SQL Queries

Top-Selling Products:

Question: What are the top 5 best-selling products in terms of quantity ordered?

```sql
--Question: What are the top 5 best-selling products in terms of quantity ordered?

SELECT product, SUM(quantity_ordered) AS total_quantity FROM sales_data
GROUP BY product
ORDER BY total_quantity DESC
LIMIT 5;
```

Expected Outcome: A list of the top-selling products with their respective quantities ordered.

product character varying (255)	total_quantity bigint
AAA Batteries (4-pack)	31017
AA Batteries (4-pack)	27635
USB-C Charging Cable	23975
Lightning Charging Cable	23217
Wired Headphones	20557

Revenue Analysis:

Question: What is the total revenue generated during the entire period covered by the dataset?

```sql
--Question: What is the total revenue generated during the entire period covered by the dataset?

SELECT  SUM(quantity_ordered * price) AS total_revenue
FROM sales_data;
```

Expected Outcome: A single value representing the total revenue from sales.

	total_revenue bigint
1	34499750

Monthly Sales Trends:

Question: How does sales revenue vary month by month?

```sql
--Question: How does sales revenue vary month by month?

SELECT month,  SUM(quantity_ordered * price) AS monthly_revenue
FROM sales_data
GROUP BY month
ORDER BY monthly_revenue;
```

Expected Outcome: A monthly breakdown of sales revenue, showing trends or seasonality.

month character varying (255)	monthly_revenue bigint
January	1822655
September	2098053
February	2202517
August	2244966
June	2578373
July	2648362
March	2807722
May	3153280
November	3200335
April	3391444
October	3737561
December	4614482

City-wise Sales Performance:

Question: Which city has the highest total sales?

```sql
--Question: Which city has the highest total sales?

SELECT city,  SUM(quantity_ordered * price) AS revenue
FROM sales_data
GROUP BY city
ORDER BY revenue DESC
LIMIT 3;
```

Expected Outcome: Identification of the city with the highest sales and the corresponding sales amount.

city character varying (255)	revenue bigint
San Francisco	8264049
Los Angeles	5453807
New York City	4665339

Order Quantity Distribution:

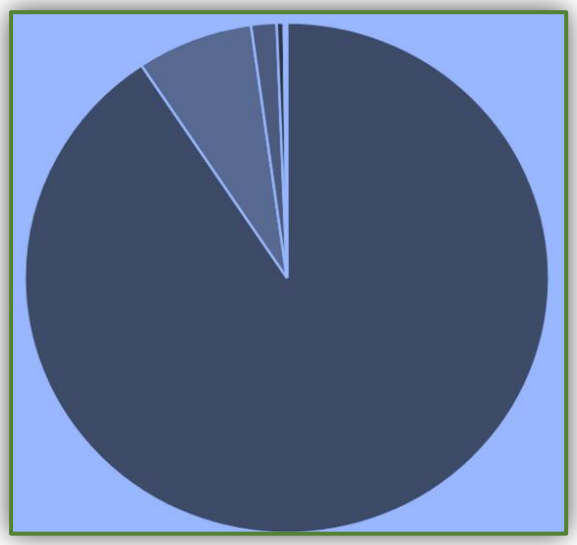
Question: What is the distribution of order quantities? Are most orders small or large?

```sql
--Question: What is the distribution of order quantities? Are most orders small or large?

SELECT quantity_ordered AS quantity, COUNT(quantity_ordered) AS Distribution
FROM sales_data
GROUP BY quantity_ordered
ORDER BY quantity_ordered;
```

Expected Outcome: A histogram or summary statistics showing the distribution of order quantities.

quantity integer	distribution bigint
1	168552
2	13324
3	2920
4	806
5	236
6	80
7	24
8	5
9	3



Average Order Value:

Question: What is the average value of an order (average order value)?

```
--Question: What is the average value of an order (average order value)?

SELECT AVG(sales) AS average_order_value
FROM sales_data;
```

Expected Outcome: A single value representing the average order value.

