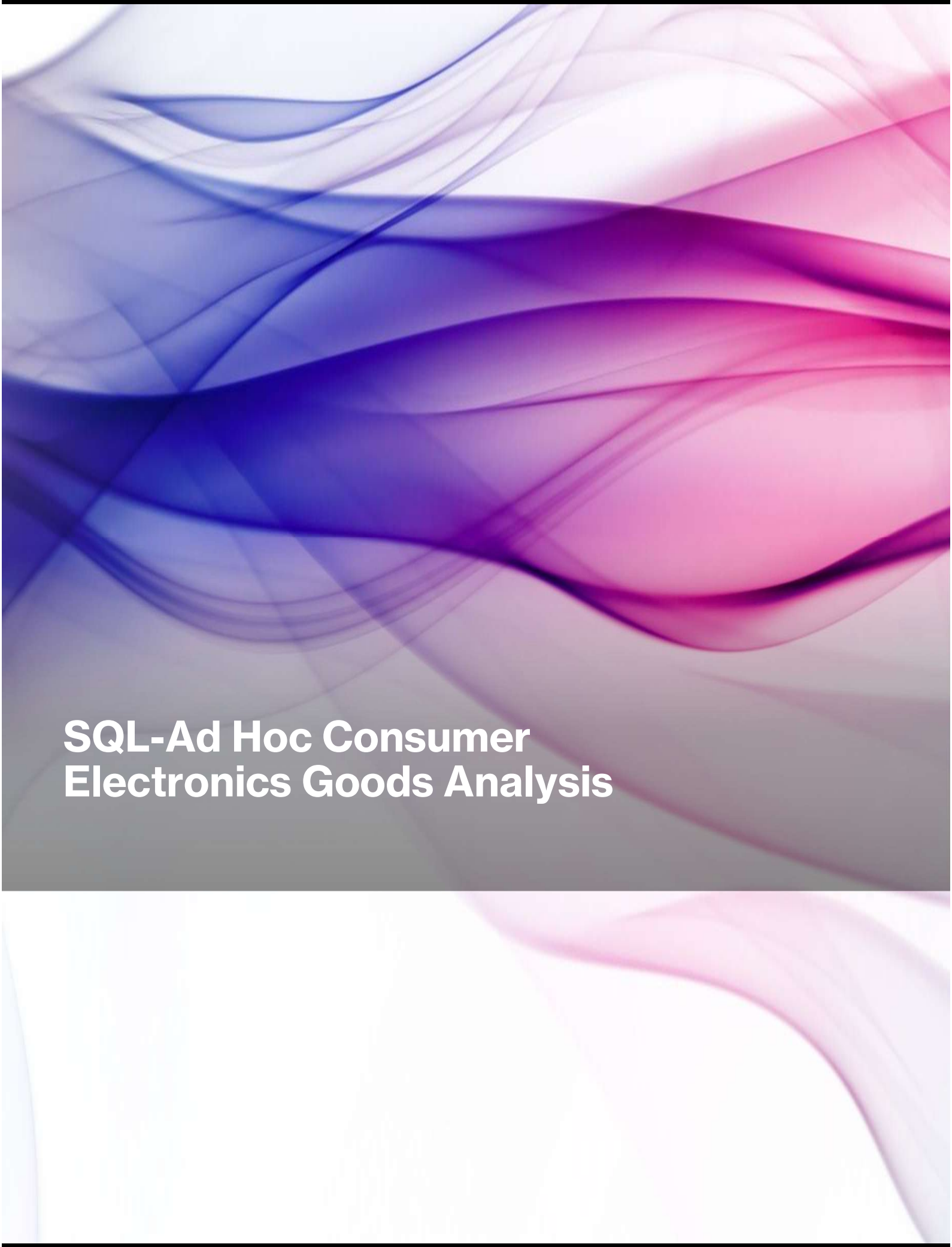

SQL-Ad Hoc Consumer
Electronics Goods Analysis

Retrieve data using text query(SELECT, WHERE, DISTINCT, LIKE)

```
SELECT market, fiscal_year, freight_pct
FROM fact_freight_cost;
```

```
SELECT *
FROM fact_gross_price
WHERE fiscal_year = 2020;
```

```
SELECT DISTINCT market
FROM fact_freight_cost;
```

```
SELECT product
FROM dim_product
WHERE product LIKE 'AQ MB%';
```

For data Filtering, aggregation, joins(Joins, Group by, Order by, Having, Between, In, If Clauses)

```
SELECT c.customer, d.market, p.pre_invoice_discount_pct
FROM fact_pre_invoice_deductions p
JOIN dim_customer c ON p.customer_code =
c.customer_code
JOIN fact_freight_cost d ON c.market = d.market AND
p.fiscal_year = d.fiscal_year;
```

```
SELECT market, COUNT(*) AS
num_customers
FROM dim_customer
GROUP BY market;
```

```
SELECT product_code, gross_price
FROM fact_gross_price
WHERE fiscal_year = 2022
ORDER BY gross_price DESC;
```

```
SELECT market, COUNT(*) AS count_customers
FROM dim_customer
GROUP BY market
HAVING COUNT(*) > 3;
```

```
SELECT *
FROM fact_freight_cost
WHERE freight_pct BETWEEN 0.02 AND
0.03;


SELECT *
FROM fact_gross_price
WHERE fiscal_year IN (2020, 2021);


SELECT *
FROM dim_customer
LIMIT 5;


SELECT *
FROM dim_customer
LIMIT 5 OFFSET 5;


SELECT customer_code,
      IF(pre_invoice_discount_pct > 0.2, 'High',
'Low') AS discount_level
FROM fact_pre_invoice_deductions
WHERE fiscal_year = 2022;


SELECT customer_code,
      CASE
        WHEN pre_invoice_discount_pct > 0.2
THEN 'High'
        WHEN pre_invoice_discount_pct > 0.1
THEN 'Medium'
        ELSE 'Low'
      END AS discount_category
FROM fact_pre_invoice_deductions
WHERE fiscal_year = 2022;
```

```
SELECT customer_code,
      CASE
       WHEN pre_invoice_discount_pct > 0.2
THEN 'High'
       WHEN pre_invoice_discount_pct > 0.1
THEN 'Medium'
       ELSE 'Low'
      END AS discount_category
FROM fact_pre_invoice_deductions
WHERE fiscal_year = 2022;
```

```
SELECT
    d.calender_date,
    YEAR(d.calender_date) AS calendar_year,
    dp.product,
    gp.gross_price
FROM dim_date d
JOIN fact_gross_price gp ON gp.fiscal_year = YEAR(d.calender_date)
JOIN dim_product dp ON gp.product_code = dp.product_code
WHERE dp.category = 'Graphic Card'
ORDER BY d.calender_date
LIMIT 10;
```

```
SELECT
    c.customer,
    p.fiscal_year,
    p.pre_invoice_discount_pct,
    CASE
        WHEN p.fiscal_year = CURYEAR() THEN 'Current Year'
        WHEN p.fiscal_year = CURYEAR() - 1 THEN 'Last Year'
        ELSE 'Older'
    END AS year_label
FROM fact_pre_invoice_deductions p
JOIN dim_customer c ON p.customer_code = c.customer_code
WHERE p.fiscal_year BETWEEN CURYEAR() - 2 AND CURYEAR()
ORDER BY c.customer
LIMIT 10;
```

