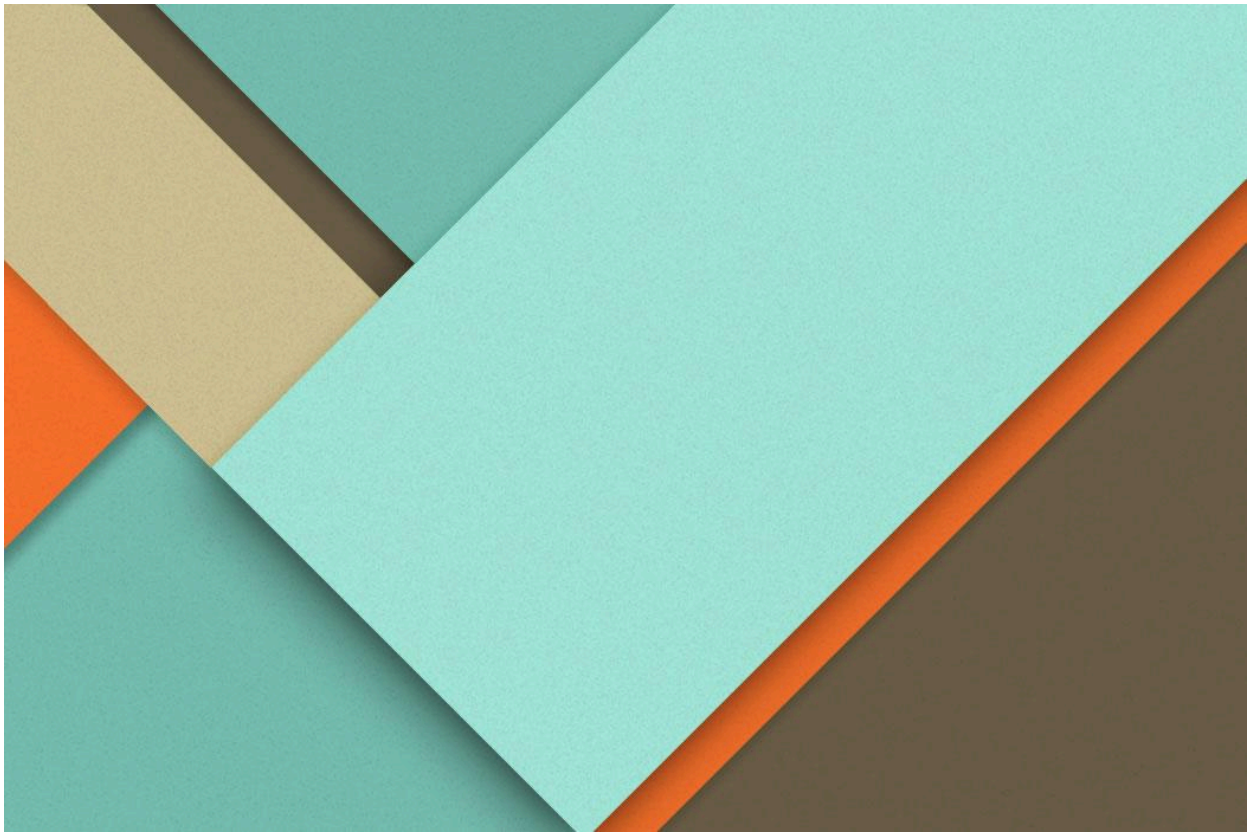

TASK 4

SQL for Data Analysis

27.06.2025

Saumya Singh

The original table :

```
USE sql_invoicing;
SELECT *
FROM clients;
```

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content:

client_id	name	address	city	state	phone
1	Vinte	3 Nevada Parkway	Syracuse	NY	315-252-7305
2	Myworks	34267 Glendale Parkway	Huntington	WV	304-659-1170
3	Yadel	096 Pawling Parkway	San Francisco	CA	415-144-6037
4	Kwideo	81674 Westerfield Circle	Waco	TX	254-750-0784
5	Topicloungе	0863 Farmco Road	Portland	OR	971-888-9129
NULL	NULL	NULL	NULL	NULL	NULL

a.Use SELECT, WHERE, ORDER BY, GROUP BY

```
USE sql_invoicing;

SELECT state, AVG(CHAR_LENGTH(name)) AS avg_name_length
FROM clients
WHERE state IS NOT NULL
GROUP BY state
```

