

# Customer Retention and Sales Optimization in Retail

## Part 1 - Data Preparation, Excel Analytics, SQL Analytics & Excel Dashboarding

### 1. Data Preparation

The screenshot shows the Microsoft Excel interface. A data table named 'transactions' is displayed in the center, with columns: transaction\_id, customer\_id, product\_id, quantity, discount\_rate, and total\_price. The table contains 17 rows of data. To the right, the 'Queries & Connections' pane is visible, showing a list of queries and their row counts: products (100 rows loaded), campaigns (10 rows loaded), customers (300 rows loaded), transactions (1,500 rows loaded), top 10 customers by revenue (10 rows loaded), and 1 vs repeat customer.

transaction_id	customer_id	product_id	quantity	discount_rate	total_price
a5504658-4fb0-44b0-9fad-c8584f3bc...	CUST0230	PROD056	3	0.15	1.15
fa7c5413-70e2-455e-a379-450d6cd0b...	CUST0156	PROD072	5	0	5.00
262b8df6-8816-4677-b790-44277cc0c...	CUST0050	PROD097	2	0.05	1.90
66dcd04-c330-4427-a924-64dacef9c...	CUST0135	PROD084	1	0.05	1.00
2ba251e7-8735-4d88-b533-8a28...	CUST0122	PROD022	5	0	5.00
958b2f8b-c90-4bdc-8f86-ab15c...	CUST0080	PROD000	4	0.15	5.60
09761b97-315e-45af-99c7-0226b79d6...	CUST0240	PROD037	1	0.05	1.00
fdcafee7-69b5-474a-9e2a-f4460...	CUST0147	PROD090	3	0.15	2.55
b6db7359-461d-48e0-b8cc-5b4e3265...	CUST0036	PROD087	2	0.1	1.80

The screenshot shows the MySQL Workbench interface. The SQL editor contains the following queries:

```
1 CREATE DATABASE retail_capstone;
2 USE retail_capstone;
3
4 CREATE TABLE customer_data (
5   customer_id VARCHAR(100) PRIMARY KEY,
6   name VARCHAR(100),
7   gender VARCHAR(10),
8   age INT,
9   region VARCHAR(100),
10  sign_up_date DATE,
11  loyalty_score INT
12 );
13
14 CREATE TABLE marketing_campaigns (
15   campaign_id VARCHAR(100) PRIMARY KEY,
16   channel VARCHAR(20),
17   start_date DATE,
18   end_date DATE
19 );
```

The output pane at the bottom shows an error message: "Error Code: 1007. Can't create database 'retail\_capstone': database exists".

The data has been cleaned using power query in the excel and then imported into SQL. A database is created and containing four tables of main dataset.

## 2. Excel Analytics

transactions					
descriptive_stats_quantity		descriptive_stats_discount_rate		descriptive_stats_total_price	
Mean	3.092667	Mean	0.0765	Mean	773.0681
Standard Error	0.037275	Standard Error	0.001442	Standard Error	15.62401
Median	3	Median	0.1	Median	638.26
Mode	5	Mode	0.15	Mode	1266.2
Standard Deviation	1.443654	Standard Deviation	0.055841	Standard Deviation	605.1155
Sample Variance	2.084136	Sample Variance	0.003118	Sample Variance	366164.7
Kurtosis	-1.34264	Kurtosis	-1.35598	Kurtosis	-0.31858
Skewness	-0.09296	Skewness	-0.03504	Skewness	0.771735
Range	4	Range	0.15	Range	2458.14
Minimum	1	Minimum	0	Minimum	8.81
Maximum	5	Maximum	0.15	Maximum	2466.95
Sum	4639	Sum	114.75	Sum	1159602
Count	1500	Count	1500	Count	1500

customers				products	
descriptive_stats_age		descriptive_stats_loyalty_score		descriptive_stats_price	
Mean	43.32	Mean	50.44	Mean	265.8916
Standard Error	0.884887	Standard Error	1.693906	Standard Error	15.52898
Median	43.5	Median	49	Median	271.615
Mode	56	Mode	33	Mode	#N/A
Standard Deviation	15.32669	Standard Deviation	29.33932	Standard Deviation	155.2898
Sample Variance	234.9073	Sample Variance	860.7957	Sample Variance	24114.92
Kurtosis	-1.17959	Kurtosis	-1.18286	Kurtosis	-1.28082
Skewness	-0.03838	Skewness	0.130822	Skewness	-0.14831
Range	51	Range	99	Range	486.16
Minimum	18	Minimum	1	Minimum	10.36
Maximum	69	Maximum	100	Maximum	496.52
Sum	12996	Sum	15132	Sum	26589.16
Count	300	Count	300	Count	100