Custom Function: `get_fiscal_year()`

```
-- Returns fiscal year for a given date. Assumes fiscal year starts
in September.

CREATE FUNCTION `get_fiscal_year`(
    calendar_date DATE) RETURNS INT
    DETERMINISTIC
BEGIN
    DECLARE fiscal_year INT;
    SET fiscal_year = YEAR(DATE_ADD(calendar_date, INTERVAL
4 MONTH));
    RETURN fiscal_year;
END
```

SQL Query

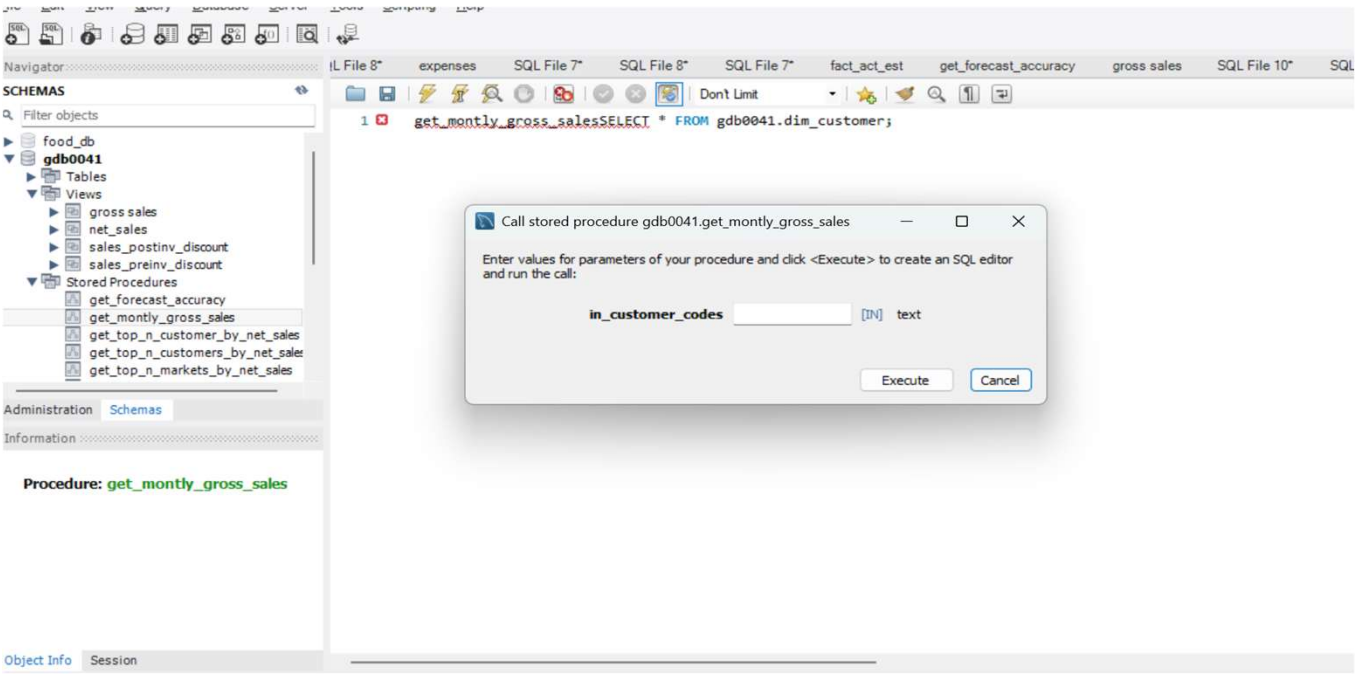
```
SELECT
    s.date,
    SUM(g.gross_price * s.sold_quantity) AS gross_price_total
FROM
    fact_sales_monthly s
JOIN
    fact_gross_price g
    ON s.product_code = g.product_code
    AND g.fiscal_year = get_fiscal_year(s.date)
WHERE
    customer_code = 90002002
GROUP BY
    s.date
ORDER BY
    s.date ASC;
```

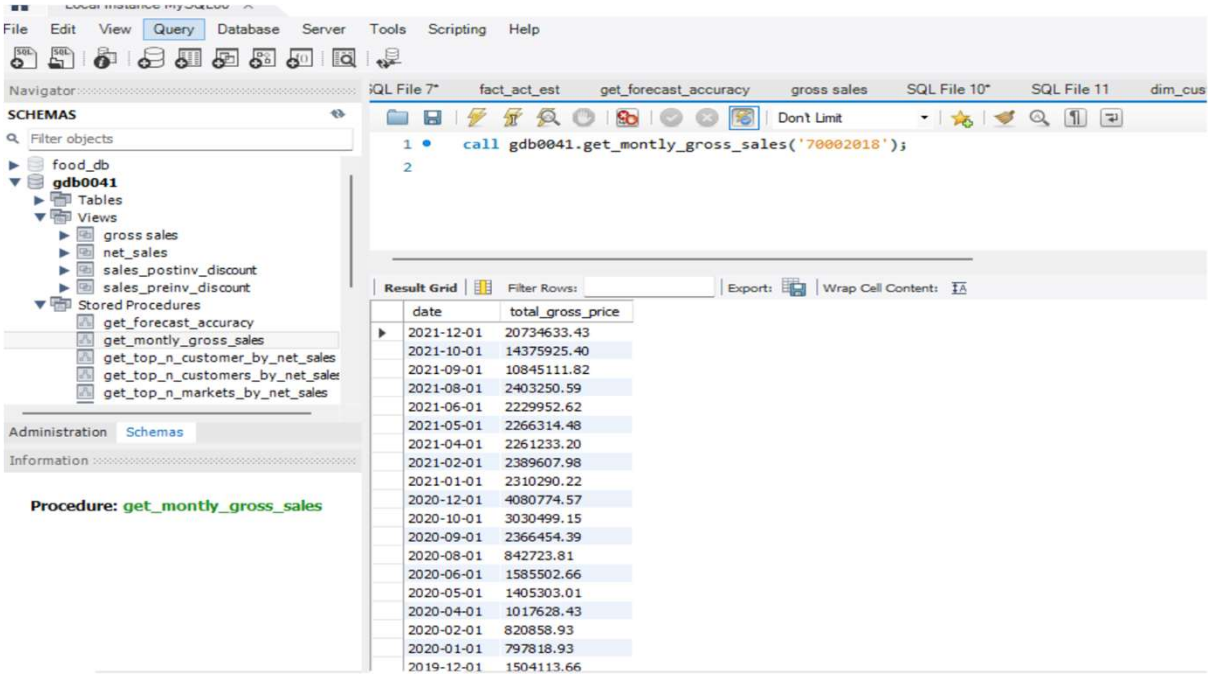
Stored Procedure – `get_monthly_gross_sales_for_customer`