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

state	avg_name_length
OR	11.0000
WV	7.0000

b.Use JOINS (INNER, LEFT, RIGHT)

```
USE sql_invoicing;

SELECT
    c.name AS client_name,
    i.number AS invoice_number,
    i.invoice_total,
    p.amount AS payment_amount,
```

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

client_name	invoice_number	invoice_total	payment_amount	payment_method
Myworks	91-953-3396	101.79	NULL	NULL
Topiclounge	03-898-6735	175.32	8.18	Credit Card
Topiclounge	20-228-0335	147.99	NULL	NULL
Topiclounge	41-666-1035	135.01	87.44	Credit Card
Topiclounge	52-269-9803	180.17	10.00	Cash
Topiclounge	52-269-9803	180.17	32.77	Credit Card
Topiclounge	77-593-0081	172.17	NULL	NULL
Topiclounge	87-052-3121	169.36	NULL	NULL
Vinte	10-451-8824	162.02	NULL	NULL
Vinte	48-266-1517	159.50	NULL	NULL
Vinte	75-587-6626	157.78	74.55	Credit Card
Vinte	78-145-1093	189.12	NULL	NULL
Vinte	83-559-4105	134.47	NULL	NULL
Yadel	20-848-0181	126.15	0.03	Credit Card
Yadel	33-615-4694	126.38	68.10	Credit Card
Yadel	55-105-9605	167.29	80.31	Credit Card
Yadel	56-934-0748	152.21	NULL	NULL
Yadel	68-093-9863	133.87	NULL	NULL

clients | Query 1 | SQL File 3* | SQL File 4*

```
USE sql_invoicing;

SELECT
    p.payment_id,
    c.name AS client_name,
    p.amount AS payment_amount
FROM payments p
RIGHT JOIN clients c ON p.client_id = c.client_id
ORDER BY p.payment_id;
```

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

payment_id	client_name	payment_amount
NULL	Myworks	NULL
NULL	Kwideo	NULL
1	Topiclounge	8.18
2	Vinte	74.55
3	Yadel	0.03
4	Topiclounge	87.44
5	Yadel	80.31
6	Yadel	68.10
7	Topiclounge	32.77
8	Topiclounge	10.00

c.Write subqueries

```sql
USE sql_invoicing;
-- Sunquery in where clause
SELECT name
FROM clients
WHERE client_id IN (
    SELECT client_id
    FROM invoices
    GROUP BY client_id
    HAVING COUNT(*) > 2
);
```

name
Vinte
Yadel
Topiclounge

```sql
-- Subquery in from clause
SELECT client_id, AVG(invoice_total) AS avg_invoice
FROM (
    SELECT client_id, invoice_total
    FROM invoices
) AS sub
GROUP BY client_id
HAVING COUNT(*) > 1;

-- Subquery in Select clause
SELECT
    name,
```

client_id	avg_invoice
1	160.578000
3	141.180000
5	163.336667

clients | Query 1 | SQL File 3* | SQL File 4* | SQL File 5*

Limit to 1000 rows

```sql
-- Subquery in Select clause
SELECT
    name,
    (SELECT SUM(amount)
     FROM payments
     WHERE payments.client_id = clients.client_id) AS total_paid
FROM clients;

--Correlated subquery
SELECT name
FROM clients
WHERE client id IN (
```

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

name	total_paid
Vinte	74.55
Myworks	NULL
Yadel	148.44
Kwideo	NULL
Topiclounge	138.39

d.Use aggregate functions (SUM, AVG)

clients | Query 1 | SQL File 3* | SQL File 4* | SQL File 6* | SQL File 7*

Limit to 1000 rows

```sql
FROM clients c
JOIN invoices i ON c.client_id = i.client_id
GROUP BY c.client_id;


SELECT
    AVG(invoice_total) AS avg_invoice_value
FROM invoices;
```

