SQL Project

NAME: VAIBHAV PAL

MAIL_ID:VAIBHAVPAL2022@GMAIL.COM

Dataset:

Sales

Language:

SQL

Dataset link:

https://drive.google.com/file/d/1W5Umk0ABVMM9--uvuAYvgPcBRsEyYh3G/view

Questions:

Various queries are written to retrieve meaningful data insights from the sales dataset.

Solutions:

Schema:

```
CREATE TABLE Sales_Data (
  order_id INT NOT NULL,
  order_date DATE,
  ship_date DATE,
  ship_mode VARCHAR(50),
  customer_name VARCHAR(100),
  segment VARCHAR(50),
  state VARCHAR(50),
  market VARCHAR(50),
  category VARCHAR(50),
  sub_category VARCHAR(50),
  product_name VARCHAR(100),
  sales DECIMAL(10, 2),
  quantity INT,
  discount DECIMAL(5, 2),
  profit DECIMAL(10, 2),
  shipping_cost DECIMAL(10, 2),
  order_priority VARCHAR(50),
  year INT,
  PRIMARY KEY (order_id)
);
```

1. Top 5 products by sales:

This query selects the top 5 products based on total sales.

```
SELECT product_name, SUM(sales) AS total_sales
FROM Sales_Data
GROUP BY product_name
ORDER BY total_sales DESC
LIMIT 5;
```

---

2. Average profit margin per category:

This query calculates the average profit margin for each product category.

```
SELECT category, AVG((profit / sales) * 100) AS avg_profit_margin
FROM Sales_Data
GROUP BY category;
```

---

3. Number of orders placed per segment in 2012:

This query retrieves the total number of orders placed for each segment in the year 2012.

```
SELECT segment, COUNT(order_id) AS total_orders
FROM Sales_Data
WHERE year = 2012
GROUP BY segment;
```

---

4. Top 3 states with highest shipping costs:

This query selects the top 3 states with the highest total shipping costs.

```
SELECT state, SUM(shipping_cost) AS total_shipping_cost
FROM Sales_Data
GROUP BY state
ORDER BY total_shipping_cost DESC
LIMIT 3;
```

---

5. Total sales and profit for each market in 2012:

This query calculates the total sales and profit for each market in the year 2012.

```
SELECT market, SUM(sales) AS total_sales, SUM(profit) AS
total_profit
FROM Sales_Data
WHERE year = 2012
GROUP BY market;
```

---

6. Percentage of orders with a discount:

This query calculates the percentage of orders that had a discount.

```
SELECT (COUNT(CASE WHEN discount > 0 THEN 1 END) * 100.0 /
COUNT(*)) AS discount_percentage
FROM Sales_Data;
```

---

7. Average shipping cost per order priority:

This query calculates the average shipping cost for each order priority.

```
SELECT order_priority, AVG(shipping_cost) AS avg_shipping_cost
FROM Sales_Data
GROUP BY order_priority;
```

---

8. Products with profit margin greater than 20%:

This query selects products that have a profit margin greater than 20%.

```
SELECT product_name, (profit / sales) * 100 AS profit_margin
FROM Sales_Data
HAVING profit_margin > 20;
```

---

9. Total quantity sold for each sub-category:

This query calculates the total quantity sold for each sub-category.

```
SELECT sub_category, SUM(quantity) AS total_quantity_sold
FROM Sales_Data
GROUP BY sub_category;
```

---

10. CTE to calculate total sales and profit per year:

This query uses a Common Table Expression (CTE) to calculate total sales and profit per year.

```
WITH yearly_sales AS (
  SELECT year, SUM(sales) AS total_sales, SUM(profit) AS total_profit
  FROM Sales_Data
  GROUP BY year
)
SELECT * FROM yearly_sales;
```

---

11. CTE for products with profit margin greater than 20%:

This query uses a CTE to select products that have a profit margin greater than 20%.

```
WITH product_margins AS (
   SELECT product_name, (profit / sales) * 100 AS profit_margin
   FROM Sales_Data
)
SELECT product_name, profit_margin
FROM product_margins
WHERE profit_margin > 20;
```
