : "US", "CostPrice" : 66, "ItemCateg
     s" }
tid("scb54de8d781fc84082e5c11"), "ItemNumber" : 4, "ItemDescription" : "Descripti
"ItemNeight" : "546 lbs", "CountryOfOrigin" : "US", "CostPrice" : 87, "ItemCatego
                                                                                                                                                                                                   .ltemprice.*InO(())
(d" : ObjectId"'scb45494d781fc84882e5c77"), "PriceId" : 1, "SalePrice" : 296, "Country'
Item_ItemNumber" : 1)
(d" : ObjectId''scb45496d781fc84882e5c78"), "PriceId" : 2, "SalePrice" : 313, "Country'
nics")
ectid("stp64de8d78ifc8d982s5c18"), "ItemNumben" : 7, "ItemSescription" : "Descriptio
, "ItemWeight" : "319 lbs", "CountryOfOrigin" : "US", "CostPrice" : 75, "ItemCategor
                                                                                                                                                                                                               [temNumber" : 2 }
ObjectId("5cb54504d781fc84082e5c79"), "PriceId" : 3, "SalePrice" : 499, "Country
            )
"Gscb544e8d78ifc84682e5c15"), "ItemNumber" : 8, "ItemDescription" : "Descriptio
emWeight" : "348 lbs", "CountryOfOrigin" : "US", "CostPrice" : 75, "ItemCategor
      ctd("5cb544e8d781fc84082e5c16"), "ItemNumber" : 9, "ItemDescription" : "Descriptio
'ItemWeight" : "325 lbs", "CountryOfOrigin" : "US", "CostPrice" : 60, "ItemCategor
     s" }
tid("Scb54de8d781fc8d682e5c17"), "ItemNumber" : 12, "ItemDescription" : "Descripti
, "ItemNeight" : "594 lbs", "CountryOfOrigin" : "US", "CostPrice" : 18, "ItemCateg
            ("5cb544e8d781fc84082e5c19"), "ItemNumber" : 11, "ItemDescription" : "Descripti
(temWeight" : "257 lbs", "CountryOfOrigin" : "US", "CostPrice" : 25, "ItemCateg
                                                                                                                                                                                                               objecting scopessuad/81fc84082e5c81"), "PriceId" : 10, "SalePrice" : 200, "Country"
[TemNumber" : 10 ]
ObjectId("5cb54504d781fc84082e5c82"), "PriceId" : 12, "SalePrice" : 375, "Country"
[TemNumber" : 12 ]
              .)
"scb54Ae8d781fc84082e5cia"), "ItemNumber" : 13, "ItemOescription" : "Oescripti
TemWeight" : "703 lbs", "CountryOfOrigin" : "US", "CostPrice" : 19, "ItemCateg
                emMelgnt: /69.00,
Scb544e8d781fc84082e5c1b"), "ItemNumber": 14, "ItemDescription": "Descripti
emMelght": "253 lbs", "CountryOfOrigin": "US", "CostPrice": 82, "ItemCategS
                                                                                                                                                                                                                                         : 12 }
cb54504d781fc84082e5c83"), "PriceId" : 13, "SalePrice" : 387, "Country"
               )
schbidde8d781fc8d682e5c1c"), "ItemNumber" : 15, "ItemDescription" : "Descripti
temWeight" : "805 lbs", "CountryOfOrigin" : "US", "CostPrice" : 80, "ItemCateg
            {
("5cb544e8d781fc84082e5cie"), "ItemNumber" : 17, "ItemDescription" : "Descriptis
[temWeight" : "865 lbs", "CountryOfOrigin" : "US", "CostPrice" : 11, "ItemCateg[
              S
"Scb544e8d781fc84082e5c1f"), "ItemNumber" : 18, "ItemDescription" : "Descripti
temWeight" : "846 lbs", "CountryOfOrigin" : "US", "CostPrice" : 59, "ItemCategS
     " }
"Id("Scb544e8d781fc84082e5c20"), "ItemNumber" : 22, "ItemDescription" : "Descriptis",
