Implementation Internship – Assignment

Name: Himansu Sekhar B
Date: 9th November 2025
Tool Used: MySQL Workbench 8.0
Database Name: client_data_project
Table Name: customers

Assignment Tasks

1. SQL & Data Familiarity

1.1 What steps would you take to review this data before importing it into a system? Please explain the process in not more than 3 lines.

Answer:

Before importing, I would:

1. Check for missing or invalid values (especially in email, phone, and dates).
2. Verify consistent date formats (DD-MM-YYYY) and correct data types.
3. Remove duplicates and ensure column names match the database schema.

1.2 Please execute each of the queries and support your queries with a screenshot of your results.

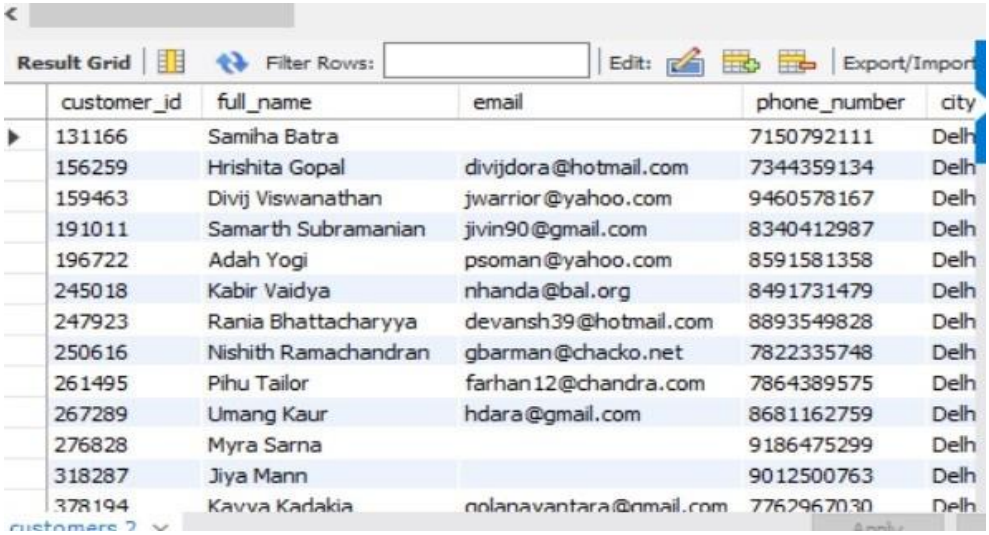
Write SQL queries for the following:

- **Write a query to display all customers from the city 'Delhi'.**

Query:-

```
SELECT * FROM customers
WHERE city = 'Delhi';
```

Output:-



Result Grid | Filter Rows: | Edit: | Export/Import

customer_id	full_name	email	phone_number	city
131166	Samiha Batra		7150792111	Delh
156259	Hrishita Gopal	divijdora@hotmail.com	7344359134	Delh
159463	Divij Viswanathan	jwarrior@yahoo.com	9460578167	Delh
191011	Samarth Subramanian	jivin90@gmail.com	8340412987	Delh
196722	Adah Yogi	psoman@yahoo.com	8591581358	Delh
245018	Kabir Vaidya	nhanda@bal.org	8491731479	Delh
247923	Rania Bhattacharyya	devansh39@hotmail.com	8893549828	Delh
250616	Nishith Ramachandran	gbarman@chacko.net	7822335748	Delh
261495	Pihu Tailor	farhan12@chandra.com	7864389575	Delh
267289	Umang Kaur	hdara@gmail.com	8681162759	Delh
276828	Myra Sarna		9186475299	Delh
318287	Jiya Mann		9012500763	Delh
378194	Kavya Kadakia	golanavantara@gmail.com	7762967030	Delh

customers 2

● Count the number of signups in the last 30 days. Assume today to be 16th April 2025