## 3. SQL Analytics

### a) Top 10 customers by revenue

The screenshot shows the MySQL Workbench interface. The SQL editor contains the following query:

```
SELECT customer_id, SUM(total_price) AS total_spent
FROM sales_transactions
GROUP BY customer_id
ORDER BY total_spent DESC
LIMIT 10;
```

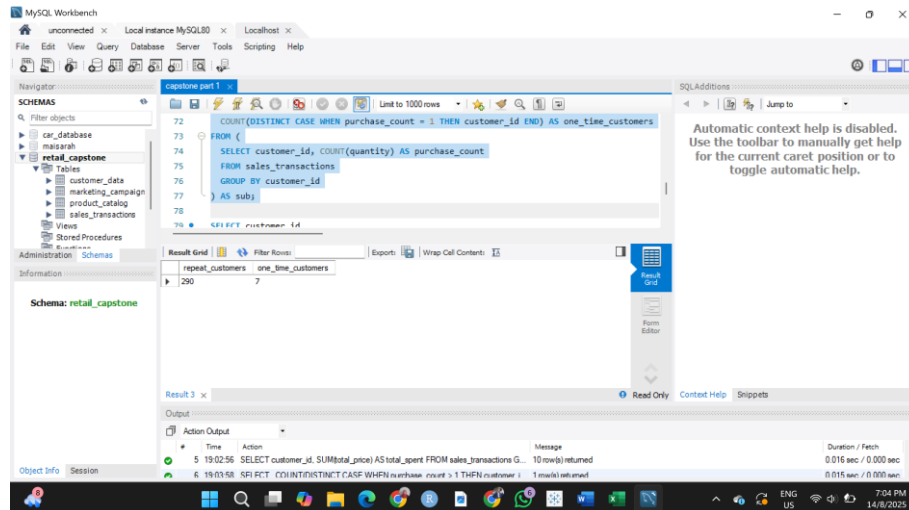
The Results grid displays the top 10 customers by total revenue:

customer_id	total_spent
CUST0085	13135.79
CUST0055	12836.38
CUST0154	12612.08
CUST0046	10680.16
CUST0079	10009.02
CUST0176	9160.16
CUST0232	9091.72
CUST0228	8610.32
CUST0184	8554.22

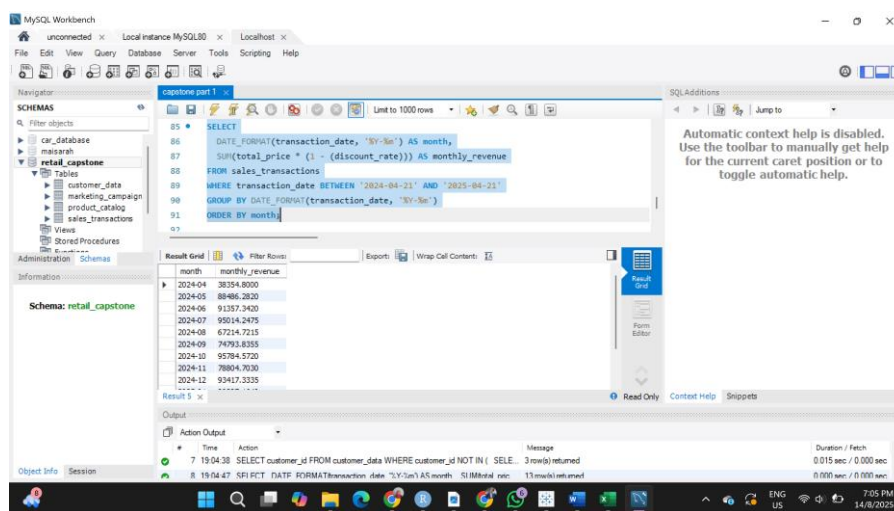
The Output tab shows the following error message:

```
4 19:02:47 ALTER TABLE sales_transactions ADD CONSTRAINT fk_customer FOREIGN KE... Error Code: 1826. Duplicate foreign key constraint name 'fk_customer'
```