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

avg_invoice_value
152.388235

clients | Query 1 | SQL File 3* | SQL File 4* | SQL File 6* | SQL File 7*

Limit to 1000 rows

```sql
Use sql_invoicing;

SELECT
    c.name AS client_name,
    SUM(i.invoice_total) AS total_invoiced
FROM clients c
JOIN invoices i ON c.client_id = i.client_id
GROUP BY c.client_id;
```

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

client_name	total_invoiced
Myworks	101.79
Topiclounge	980.02
Yadel	705.90
Vinte	802.89

e.Create views for analysis

clients | Query 1 | SQL File 3* | SQL File 4* | SQL File 6* | SQL File 7* | SQL File 8*

Limit to 1000 rows

```sql
USE sql_invoicing;

-- creating view: Totalpayments per customer
CREATE  or REPLACE VIEW v2_client_payment_totals AS
SELECT
    c.client_id,
    c.name AS client_name,
    SUM(p.amount) AS total_paid,
    COUNT(p.payment_id) AS num_payments
FROM clients c
JOIN payments p ON c.client_id = p.client_id
GROUP BY c.client_id, c.name;
```

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

client_id	client_name	total_paid	num_payments
5	Topiclounge	138.39	4
1	Vinte	74.55	1
3	Yadel	148.44	3

clients | Query 1 | SQL File 3* | SQL File 4* | SQL File 6* | SQL File 7* | SQL File 8*

Limit to 1000 rows

```sql
-- View:  Total Invoice Amount for Each Client
CREATE VIEW v2_client_invoice_totals AS
SELECT
    c.client_id,
    c.name AS client_name,
    SUM(i.invoice_total) AS total_invoiced,
    COUNT(i.invoice_id) AS invoice_count
FROM clients c
JOIN invoices i ON c.client_id = i.client_id
GROUP BY c.client_id, c.name;

SELECT * FROM v2_client_invoice_totals;
```

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

client_id	client_name	total_invoiced	invoice_count
2	Myworks	101.79	1
5	Topiclounge	980.02	6
3	Yadel	705.90	5
1	Vinte	802.89	5

v2_client_payment_totals 2 | v2_client_invoice_totals 3 | v2_payments_by_method 4

Output

clients | Query 1 | SQL File 3* | SQL File 4* | SQL File 6* | SQL File 7* | SQL File 8*

Limit to 1000 rows

```sql
-- views: PAyments by payment methods

CREATE VIEW v2_payments_by_method AS
SELECT
    pm.name AS payment_method,
    COUNT(p.payment_id) AS number_of_payments,
    SUM(p.amount) AS total_collected
FROM payments p
JOIN payment_methods pm ON p.payment_method = pm.payment_method_id
GROUP BY p.payment_method;

SELECT * FROM v2_payments_by_method;
```

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

payment_method	number_of_payments	total_collected
Credit Card	7	351.38
Cash	1	10.00

v2_client_payment_totals 2 | v2_client_invoice_totals 3 | v2_payments_by_method 4

Output

f.Optimize queries with indexes

clients | Query 1 | SQL File 3* | SQL File 4* | SQL File 6* | SQL File 7* | SQL File 8* | SQL File 9* | SQL File 10*

Limit to 1000 rows

```sql
use sql_invoicing;
SHOW INDEXES FROM invoices;

-- Show all indexes on the payments table
SHOW INDEXES FROM payments;

-- Show all indexes on the clients table
SHOW INDEXES FROM clients;
```