To improve scalability and reusability of monthly gross sales reporting, I created a stored procedure that accepts **multiple customer codes** and returns their **aggregated monthly gross sales**.
Purpose
This stored procedure enables dynamic gross sales reporting across one or more customers without rewriting the query logic. Useful for both **manual analysis** and **automated pipelines**.

SQL Code

```
CREATE PROCEDURE `get_monthly_gross_sales_for_customer`(
   in_customer_codes TEXT)
BEGIN
   SELECT s.date,
      SUM(g.gross_price * s.sold_quantity) AS gross_price_total
   FROM   fact_sales_monthly s
   JOIN        fact_gross_price g
      ON s.product_code = g.product_code
      AND g.fiscal_year = get_fiscal_year(s.date)
   WHERE       FIND_IN_SET(s.customer_code,
in_customer_codes) > 0
   GROUP BY
      date;
END
```





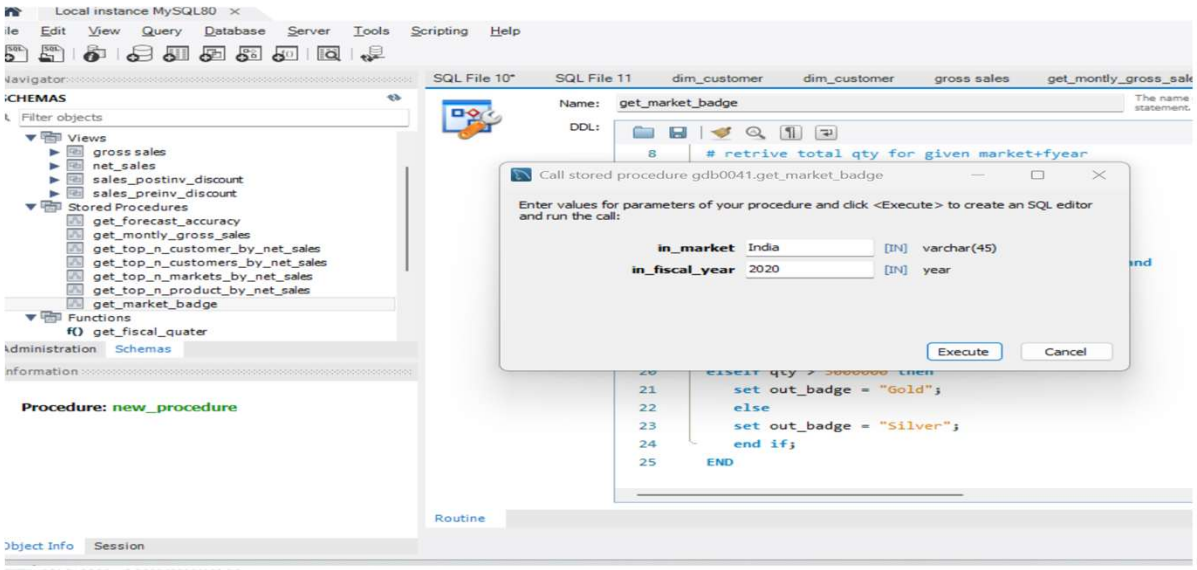
Create a stored procedure that can determine the market badge based on the following logic:

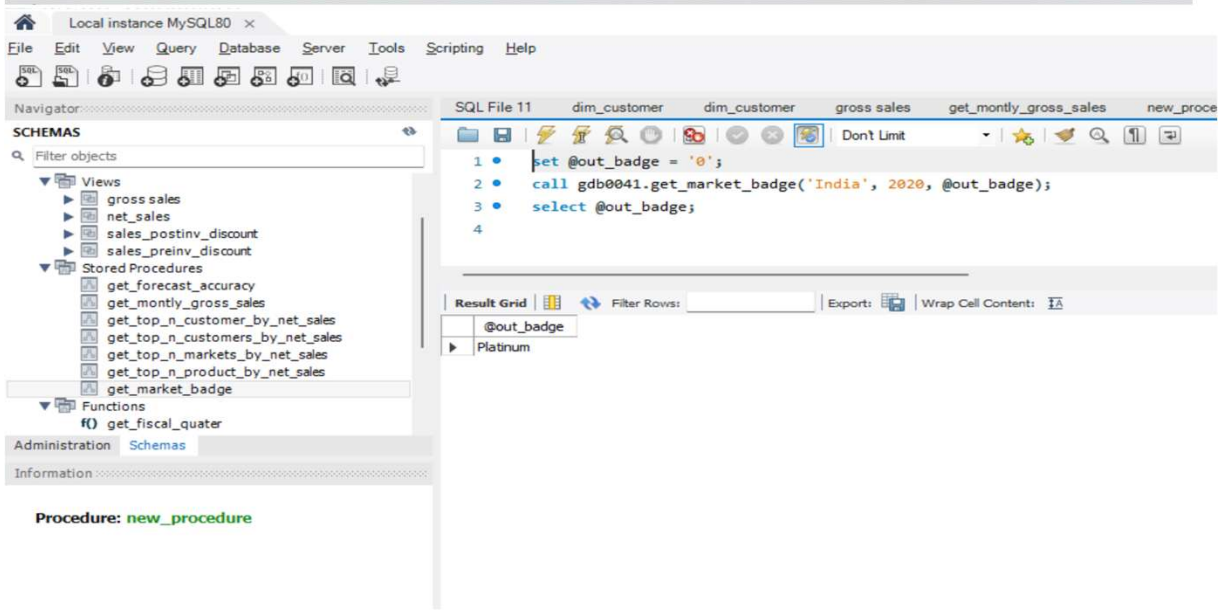
If total sold quantity > 5 million that market is considered Gold else it is Silver.