<u>Query:-</u>

```
18 •  SELECT COUNT(*) AS signups_last_30_days
19    FROM customers
20    WHERE STR_TO_DATE(signup_date, '%d-%m-%Y')
21          BETWEEN DATE_SUB('2025-04-16', INTERVAL 30 DAY) AND '2025-04-16';
22
```

Output:-

```
18 •  SELECT COUNT(*) AS signups_last_30_days
19    FROM customers
20    WHERE STR_TO_DATE(signup_date, '%d-%m-%Y')
21          BETWEEN DATE_SUB('2025-04-16', INTERVAL 30 DAY) AND '2025-04-16';
22
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

signups_last_30_days
85

Result 3 | Read Only

- **List unique cities where customers are based**

Query:-

```
22
23 •  SELECT DISTINCT city FROM customers;
24
```

Output:-

```
23 •  SELECT DISTINCT city FROM customers;
24
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

city
Chennai
Bangalore
Pune
Hyderabad
Kolkata
Delhi
Mumbai
Ahmedabad

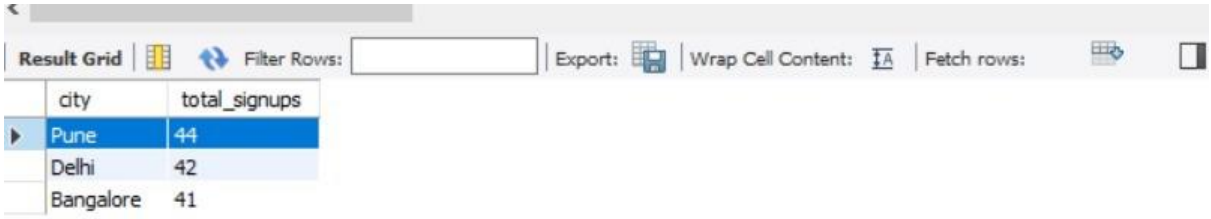
customers 4

● **List the top 3 cities by number of signups.**

Query:-

```
24
25 •  SELECT city, COUNT(*) AS total_signups
26    FROM customers
27    GROUP BY city
28    ORDER BY total_signups DESC
29    LIMIT 3;
```

Output:-



city	total_signups
Pune	44
Delhi	42
Bangalore	41

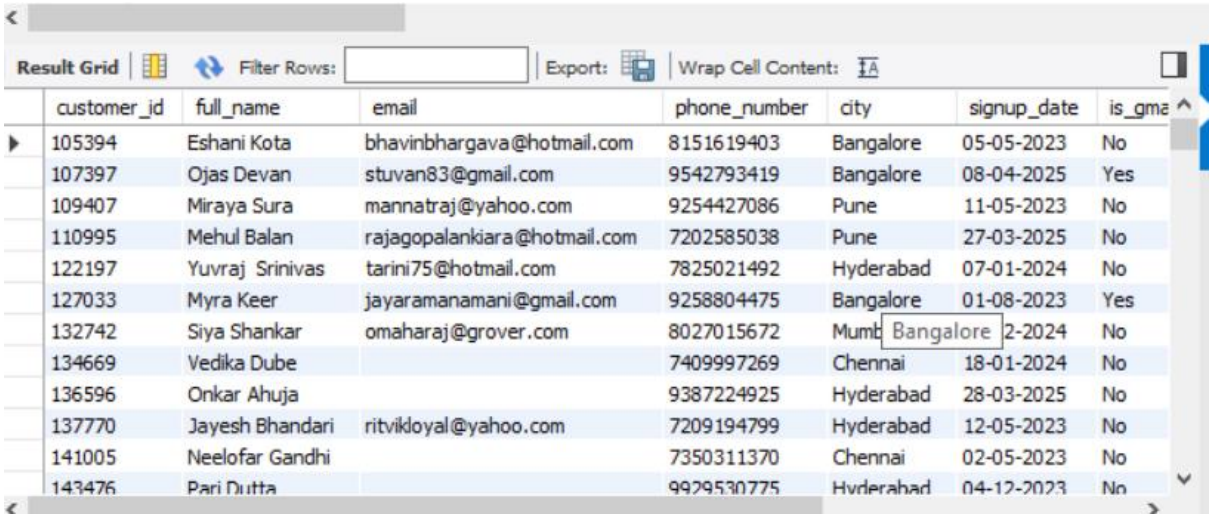
● **Assume there's another table orders (customer_id, order_id, amount). How would you find customers who have never placed an order?**

