MOP – Essential SQL Operations for Telecom Analytics

1. Objective

Students will learn how to:

- Retrieve, filter, and aggregate telecom usage data using SQL.
- Combine multiple tables (JOINs).
- Identify high-usage customers for targeted marketing or capacity planning.

---

2. Prerequisites

Knowledge:

- Basic understanding of databases and SQL syntax (SELECT, WHERE).
- Awareness of what "data usage," "customer_id," and "billing" mean in telecom context.

System Requirements:

- Laptop with Windows/macOS/Linux
- Internet connection

Software Setup:

Tool	Purpose	Installation Notes
MySQL	Database execution	Use local MySQL Workbench (portable, no admin rights needed).
Sample Dataset	Input data	Provided as `usage_data.csv` and `customers.csv`.

---

3. Files Required

Create or import the following simple datasets:

a) **`usage_data.csv`**

customer_id	data_used_in_GB	calls_made	revenue_inr	region
1001	5.2	25	180	Delhi
1002	12.5	40	280	Mumbai
1003	7.8	32	210	Chennai

customer_id	data_used_in_GB	calls_made	revenue_inr	region
1004	15.6	55	320	Delhi
1005	3.4	18	120	Kolkata

b) `customers.csv`

customer_id	name	plan_type	join_date
1001	Asha Mehta	Prepaid	2023-05-12
1002	Ravi Kumar	Postpaid	2022-12-20
1003	Sneha Rao	Prepaid	2023-01-18
1004	Manoj Singh	Postpaid	2021-11-05
1005	Divya Jain	Prepaid	2023-03-28

---

4. Procedure

Step 1 – Database Setup

Launch **MySQL Workbench**

Create a new database:

```
CREATE DATABASE testdb;
USE testdb;
```

Create & Delete two tables:

```
CREATE TABLE usage_data (
    customer_id INT,
    data_used_in_GB FLOAT,
    calls_made INT,
    revenue_inr FLOAT,region VARCHAR(20));

DROP TABLE usage_data;

Then again create table – USAGE_DATA using create command
```

Import or insert data manually:

```
INSERT INTO usage_data VALUES
(1001, 5.2, 25, 180, 'Delhi'),
(1002, 12.5, 40, 280, 'Mumbai'),
(1003, 7.8, 32, 210, 'Chennai'),
(1004, 15.6, 55, 320, 'Delhi'),
(1005, 3.4, 18, 120, 'Kolkata');
```

```
CREATE TABLE customers (
  customer_id INT,
  name VARCHAR(50),
  plan_type VARCHAR(20),
  join_date DATE
);

INSERT INTO customers VALUES
(1001, 'Asha Mehta', 'Prepaid', '20230512'),
(1002, 'Ravi Kumar', 'Postpaid', '20221220'),
(1003, 'Sneha Rao', 'Prepaid', '20230118'),
(1004, 'Manoj Singh', 'Postpaid', '20211105'),
(1005, 'Divya Jain', 'Prepaid', '20230328');
```

Step 2 – SELECT & WHERE Query

Goal: Filter customers exceeding 10 GB usage.

```
SELECT customer_id, data_used_in_GB
FROM usage_data
WHERE data_used_in_GB > 10
ORDER BY data_used_in_GB DESC;
```

Expected Output:

customer_id	data_used_in_GB
1004	15.6
1002	12.5

☑ **Learning Point:**
You filtered and sorted structured telecom usage data to identify **high-usage customers.**

Step 3 – JOIN Operation

Goal: Combine customer information with usage details.

```
SELECT c.name, c.plan_type, u.data_used_in_GB, u.region
FROM customers c
JOIN usage_data u ON c.customer_id = u.customer_id;
```

Expected Output:

name	plan_type	data_used_in_GB	region
Asha Mehta	Prepaid	5.2	Delhi
Ravi Kumar	Postpaid	12.5	Mumbai
Sneha Rao	Prepaid	7.8	Chennai
Manoj Singh	Postpaid	15.6	Delhi
Divya Jain	Prepaid	3.4	Kolkata

☑ **Learning Point:**
JOIN merges data from two tables → more complete analysis (e.g., who uses the most data by plan type).

Step 4 – GROUP BY Aggregations

Goal: Find **average data usage per region.**

```
SELECT region, AVG(data_used_in_GB) AS avg_usage
FROM usage_data
GROUP BY region;

SELECT region, ROUND(AVG(data_used_in_GB),2) AS avg_usage
FROM usage_data
GROUP BY region;
```

Expected Output:

region	avg_usage
Delhi	10.4
Mumbai	12.5
Chennai	7.8
Kolkata	3.4

☑ **Learning Point:**
Aggregation (GROUP BY) allows summarization — a critical skill for business intelligence.

EXERCISE – Practical Analysis

Business requirement: Identify which are the highest generating revenue plans and regions -> which we can target for premium upsells.

```
SELECT c.plan_type, u.region, AVG(u.revenue_inr) AS avg_revenue
FROM customers c
JOIN usage_data u ON c.customer_id = u.customer_id
GROUP BY u.region, c.plan_type
ORDER BY avg_revenue DESC;
```

---

5. Expected Outcomes

By the end of the exercise, students will be able to:

- Create, populate, and query telecom data tables.
- Use **SELECT, WHERE, JOIN, and GROUP BY** effectively.
- Derive actionable insights (e.g., identifying heavy users or regional revenue trends).
- Present query results as mini business insights.

---

6. Safety & Good Practices

- Double-check column names before running queries.
- Save your SQL scripts (.sql file) for submission.
- Never edit live production data (use local databases only).