SQL Code

```
CREATE PROCEDURE get_market_badge(
    IN in_fiscal_year YEAR,
    IN in_market VARCHAR(45),
    OUT out_badge VARCHAR(45)
)
BEGIN
    DECLARE qty INT DEFAULT 0;

    -- Set default market to India if input is empty
    IF in_market = '' THEN
        SET in_market = 'India';
    END IF;
```

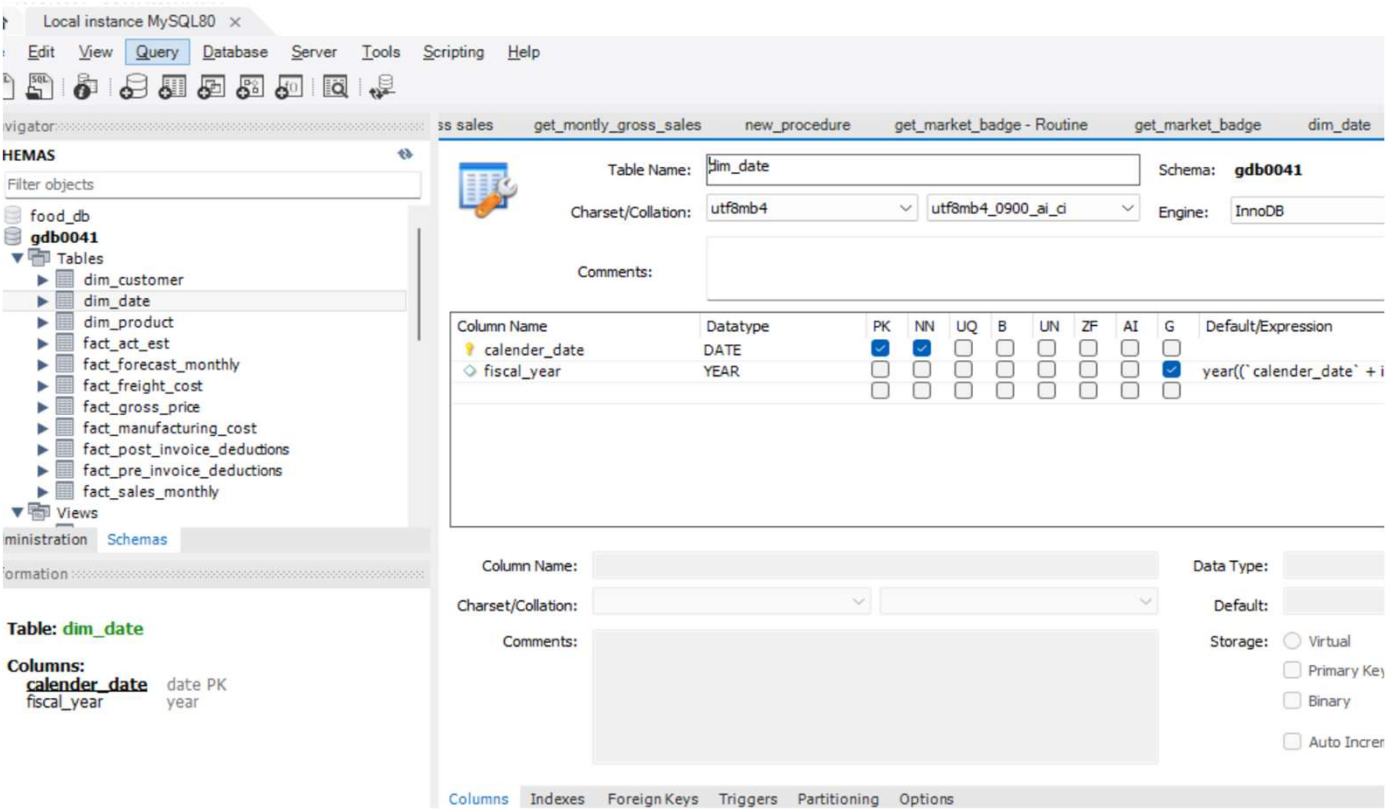
```
    -- Get total sold quantity for given market and fiscal year
    SELECT
        SUM(s.sold_quantity) INTO qty
    FROM
        fact_sales_monthly s
    JOIN
        dim_customer c ON s.customer_code = c.customer_code
    WHERE  get_fiscal_year(s.date) = in_fiscal_year
        AND c.market = in_market;

    -- Classify market badge based on quantity
    IF qty > 5000000 THEN
        SET out_badge = 'Gold';
    ELSE
        SET out_badge = 'Silver';
    END IF;
END
```