Query:-

```
31 • CREATE TABLE orders (
32       order_id INT PRIMARY KEY,
33       customer_id INT,
34       amount DECIMAL(10, 2)
35   );
36 • INSERT INTO orders (order_id, customer_id, amount)
37   VALUES
38   (1, 105156, 1200.50),
39   (2, 131166, 550.00),
40   (3, 245018, 999.99),
41   (4, 128097, 150.00);
```

```sql
SELECT c.*
FROM customers c
LEFT JOIN orders o ON c.customer_id = o.customer_id
WHERE o.customer_id IS NULL;
```

Output:-



Result Grid | Filter Rows: | Export: | Wrap Cell Content:

customer_id	full_name	email	phone_number	city	signup_date	is_gma
105394	Eshani Kota	bhavinbhargava@hotmail.com	8151619403	Bangalore	05-05-2023	No
107397	Ojas Devan	stuvan83@gmail.com	9542793419	Bangalore	08-04-2025	Yes
109407	Miraya Sura	mannatraj@yahoo.com	9254427086	Pune	11-05-2023	No
110995	Mehul Balan	rajagopalankiara@hotmail.com	7202585038	Pune	27-03-2025	No
122197	Yuvraj Srinivas	tarini75@hotmail.com	7825021492	Hyderabad	07-01-2024	No
127033	Myra Keer	jayaramanamani@gmail.com	9258804475	Bangalore	01-08-2023	Yes
132742	Siya Shankar	omaharaj@grover.com	8027015672	Mumb Bangalore	2-2024	No
134669	Vedika Dube		7409997269	Chennai	18-01-2024	No
136596	Onkar Ahuja		9387224925	Hyderabad	28-03-2025	No
137770	Jayesh Bhandari	ritvikloyal@yahoo.com	7209194799	Hyderabad	12-05-2023	No
141005	Neelofar Gandhi		7350311370	Chennai	02-05-2023	No
143476	Pari Dutta		9929530775	Hyderabad	04-12-2023	No

2. Data Transformation & Enrichment

The operations teams want a few more details to ensure smooth operations. Questions:

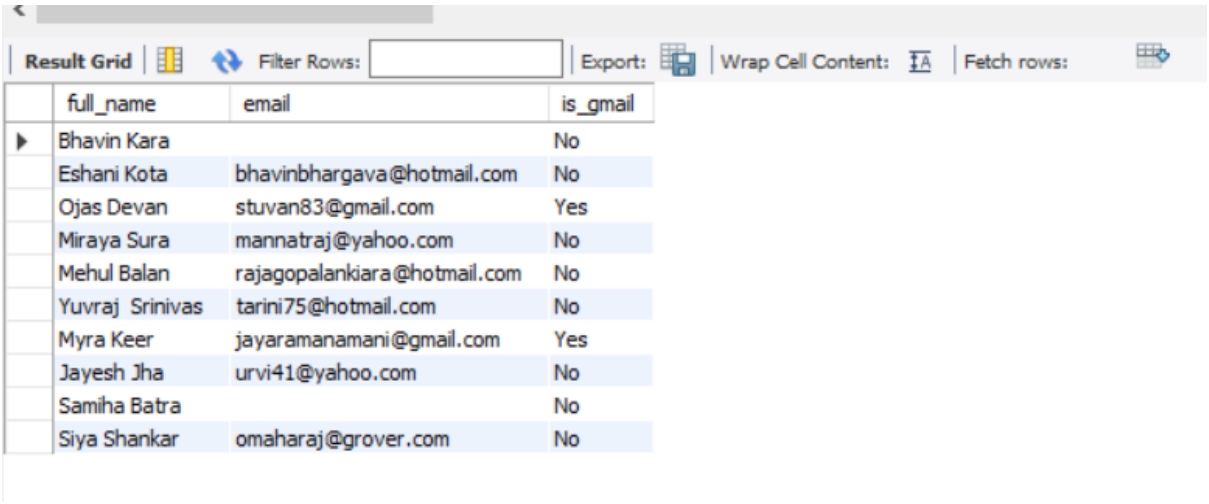
● Add a new column to show if the email domain is ‘gmail.com’ or not. Fill it with ‘Yes’ or‘No’.