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

Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Vis
invoices	0	PRIMARY	1	invoice_id	A	17	NULL	NULL		BTREE			YES
invoices	1	FK_client_id	1	client_id	A	4	NULL	NULL		BTREE			YES
invoices	1	idx_invoices_client_id	1	client_id	A	4	NULL	NULL		BTREE			YES
invoices	1	idx_invoices_invoice_date	1	invoice_date	A	17	NULL	NULL		BTREE			YES
invoices	1	idx_inv_client_id	1	client_id	A	4	NULL	NULL		BTREE			YES
invoices	1	idx_inv_date	1	invoice_date	A	17	NULL	NULL		BTREE			YES
invoices	1	idx_invoices_client_id_custom	1	client_id	A	4	NULL	NULL		BTREE			YES
invoices	1	idx_invoices_invoice_date_custom	1	invoice_date	A	17	NULL	NULL		BTREE			YES
invoices	1	idx_inv_client_id_20250627a	1	client_id	A	4	NULL	NULL		BTREE			YES
invoices	1	idx_inv_date_20250627d	1	invoice_date	A	17	NULL	NULL		BTREE			YES
invoices	1	idx_inv_date_final	1	invoice_date	A	17	NULL	NULL		BTREE			YES
invoices	1	idx_invoice_date_final	1	invoice_date	A	17	NULL	NULL		BTREE			YES
invoices	1	idx_invoice_total_final	1	invoice_total	A	17	NULL	NULL		BTREE			YES

Result Grid | Form Editor | Field Types | Query Stats

```sql
-- Show all indexes on the payments table
SHOW INDEXES FROM payments;
```

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

Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment
payments	0	PRIMARY	1	payment_id	A	8	NULL	NULL		BTREE		
payments	1	fk_client_id_idx	1	client_id	A	3	NULL	NULL		BTREE		
payments	1	fk_invoice_id_idx	1	invoice_id	A	7	NULL	NULL		BTREE		
payments	1	fk_payment_payment_method_idx	1	payment_method	A	2	NULL	NULL		BTREE		
payments	1	idx_payments_client_id	1	client_id	A	3	NULL	NULL		BTREE		
payments	1	idx_payments_invoice_id	1	invoice_id	A	7	NULL	NULL		BTREE		
payments	1	idx_payments_payment_method	1	payment_method	A	2	NULL	NULL		BTREE		
payments	1	idx_pay_client_id	1	client_id	A	3	NULL	NULL		BTREE		
payments	1	idx_pay_invoice_id	1	invoice_id	A	7	NULL	NULL		BTREE		
payments	1	idx_pay_method	1	payment_method	A	2	NULL	NULL		BTREE		
payments	1	idx_payments_client_id_custom	1	client_id	A	3	NULL	NULL		BTREE		
payments	1	idx_payments_invoice_id_custom	1	invoice_id	A	7	NULL	NULL		BTREE		
payments	1	idx_payments_method_custom	1	payment_method	A	2	NULL	NULL		BTREE		
payments	1	idx_pay_client_id_20250627b	1	client_id	A	3	NULL	NULL		BTREE		
payments	1	idx_pay_invoice_id_20250627c	1	invoice_id	A	7	NULL	NULL		BTREE		
payments	1	idx_pay_method_20250627e	1	payment_method	A	2	NULL	NULL		BTREE		
payments	1	idx_payments_date_final	1	date	A	6	NULL	NULL		BTREE		

Result Grid | Form Editor | Field Types | Query Stats | Execution Plan

```
7    -- Show all indexes on the clients table
8 •  SHOW INDEXES FROM clients;
```

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

Table	Non_unique	Key_name	Seq_in_index	Column_name	Collation	Cardinality	Sub_part	Packed	Null	Index_type	Comment	Index_comment	Visible
clients	0	PRIMARY	1	client_id	A	5	NULL	NULL		BTREE			YES
clients	1	idx_clients_city	1	city	A	5	NULL	NULL		BTREE			YES
clients	1	idx_clients_state	1	state	A	5	NULL	NULL		BTREE			YES
clients	1	idx_clients_city_20250627f	1	city	A	5	NULL	NULL		BTREE			YES
clients	1	idx_clients_state_20250627g	1	state	A	5	NULL	NULL		BTREE			YES
clients	1	idx_clients_city_final	1	city	A	5	NULL	NULL		BTREE			YES
clients	1	idx_clients_state_final	1	state	A	5	NULL	NULL		BTREE			YES

Result Grid | Form Editor | Field Types

Result 1 | Result 2 | Result 3

Read Only | Context