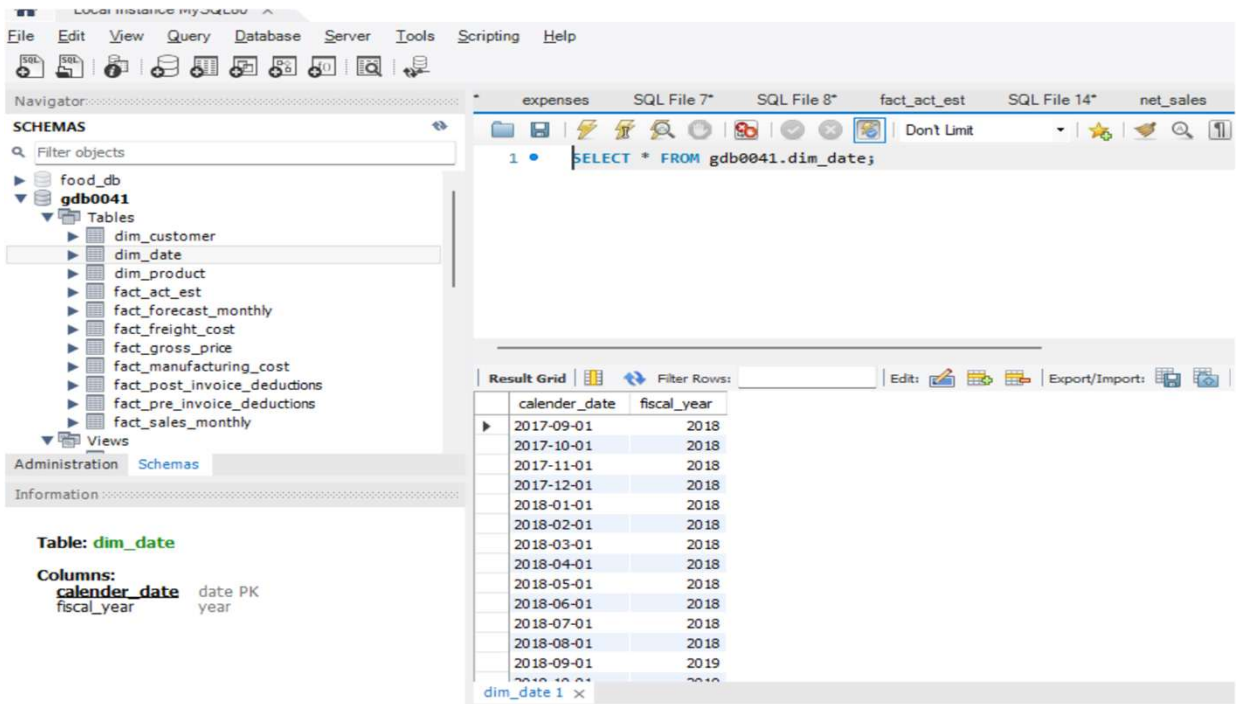




Performance Improvement #1: Add `dim_date` Table

By introducing the `dim_date` table, we avoid applying functions directly on date columns in the `WHERE` clause and improve join logic across time-based tables. This enhances query performance and supports better filtering and aggregation.