Query:-

```sql
ALTER TABLE customers
ADD COLUMN is_gmail VARCHAR(3);
SET SQL_SAFE_UPDATES = 0;
UPDATE customers
SET is_gmail =
    CASE
        WHEN email LIKE '%@gmail.com' THEN 'Yes'
```

```sql
    CASE
        WHEN email LIKE '%@gmail.com' THEN 'Yes'
        ELSE 'No'
    END;
SELECT full_name, email, is_gmail
FROM customers
LIMIT 10;
```

Output:-



Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

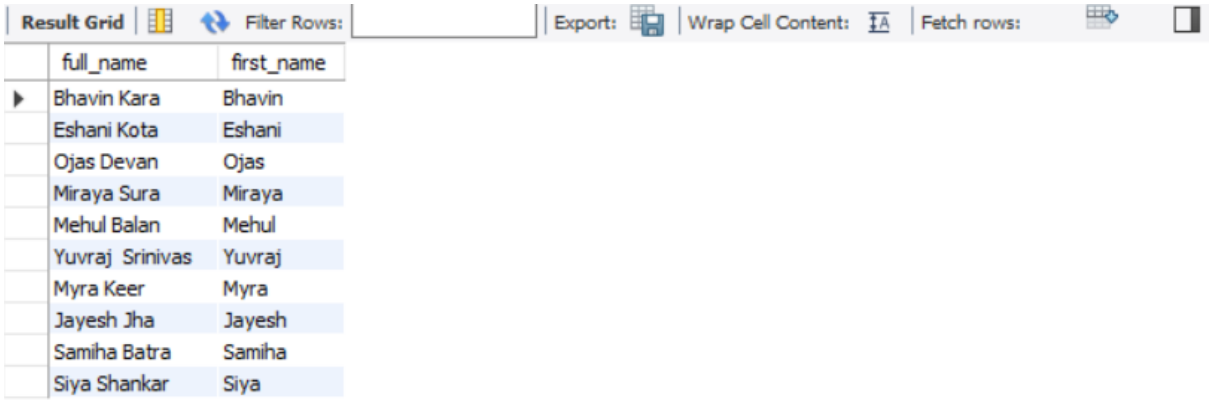
full_name	email	is_gmail
Bhavin Kara		No
Eshani Kota	bhavinbhargava@hotmail.com	No
Ojas Devan	stuvan83@gmail.com	Yes
Miraya Sura	mannatraj@yahoo.com	No
Mehul Balan	rajagopalankiara@hotmail.com	No
Yuvraj Srinivas	tarini75@hotmail.com	No
Myra Keer	jayaramanamani@gmail.com	Yes
Jayesh Jha	urvi41@yahoo.com	No
Samiha Batra		No
Siya Shankar	omaharaj@grover.com	No

- **Extract the first name from the name column and store it in ‘first_name’**

Query:-

```sql
ALTER TABLE customers ADD COLUMN first_name VARCHAR(100);
UPDATE customers
SET first_name = SUBSTRING_INDEX(full_name, ' ', 1);
SELECT full_name, first_name
FROM customers
LIMIT 10;
```