, "ItemMeight" : "461 lbs", "CountryOfOrigin" : "US", "CostPrice" : 94, "ItemCategType
```

```
rderHeader.find(())
": ObjectId("Scb54526d781fc84082eSce8"), "OrderNumber": 1, "OrderDate": "2017-12-22 00:0
"CustomerTable_CustomerID": 96, "OrderType": "Sales", "Address_AddressID": 48 }
": ObjectId("Scb54526d781fc84082eSce"), "OrderNumber": 2, "OrderDate": "2017-02-04 00:0
"CustomerTable_CustomerID": 44, "OrderType": "Sales", "Address_AddressID": 29 0:0
"CustomerTable_CustomerID": 29, "OrderType": "Sales", "Address_AddressID": 29 0:0
"CustomerTable_CustomerID": 29, "OrderType": "Sales", "Address_AddressID": 76 }
"CustomerTable_CustomerID": 29, "OrderType": "Sales", "Address_AddressID": 78 0:0
"CustomerTable_CustomerID": 56, "OrderType": "Sales", "Address_AddressID": 29 0:0
"CustomerTable_CustomerID": 69, "OrderType": "Sales", "Address_AddressID": 29 0:0
"CustomerTable_CustomerID": 69, "OrderType": "Sales", "Address_AddressID": 29 1:0
"SolectId("Scb54526d781fc40802eSce8"), "OrderNumber": 7, "OrderDate": "2017-03-09 00:0
"CustomerTable_CustomerID": 83, "OrderType": "Sales", "Address_AddressID": 29 1:0
"SolectId("Scb54526d781fc40802eSce8"), "OrderNumber": 4, "OrderDate": "2017-03-08 00:0
"CustomerTable_CustomerID": 83, "OrderType": "Sales", "Address_AddressID": 38 9:0
"CustomerTable_CustomerID": 83, "OrderType": Sales", "Address_AddressID": 38 9:0
"CustomerTable_CustomerID": 80, "OrderType": "Sales", "Address_AddressID": 70 9:0
"CustomerTable_CustomerID": 80, "OrderType": "Sales", "Address_AddressID": 18 9:0
"CustomerTable_CustomerID": 18, "OrderTy
           d: 0bjectId("Scb8d35dd781fc8d892s6f2"), OrderNumber" : 26, UnderNumte : .017.12% or .
      id" : ObjectId("Scb5458ed781fc84082e64bb"), "PaymentId" : 9, "PaymentType" : "CreditCard",
Number" : NumberLong("9524020328898636"), "AmountCharged" : 5834, "OrderHeader_OrderNumber"
     _id" : ObjectId("5cb5458ed/81fc84082e64bc"), "PaymentId" : 10, "PaymentType" : "CreditCard",
dNumber" : NumberLong("6405126901355931"), "AmountCharged" : 10314, "OrderHeader_OrderNumber
                     : ObjectId("5cb5458ed781fc84082e64bd"), "PaymentId" : 11, "PaymentType" : "CreditCard"
ber" : NumberLong("9687459372803632"), "AmountCharged" : 8004, "OrderHeader_OrderNumber
      }
id" : ObjectId("Scb5458ed781fc84082e64be"), "PaymentId" : 12, "PaymentType" : "CreditCard"
|Number" : NumberLong("6500886117775619"), "AmountCharged" : 7270, "OrderHeader_OrderHumber
     }
_id" : ObjectId("5cb5458ed781fc84082e64bf"), "PaymentId" : 13, "PaymentType" : "CreditCard"
dNumber : NumberLong("6443948206397203"), "AmountCharged" : 7950, "OrderHeader_OrderNumber
       }
id" : ObjectId("5cb5458ed781fc84082e64c8"), "PaymentId" : 14, "PaymentType" : "CreditCard"
Number" : NumberLong("6121739292121567"), "AmountCharged" : 3146, "OrderHeader_OrderNumber
      }
id": ObjectId("Scb5458ed781fc84082e64c1"), "PaymentId": 15, "PaymentType": "CreditCard"
iNumber": NumberLong("9286738143604700"), "AmountCharged": 12072, "OrderHeader_OrderNumber
 15 )
__dd" . ObjectId("Scb54S8ed781fc84082e64c2"), "PaymentId" : 16, "PaymentType" : "CreditCard"
rdNumber" : NumberLong("8397272055081098"), "AnountCharged" : 11014, "OrderHeader_OrderNumber
  [6 ]
__dd : ObjectId("Scb5458ed/78Ifc84082e66c3"), "PaymentId" : 17, "PaymentType" : "CreditCard"
"dNumber" : NumberLong("6790127286666767"), "AmountCharged" : 8600, "OrderHeader_OrderNumber
     .ld": ObjectId("5cb5458ed781fc84082e64c4"), "PaymentId": 4, "PaymentType": "CreditCard",
Number": NumberLong("5360488591662601"), "AmountCharged": 2440, "OrderHeader_OrderNumber"