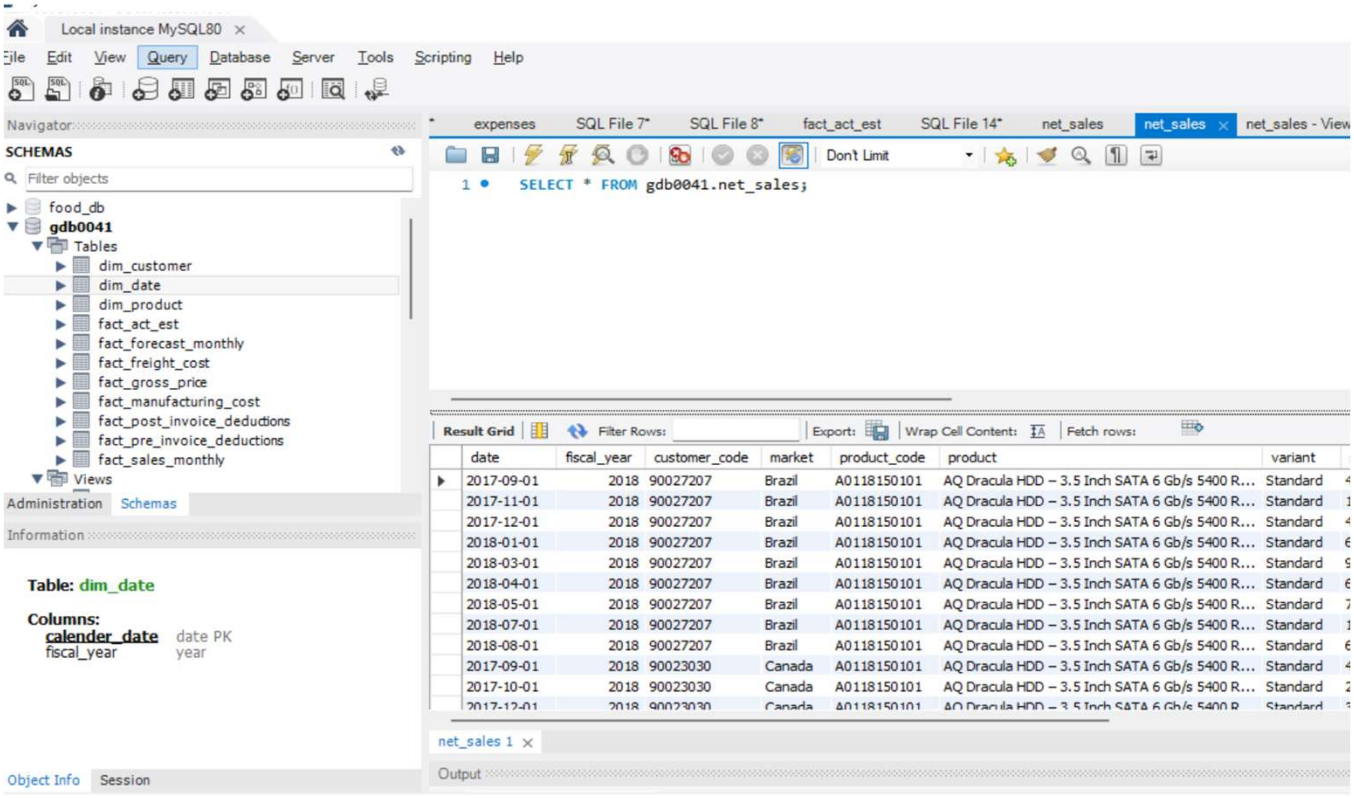




Performance Improvement #2: Add `fiscal_year` Column in `fact_sales_monthly`

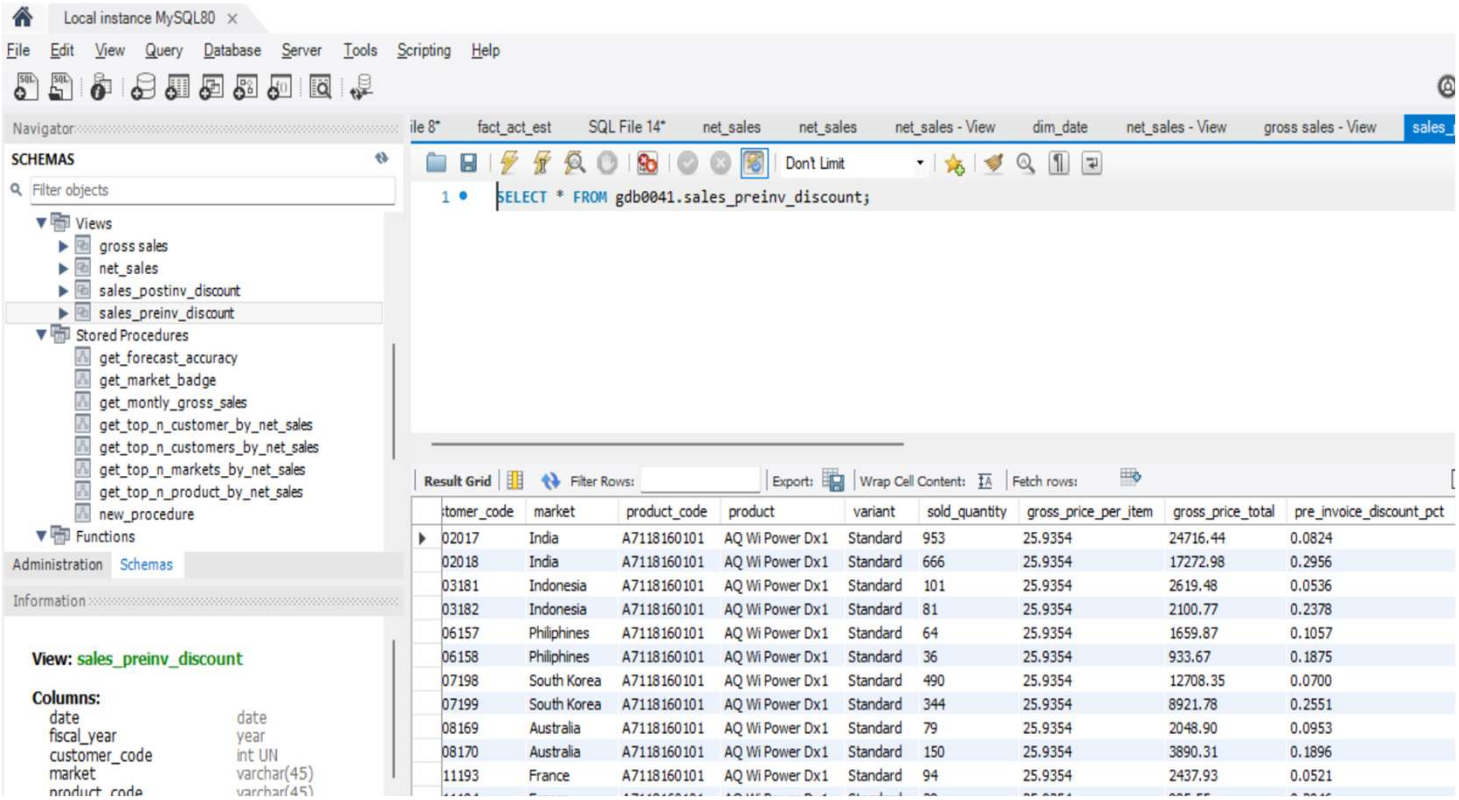
Adding a `fiscal_year` column to the `fact_sales_monthly` table simplifies joins with other fact tables and eliminates the need for a date dimension join in certain use cases, improving performance and clarity.

```
SELECT
    s.date,
    s.product_code,
    p.product,
    p.variant,
    s.sold_quantity,
    ROUND(g.gross_price, 2) AS gross_price,
    ROUND(g.gross_price * s.sold_quantity, 2) AS gross_price_total,
    pre.pre_invoice_discount_pct
FROM
    fact_sales_monthly s
JOIN dim_product p
    ON s.product_code = p.product_code
JOIN fact_gross_price g
    ON g.product_code = s.product_code AND g.fiscal_year = s.fiscal_year
JOIN fact_pre_invoice_deductions pre
    ON pre.customer_code = s.customer_code AND pre.fiscal_year = s.fiscal_year
WHERE
    s.fiscal_year = 2021
LIMIT 1000000;
```