Output:-



Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

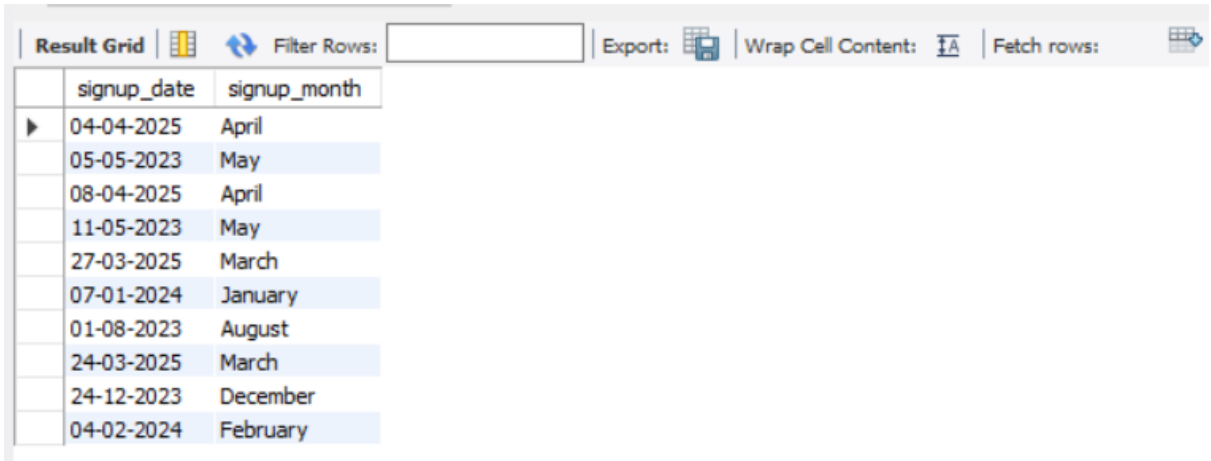
full_name	first_name
Bhavin Kara	Bhavin
Eshani Kota	Eshani
Ojas Devan	Ojas
Miraya Sura	Miraya
Mehul Balan	Mehul
Yuvraj Srinivas	Yuvraj
Myra Keer	Myra
Jayesh Jha	Jayesh
Samiha Batra	Samiha
Siya Shankar	Siya

● Add a column 'signup_month' to capture the month name the customer signed up.

Query:-

```
76 •  ALTER TABLE customers ADD COLUMN signup_month VARCHAR(20);
77 •  UPDATE customers
78    SET signup_month = MONTHNAME(STR_TO_DATE(signup_date, '%d-%m-%Y'));
79
```

Output:-



Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

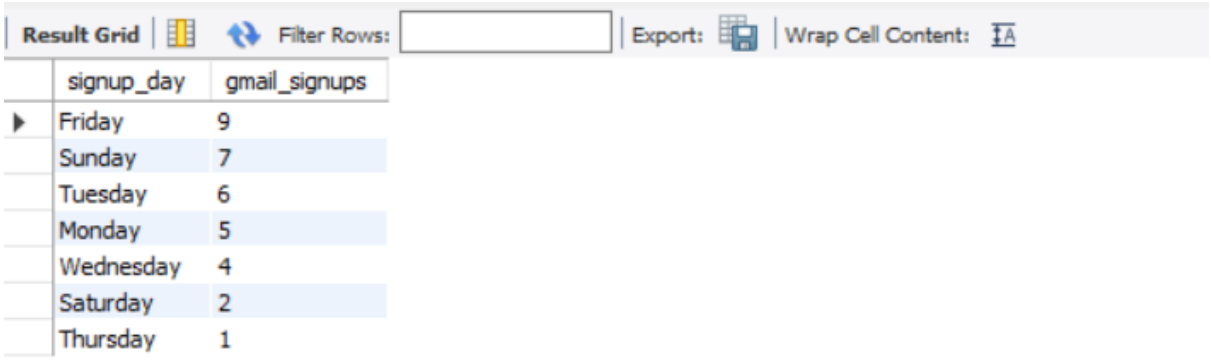
signup_date	signup_month
04-04-2025	April
05-05-2023	May
08-04-2025	April
11-05-2023	May
27-03-2025	March
07-01-2024	January
01-08-2023	August
24-03-2025	March
24-12-2023	December
04-02-2024	February

