

# Queries ran in Athena

## ***Query 1:***

```
SELECT
    "date",
    SUM(CAST(amount AS DOUBLE)) AS daily_sales,
    SUM(SUM(CAST(amount AS DOUBLE))) OVER (ORDER BY "date" ASC)
AS cumulative_sales
FROM
    "AwsDataCatalog"."output_db"."raw"
WHERE
    status IN ('Shipped', 'Shipped - Delivered to Buyer')
    AND substr("date", 7, 2) = '22' -- Assuming 2022 (based on '04-30-22' format)
GROUP BY
    "date"
ORDER BY
    "date" ASC
LIMIT 10;
```

---

## ***Query 2:***

```
SELECT
    COALESCE(NULLIF("ship-state", ""), 'Unknown') AS state,
    SUM(
        CASE
            WHEN lower(status) IN ('cancelled','canceled','refunded','returned','return
requested','refund initiated')
            THEN -ABS(amount)
            ELSE 0
        END
    ) AS total_negative_profit
FROM raw
GROUP BY COALESCE(NULLIF("ship-state", ""), 'Unknown')
```

```

HAVING SUM(
  CASE
    WHEN lower(status) IN ('cancelled','canceled','refunded','returned','return
requested','refund initiated')
    THEN -ABS(amount)
    ELSE 0
  END
) < 0
ORDER BY total_negative_profit ASC
LIMIT 10;

```

---

### ***Query 3:***

```

SELECT
  category,
  sku AS sub_category,
  ROUND(AVG(rand() * 0.3), 2) AS avg_discount,          -- simulate
discount between 0%–30%
  CAST(SUM(CAST(amount AS DOUBLE)) AS DOUBLE) AS total_sales,
  CAST(SUM(CAST(amount AS DOUBLE)) * (1 - AVG(rand() * 0.3)) AS
DOUBLE) AS est_profit,
  (CAST(SUM(CAST(amount AS DOUBLE)) * (1 - AVG(rand() * 0.3)) AS
DOUBLE) /
  NULLIF(CAST(SUM(CAST(amount AS DOUBLE)) AS DOUBLE), 0)) AS
profit_ratio
FROM raw
GROUP BY category, sku
ORDER BY profit_ratio DESC
LIMIT 10;

```

---

### ***Query 4:***

```

WITH ranked_products AS (
  SELECT
    category,

```

```

        sku AS product_name,
        CAST(SUM(CAST(amount AS DOUBLE) * 0.25) AS DOUBLE) AS
total_profit,
        RANK() OVER (PARTITION BY category ORDER BY SUM(CAST(amount
AS DOUBLE) * 0.25) DESC) AS rank_in_category
FROM raw
GROUP BY category, sku
)
SELECT
    category,
    product_name,
    total_profit,
    rank_in_category
FROM ranked_products
WHERE rank_in_category <= 3
ORDER BY category, rank_in_category
LIMIT 10

```

---

***Query 5:***

```

WITH monthly_summary AS (
    SELECT
        date_format(date_parse(date, '%m-%d-%y'), '%Y-%m') AS month,
        CAST(SUM(CAST(amount AS DOUBLE)) AS DOUBLE) AS total_sales,
        CAST(SUM(CAST(amount AS DOUBLE) * 0.25) AS DOUBLE) AS
total_profit -- assume 25% margin
    FROM raw
    GROUP BY date_format(date_parse(date, '%m-%d-%y'), '%Y-%m')
)
SELECT
    month,
    total_sales,
    total_profit,
    ROUND(
        ((total_sales - LAG(total_sales) OVER (ORDER BY month)) /

```

```
    NULLIF(LAG(total_sales) OVER (ORDER BY month), 0)) * 100, 2
) AS sales_growth_rate,
ROUND(
    ((total_profit - LAG(total_profit) OVER (ORDER BY month)) /
    NULLIF(LAG(total_profit) OVER (ORDER BY month), 0)) * 100, 2
) AS profit_growth_rate
FROM monthly_summary
ORDER BY month
LIMIT 10;
```

---