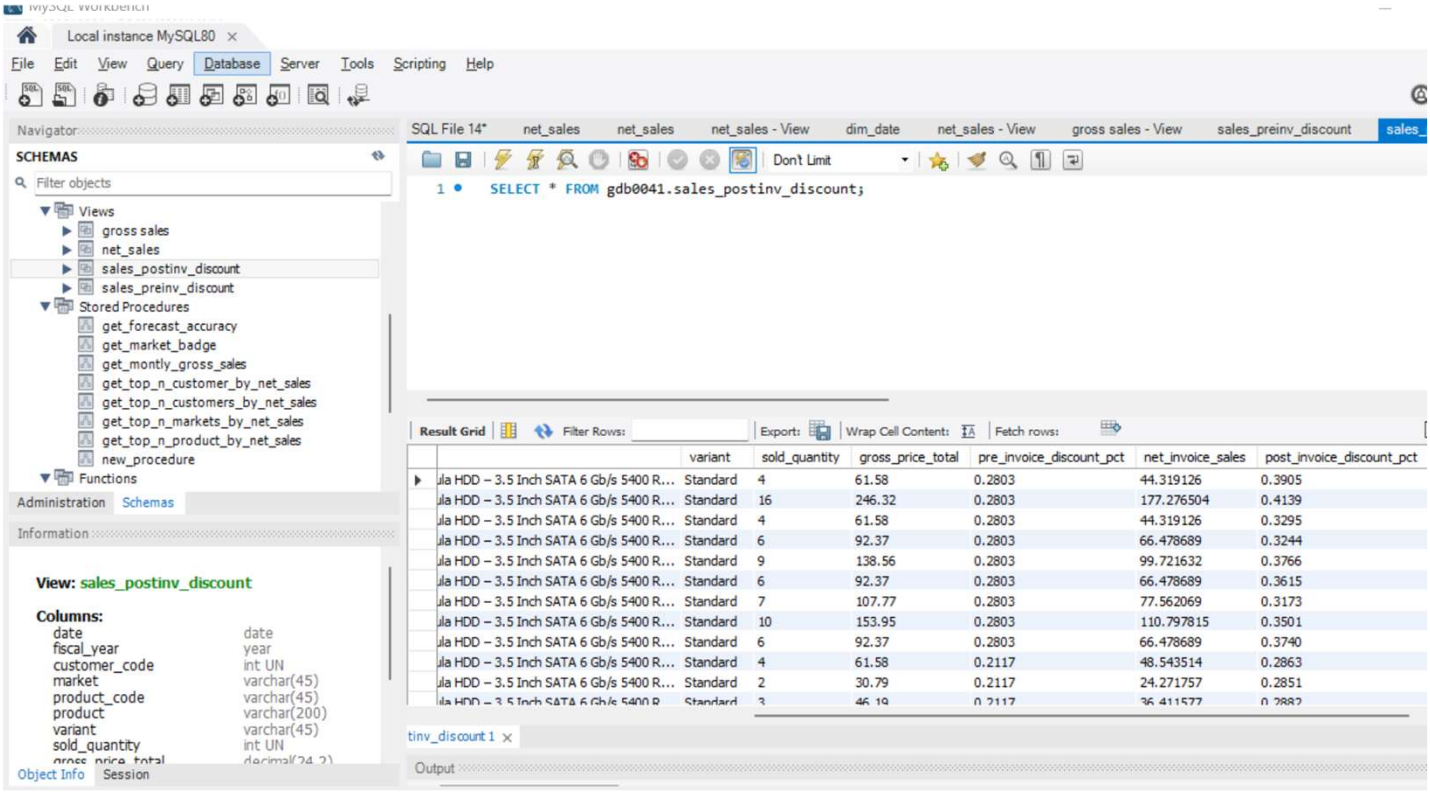


Create `sales_preinv_discount` View

```
CREATE VIEW sales_preinv_discount AS
SELECT
    s.date,
    s.fiscal_year,
    s.customer_code,
    c.market,
    s.product_code,
    p.product,
    p.variant,
    s.sold_quantity,
    ROUND(g.gross_price, 2) AS gross_price,
    ROUND(g.gross_price * s.sold_quantity, 2) AS gross_price_total,
    pre.pre_invoice_discount_pct
FROM
    fact_sales_monthly s
JOIN dim_customer c
    ON c.customer_code = s.customer_code
JOIN dim_product p
    ON s.product_code = p.product_code
JOIN fact_gross_price g
    ON g.product_code = s.product_code AND g.fiscal_year = s.fiscal_year
JOIN fact_pre_invoice_deductions pre
    ON pre.customer_code = s.customer_code AND pre.fiscal_year =
s.fiscal_year;
```

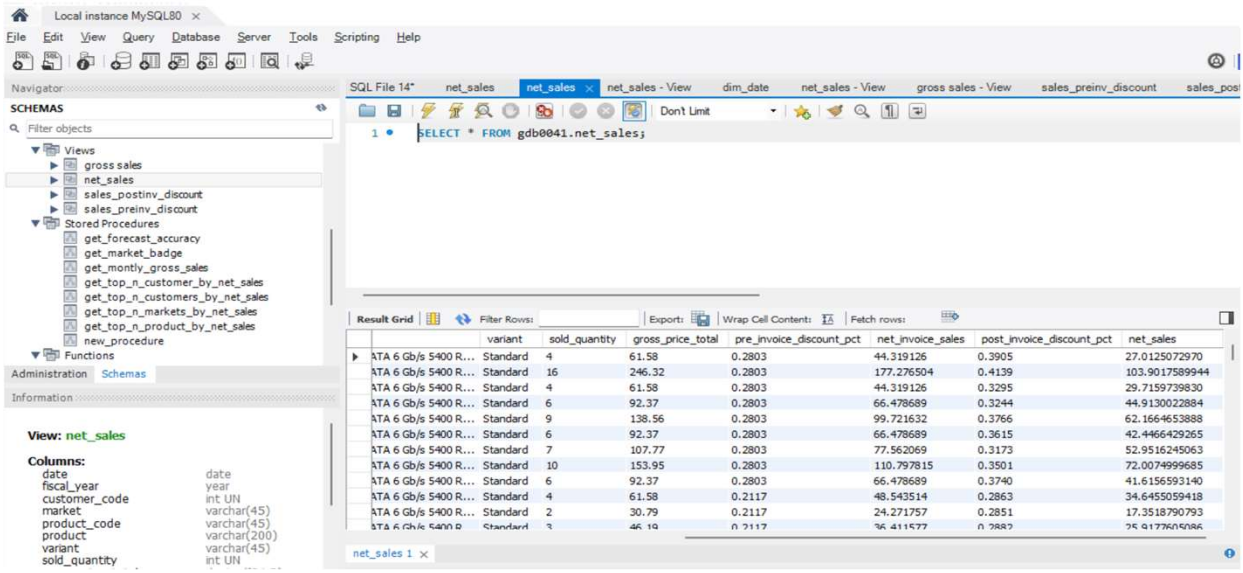


Create `sales_postinv_discount` View
```sql
CREATE VIEW sales_postinv_discount AS
SELECT
    s.date,
    s.fiscal_year,
    s.customer_code,
    s.market,
    s.product_code,
    s.product,
    s.variant,
    s.sold_quantity,
    s.gross_price,
    s.gross_price_total,
    s.pre_invoice_discount_pct,
    (1 - s.pre_invoice_discount_pct) * s.gross_price_total AS net_invoice_sales,
    (po.discounts_pct + po.other_deductions_pct) AS post_invoice_discount_pct
FROM
    sales_preinv_discount s
JOIN fact_post_invoice_deductions po
    ON po.customer_code = s.customer_code
    AND po.product_code = s.product_code
    AND po.date = s.date;
```



Calculate Final Net Sales

```
SELECT
    *,
    (1 - post_invoice_discount_pct) * net_invoice_sales AS net_sales
FROM
    sales_postinv_discount;
```