● **Create a report which shows the no. of GMAIL customers who have signed up for each day of the week.**

Query:-

```
94 •  SET sql_mode = '';
95 •  SELECT
96        DAYNAME(STR_TO_DATE(signup_date, '%d-%m-%Y')) AS signup_day,
97        COUNT(*) AS gmail_signups
98    FROM customers
99    WHERE email LIKE '%@gmail.com'
100   GROUP BY signup_day
```

Output:-



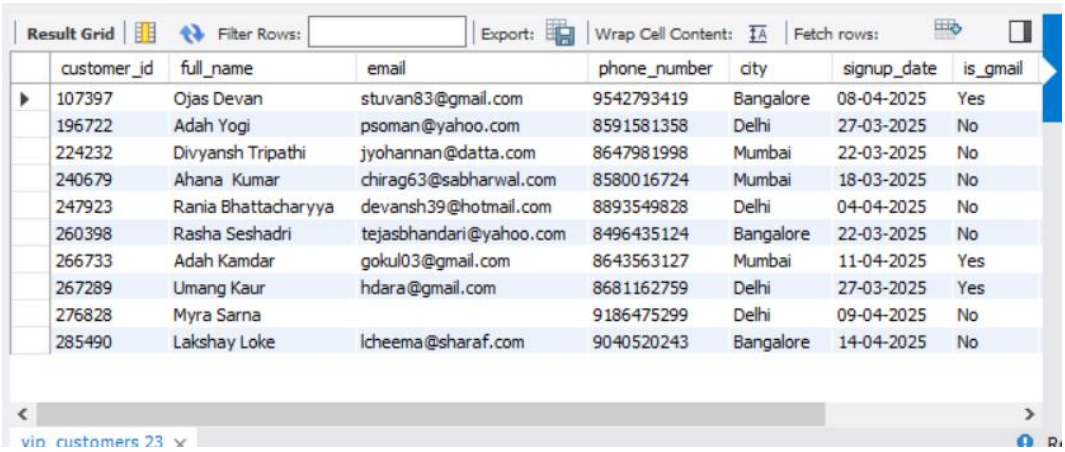
signup_day	gmail_signups
Friday	9
Sunday	7
Tuesday	6
Monday	5
Wednesday	4
Saturday	2
Thursday	1

● **Create a new table 'vip_customers' with customers from Delhi, Mumbai and Bangalore who have signed up in the last 60 days from 16th April 2025.**

Query:-

```
104 •  CREATE TABLE vip_customers AS
105    SELECT *
106    FROM customers
107    WHERE city IN ('Delhi', 'Mumbai', 'Bangalore')
108      AND STR_TO_DATE(signup_date, '%d-%m-%Y')
109          BETWEEN DATE_SUB('2025-04-16', INTERVAL 60 DAY) AND '2025-04-16';
```