      id" : ObjectId("5cb5458ed781fc84082e64c5"), "PaymentId" : 18, "PaymentType" : "CreditCard"
Mumber" : NumberLong("8960393662908723"), "AmountCharged" : 12975, "OrderHeader_OrderNumber
                   : ObjectId("5cb5458ed781fc84082e64c6"), "PaymentId" : 19, "PaymentType" : "CreditCard",
ber" : NumberLong("7807438357367336"), "AmountCharged" : 9919, "OrderHeader_OrderNumber
       }
id" : ObjectId("5cb5458ed781fc84082e64c7"), "PaymentId" : 20, "PaymentType" : "CreditCard"
Number" : NumberLong("7302230338556467"), "AmountCharged" : 2160, "OrderHeader_OrderHumber
     ]
id": ObjectId("5cb5458ed781fc84082e64c8"), "PaymentId": 22, "PaymentType": "CreditCard"
dNumber": NumberLong("8350343060914202"), "AmountCharged": 2574, "OrderHeader_OrderNumber
     }
_id" : ObjectId("5cb5458ed781fc84082e64c9"), "PaymentId" : 21, "PaymentType" : "CreditCard",
JNumber" : NumberLong("5661670417397937"), "AmountCharged" : 12375, "OrderHeader_OrderNumber
                     : ObjectId("5cb5458ed781fc84082e64ca"), "PaymentId" : 24, "PaymentType" : "CreditCard"
ber" : NumberLong("8853925379632267"), "AmountCharged" : 12594, "OrderHeader_OrderNumber
24 }
".id" : ObjectId("ScbS4S8ed781fc84082e6dcb"), "PaymentId" : 23, "PaymentType" : "CreditCard"
rdNumber" : NumberLong("9964758940487483"), "AmountChanged" : 2882, "OrderHeader_OrderNumber
       }
id" : ObjectId("Scb5458ed781fc84082e64cc"), "PaymentId" : 25, "PaymentType" : "CreditCard"
Number" : NumberLong("5797463328299096"), "AmountCharged" : 6958, "OrderHeader_OrderNumber
```

```
> db.OrderLine.find({)}
( ".id': ObjectId("5c654542/8781fc84882e60cd"), "OrderLineNumber": 1, "Oundrivelader_OrderNumber: 1, "Quantity": 2, "PricePerQty': 323, "Item_ItemNumber": 8 }
( ".id': ObjectId("5c654542/8781fc84882e60cd"), "OrderLineNumber": 2, "OrderHeader_OrderNumber: 1, "Quantity": 3, "PricePerQty': 455, "Item_ItemNumber": 3, "OrderHeader_OrderNumber: 1, "Quantity": 5, "PricePerQty': 459, "Item_ItemNumber": 3, "OrderHeader_OrderNumber: 2, "Quantity": 5, "PricePerQty': 123, "Item_ItemNumber": 3, "OrderHeader_OrderNumber: 2, "Quantity": 18, "PricePerQty': 138, "Item_ItemNumber": 3, "OrderHeader_OrderNumber: 2, "Quantity": 18, "PricePerQty': 138, "Item_ItemNumber": 4, "OrderHeader_OrderNumber: 2, "Quantity": 18, "PricePerQty': 318, "Item_ItemNumber": 5, "OrderHeader_OrderNumber: 2, "Quantity": 5, "PricePerQty': 318, "Item_ItemNumber: 3, "OrderHeader_OrderNumber: 2, "Quantity": 6, "PricePerQty': 317, "Item_ItemNumber: 3, "OrderHeader_OrderNumber: 2, "Quantity": 6, "PricePerQty': 37, "Item_ItemNumber: 3, "OrderHeader_OrderNumber: 2, "Quantity": 6, "PricePerQty': 37, "Item_ItemNumber: 3, "OrderHeader_OrderNumber: 2, "Quantity": 5, "PricePerQty': 37, "Item_ItemNumber: 3, "OrderHeader_OrderNumber: 4, "OrderHeader_OrderNumber: 4, "Quantity": 5, "PricePerQty': 37, "Item_ItemNumber: 5, "OrderHeader_OrderNumber: 4, "Quantity": 1, "PricePerQty": 244, "Item_ItemNumber: 5, "Quantity": 1, "PricePerQty": 274, "Item_ItemNumber: 1, "OrderHeader_OrderNumber: 4, "Quantity": 1, "PricePerQty": 274, "Item_ItemNumber: 1, "OrderHeader_OrderNumber: 5, "Quantity": 1, "PricePerQty": 374, "Item_ItemNumber: 1, "OrderHeader_OrderNumber: 1, "OrderHeader_OrderNumber: 5, "Quantity": 1, "PricePerQty": 274, "Item_ItemNumber: 1, "OrderHeader_OrderNumber: 1, "OrderHeader_OrderNumber
```

In 1-2 pages, compare and contrast the suitability of MySQL and MongoDB for the eCommerce database. Be sure to identify both advantages and disadvantages of each database.