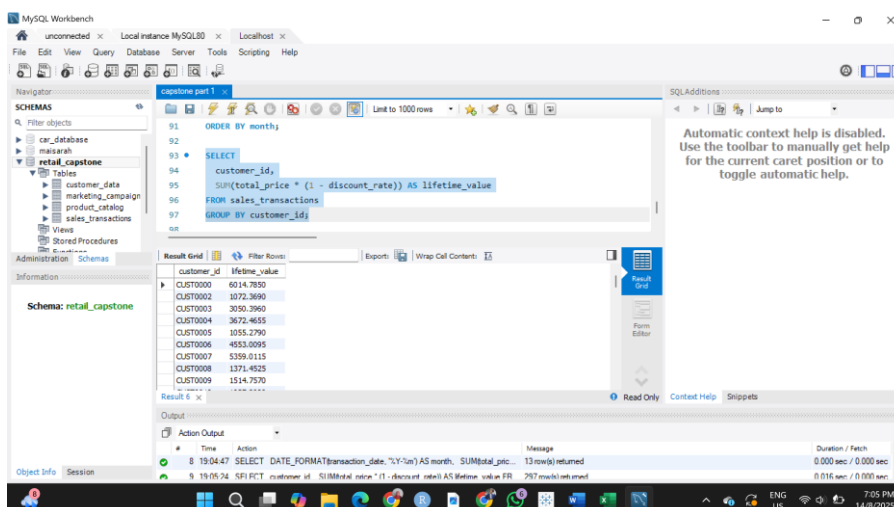
### b) Repeat vs one-time customers



### c) Monthly revenue growth

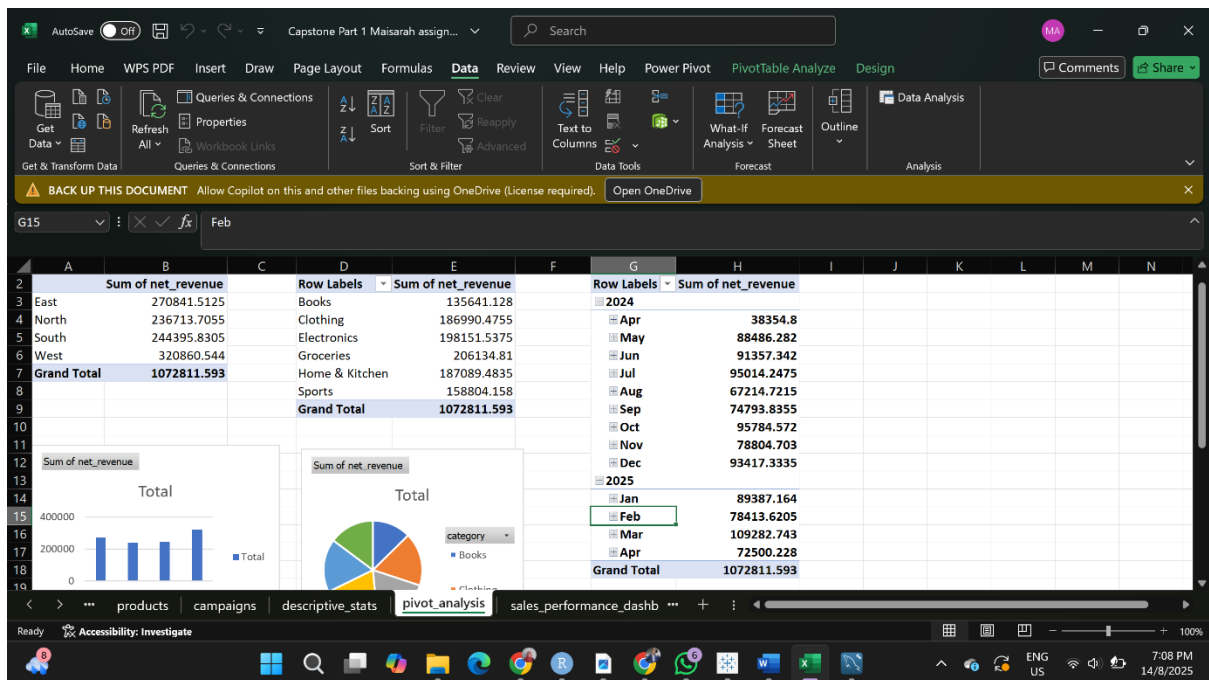


### d) Lifetime value



## 4. Excel Dashboarding

### a) Pivot analysis



### b) Dashboard