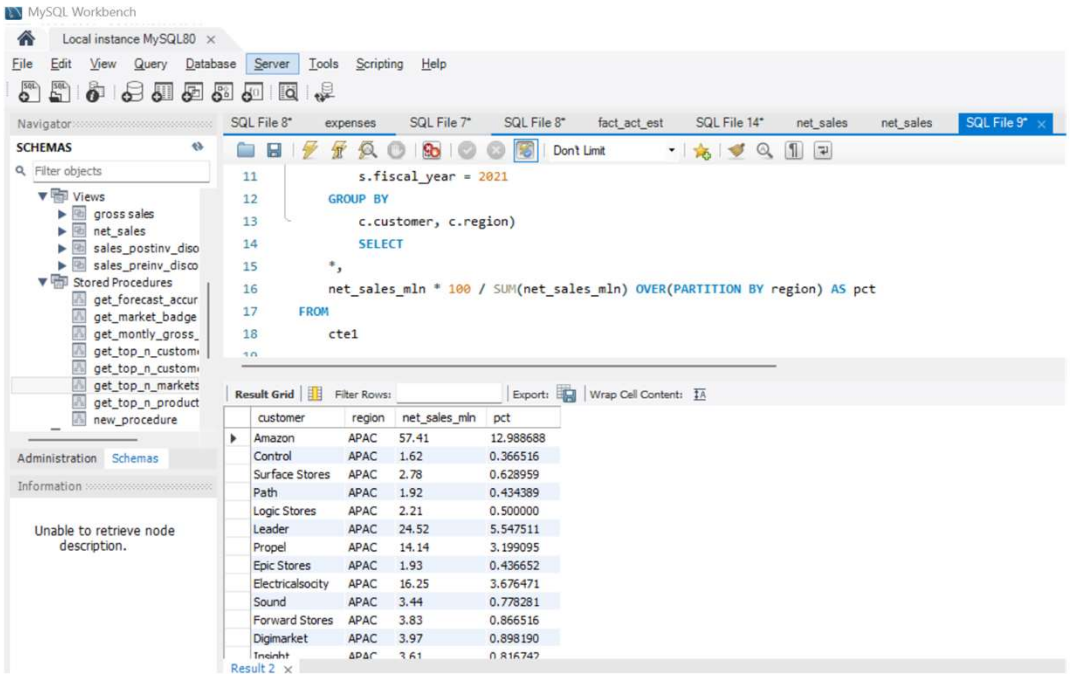


****Generate a Region-wise % Net Sales Breakdown by Customers.****

This query generates a **region-wise percentage breakdown of net sales by customers for a given fiscal year.**
It supports **regional financial analysis by showing how much each customer contributes to the total sales in their respective region.**

SQL Query

```
WITH cte1 AS (
    SELECT
        c.customer,
        c.region,
        ROUND(SUM(net_sales) / 1000000, 2) AS net_sales_mln
    FROM
        net_sales s
    JOIN
        dim_customer c USING (customer_code)
    WHERE
        s.fiscal_year = 2021
    GROUP BY
        c.customer, c.region
)
SELECT
    *,
    net_sales_mln * 100 / SUM(net_sales_mln) OVER(PARTITION BY region) AS pct
FROM
    cte1
ORDER BY
    region, net_sales_mln DESC;
```

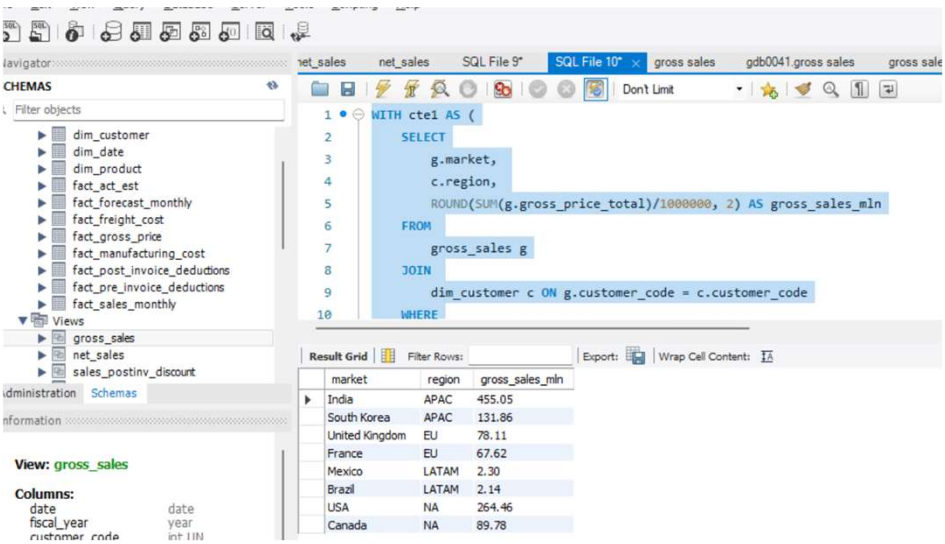


****Retrieve the Top 2 Markets in Every Region by Gross Sales.****

This query identifies the **top 2 performing markets within each region based on their **gross sales** for the fiscal year **2021**.**
It uses the `DENSE_RANK()` window function to handle ties in rankings accurately.

SQL Query

```
WITH cte1 AS (
    SELECT
        g.market,
        c.region,
        ROUND(SUM(g.gross_price_total)/1000000, 2) AS gross_sales_mln
    FROM
        gross_sales g
    JOIN
        dim_customer c ON g.customer_code = c.customer_code
    WHERE
        g.fiscal_year = 2021
    GROUP BY
        g.market, c.region
),
cte2 AS (
    SELECT
        *,
        DENSE_RANK() OVER (PARTITION BY region ORDER BY gross_sales_mln
DESC) AS drnk
    FROM
        cte1
)
SELECT
    market,
    region,
    gross_sales_mln
FROM
    cte2
WHERE
    drnk <= 2;
```



****Generate an Aggregate Forecast Accuracy Report for Customers for a Given Fiscal Year.****

Forecast Accuracy Calculation

Using the helper table, the query computes:
- **Total sold quantity** and **forecast quantity** per customer
- **Net forecast error** and **net error percentage**
- **Absolute forecast error** and **absolute error percentage**
- **Forecast accuracy** as:
  `Forecast Accuracy = 100 - Absolute Error %` (capped at 100%)

Create helper table combining actual sales and forecast quantities

```
(
  SELECT
    s.date,  s.fiscal_year,  s.product_code,  s.customer_code,
    s.sold_quantity,  f.forecast_quantity
  FROM
    fact_sales_monthly s
  LEFT JOIN
    fact_forecast_monthly f
  USING (date, customer_code, product_code)
)
UNION
(
  SELECT
    f.date,  f.fiscal_year,  f.product_code,
    f.customer_code,  s.sold_quantity,  f.forecast_quantity
  FROM fact_forecast_monthly f
  LEFT JOIN fact_sales_monthly s
  USING (date, product_code, customer_code)
);
```

Calculate forecast accuracy metrics per customer for fiscal year 2021
WITH forecast_err_table

```
(
   SELECT
      customer_code,
      SUM(sold_quantity) AS total_sold_quantity,
      SUM(forecast_quantity) AS total_forecast_quantity,
      SUM(forecast_quantity - sold_quantity) AS net_err,
      SUM(forecast_quantity - sold_quantity) * 100 / SUM(forecast_quantity)
AS net_err_pct,
      SUM(ABS(forecast_quantity - sold_quantity)) AS abs_err,
      SUM(ABS(forecast_quantity - sold_quantity)) * 100 /
SUM(forecast_quantity) AS abs_err_pct
   FROM
      fact_act_est
   WHERE
      fiscal_year = 2021
   GROUP BY
      customer_code
)
SELECT
   e.*,
   c.market,
   c.customer,
   IF(abs_err_pct > 100, 0, 100 - abs_err_pct) AS forecast_accuracy
FROM
   forecast_err_table e
JOIN
   dim_customer c
USING
   (customer_code)
ORDER BY
   forecast_accuracy DESC;
```