The eCommerce data is extremely structured and perfectly suited for a relational database (MySQL). The advantages of MySQL are that the relationships are preserved and clearly defined. It is easy to see which table has a relationship to which table. The relationships thus make for querying the data much simpler. In addition to the clear relationships, MySQL allows for line item views across multiple tables. This allows the user to query the data and perform operations for analytical/mathematical purposes. Furthermore, the joins across the data also allow for easier queries which don't require sophisticated SQL.

However, MySQL does have its limitations. Based on the current design, there are some serious constraints in scalability. Currently, the database doesn't allow for differentiation between physical, billing, mailing, and shipping address among other critical aspects. While this is acceptable based on its current state (assuming that is how the business owner wants it), it will pose problems as the business expands. Should the company decide to allow for products to have reviews attached to the products and have those reviews linked to user/customer profiles will add a complexity to the database it is not currently designed to handle. Surely, tables may be added and new relations may be created but as the business grows the complexities of the relationship begin to grow much more complex and sophisticated, which have the ability to make the database extremely messy. The problem with having many relationships across the database mean that a query would have to read in every single table in order to get to the end/designed table/tuple. As a result, MySQL meets the current needs of the database, but it will outgrown its current design eventually as typical ecommerce businesses mature into its desired level.

Scalability is a strong point for MongoDB. Unlike MySQL, MongoDB uses documents to store the tables without preserving the relational constraints. This allows the owner to scale the database by simply adding new "documents" to the database in MongoDB. This enhances the database's versatility but also allows the business to grow horizontally or vertically without having database constraints. The flexibility of MongoDB helps the database also perform at a much more efficient level than MySQL. Whereas traditional MySQL queries read in every join they perform, documents allow the user to bypass the unnecessary data by simply joining data from the desired documents. However, the eCommerce database is considerably small performance isn't at issue at its current state. There are only a handful of tables which may require joins, which essentially translate into nanoseconds in performance difference.

As efficient as MongoDB is and its promise for the future scale, there are downsides to the holding eCommerce in MongoDB is its complexity to perform joins. Considering the current data was imported into different collections using the tables as collections, data redundancy was minimized. However, this unstructured design doesn't allow for standard joins in the data. While manual joins are possible in MongoDB between collections, they must be manually coded into the query and could become complex when attempting to join multiple documents from various collections. As the code becomes sophisticated, the efficiency of MongoDB is sacrificed. Furthermore, as the database continues to grows, we would essentially end up seeing a lot of redundancy in order to preserve the ability to join in the future. While redundancy isn't

necessarily a deal-breaker, the data consumption as the database grows is something that must be considered. Storage space is valuable and expensive and excessive unstructured data means a company must purchase/lease physical space in order to store the data. This is echoed by the fact that a document is limited to 16MB in size. While under its present state the size limit doesn't pose a problem, as the database grows (especially unstructured data) then there is a greater chance of documents getting maximized much quicker. This would only add to the complexity when attempting to join collections/documents into the query.

Ultimately, as the eCommerce database currently stands it is better suited under MySQL. Relationships are preserved and queries are just as effective as they would be in a MongoDB database. The limitations of MySQL don't hinder the current state of affairs at this point, which mean that MySQL would suffice for the eCommerce database. However, if the business begins to grow and collect additional complex data then it would worth considering a migration over to MongoDB. It is important for the owner to recognize the direction in which they want to take the database and prepare for how its going suit its business needs.