average_order_value numeric
185.5305888679752622

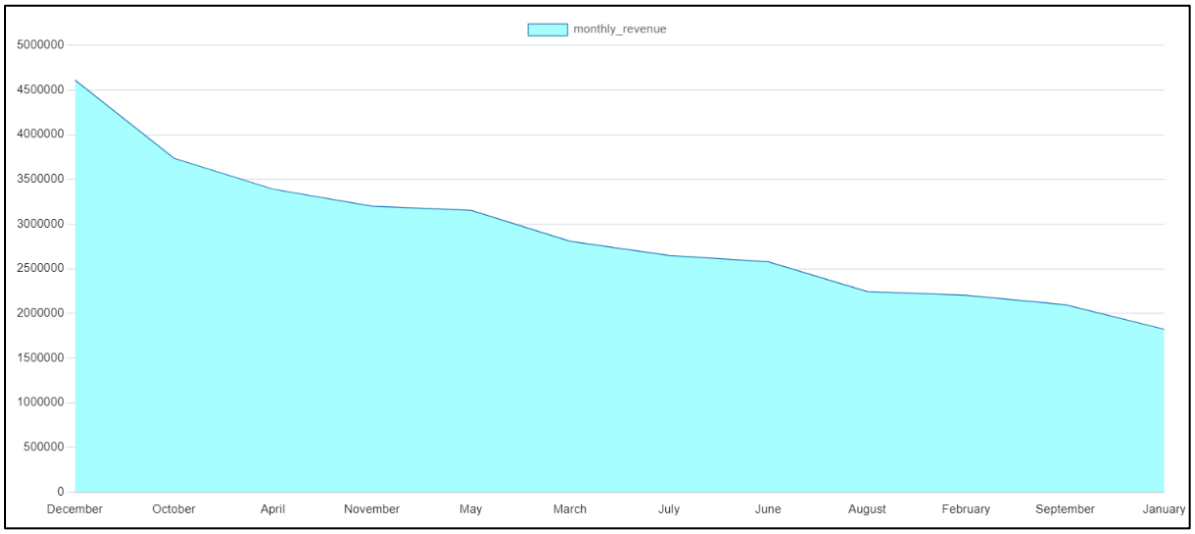
Seasonal Sales Analysis:

Question: Are there any noticeable patterns or trends in sales during different months?

```sql
--Question: Are there any noticeable patterns or trends in sales during different months?

SELECT month, SUM(quantity_ordered * price) AS monthly_revenue
FROM sales_data
GROUP BY month
ORDER BY monthly_revenue DESC;
```

Expected Outcome: Visualization or summary highlighting any seasonal trends in sales.



Correlation Between Price and Sales:

Question: Is there a correlation between the product price and the quantity ordered?

```sql
--Question: Is there a correlation between the product price and the quantity ordered?

SELECT AVG(price) AS avg_price, AVG(quantity_ordered) AS avg_quantity
FROM sales_data;
```

Expected Outcome: Insights into whether higher-priced products tend to have lower quantities ordered or vice versa.

avg_price numeric	avg_quantity numeric
184.4335090077977951	1.1243828986286636

Customer Segmentation:

Question: Can we identify different customer segments based on their purchasing behavior?

```sql
--Question: Can we identify different customer segments based on their purchasing behavior?

SELECT
CASE
WHEN quantity_ordered >=7 THEN 'High'
WHEN quantity_ordered >=4 THEN 'Medium'
ELSE 'Low'
END AS customer_segment,
COUNT(*) AS customer_count
FROM sales_data
GROUP BY customer_segment;
```

Expected Outcome: Clusters or groups of customers with similar purchasing patterns.

customer_segment text	customer_count bigint
High	32
Low	184796
Medium	1122

Geographical Analysis:

Question: How do sales vary across different regions or addresses?

```sql
--Question: How do sales vary across different regions or addresses?

SELECT city, SUM(quantity_ordered * price) AS Revenue_by_region
FROM sales_data
GROUP BY city;
```

Expected Outcome: Regional analysis highlighting areas with higher and lower sales.

	city character varying (255)	revenue_by_region bigint
1	Atlanta	2796112
2	Austin	1819987
3	Boston	3662478
4	Dallas	2768598
5	Los Angeles	5453807
6	New York City	4665339
7	Portland	2321019
8	San Francisco	8264049
9	Seattle	2748361