Output:-



customer_id	full_name	email	phone_number	city	signup_date	is_gmail
107397	Ojas Devan	stuvan83@gmail.com	9542793419	Bangalore	08-04-2025	Yes
196722	Adah Yogi	psoman@yahoo.com	8591581358	Delhi	27-03-2025	No
224232	Divyansh Tripathi	jyohannan@datta.com	8647981998	Mumbai	22-03-2025	No
240679	Ahana Kumar	chirag63@sabharwal.com	8580016724	Mumbai	18-03-2025	No
247923	Rania Bhattacharyya	devansh39@hotmail.com	8893549828	Delhi	04-04-2025	No
260398	Rasha Seshadri	tejasbhandari@yahoo.com	8496435124	Bangalore	22-03-2025	No
266733	Adah Kamdar	gokul03@gmail.com	8643563127	Mumbai	11-04-2025	Yes
267289	Umang Kaur	hdara@gmail.com	8681162759	Delhi	27-03-2025	Yes
276828	Myra Sarna		9186475299	Delhi	09-04-2025	No
285490	Lakshay Loke	lcheema@sharaf.com	9040520243	Bangalore	14-04-2025	No

3. Analytics & Reporting

The business team wants insights from the customer data.

Questions:

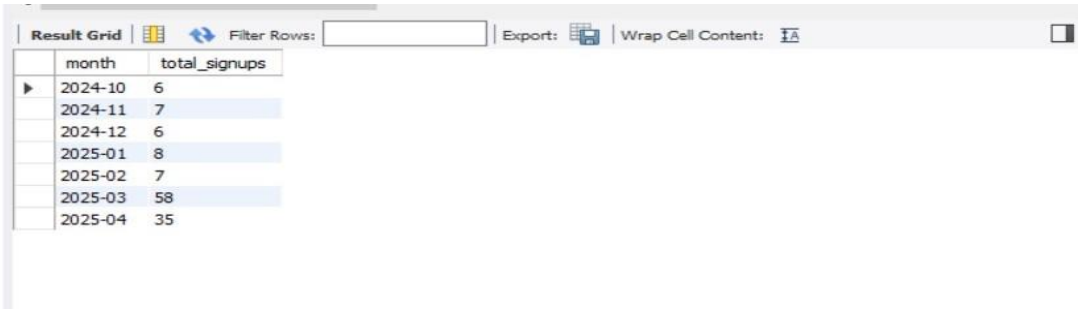
Write SQL queries for the following:

● Show a monthly signup count for the past 6 months.

Query:-

```
113 •   SELECT DATE_FORMAT(STR_TO_DATE(signup_date, '%d-%m-%Y'), '%Y-%m') AS month,
114             COUNT(*) AS total_signups
115     FROM customers
116     WHERE STR_TO_DATE(signup_date, '%d-%m-%Y') >= DATE_SUB('2025-04-16', INTERVAL 6 MONTH)
117     GROUP BY month
118     ORDER BY month;
```

Output:-



month	total_signups
2024-10	6
2024-11	7
2024-12	6
2025-01	8
2025-02	7
2025-03	58
2025-04	35

● **Get a list of cities with more than 20 customers.**

Query:-

```
120 •   SELECT city, COUNT(*) AS total_customers
121     FROM customers
122     GROUP BY city
123     HAVING total_customers > 20
124     ORDER BY total_customers DESC;
```

Output:-

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

city	total_customers
Pune	44
Delhi	42
Bangalore	41
Kolkata	41
Hyderabad	39
Ahmedabad	36
Chennai	33
Mumbai	24

● **Find the date with the highest number of signups.**

Query:-

```
126 •   SELECT signup_date, COUNT(*) AS signup_count
127     FROM customers
128     GROUP BY signup_date
129     ORDER BY signup_count DESC
130     LIMIT 1;
```

Output:-

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

signup_date	signup_count
28-03-2025	5

● **Add a new column to show the day of the signup date. Find the day with the highest number of signups.**

Query:-

```
133 •  ALTER TABLE customers ADD COLUMN signup_day VARCHAR(15);
134 •  UPDATE customers
135    SET signup_day = DAYNAME(STR_TO_DATE(signup_date, '%d-%m-%Y'));
136 •  SELECT signup_day, COUNT(*) AS total_signups
137    FROM customers
138    GROUP BY signup_day
139    ORDER BY total_signups DESC
```

Output:-

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

signup_day	total_signups
Friday	51