Analytical Queries Documentation

Type 1: Sales Performance Analysis

Query 1.1: Sales Performance by Product Category and Quarter

This query looks at how sales have changed over time for different types of products. It does this to find seasonal patterns and performance metrics that are special to each type of product.

a. SQL Script:

```
-- Sales Performance by Product Category and Quarter
SELECT
    dp.CategoryName,
    dd.Year,
    dd.Quarter,
    SUM(fs.Quantity) AS TotalQuantity,
    SUM(fs.Sales) AS TotalSales,
    COUNT(DISTINCT fs.OrderID) AS NumberOfOrders,
    SUM(fs.Sales)/COUNT(DISTINCT fs.OrderID) AS AvgOrderValue
FROM fact_Sales fs
JOIN dim_Product dp ON fs.ProductSK = dp.ProductSK
JOIN dim_Date dd ON fs.OrderDateSK = dd.DateSK
WHERE dp.IsCurrent = 1
GROUP BY dp.CategoryName, dd.Year, dd.Quarter
ORDER BY dp.CategoryName, dd.Year, dd.Quarter;
```

b. Tabular Format(Output)

```sql
--Analytical Query Type 1: Sales Performance Analysis
--Query 1.1: Sales Performance by Product Category and Quarter

-- Sales Performance by Product Category and Quarter
SELECT
    dp.CategoryName,
    dd.Year,
    dd.Quarter,
    SUM(fs.Quantity) AS TotalQuantity,
    SUM(fs.Sales) AS TotalSales,
    COUNT(DISTINCT fs.OrderID) AS NumberOfOrders,
    SUM(fs.Sales)/COUNT(DISTINCT fs.OrderID) AS AvgOrderValue
FROM fact_Sales fs
JOIN dim_Product dp ON fs.ProductSK = dp.ProductSK
JOIN dim_Date dd ON fs.OrderDateSK = dd.DateSK
WHERE dp.IsCurrent = 1
```

100 %

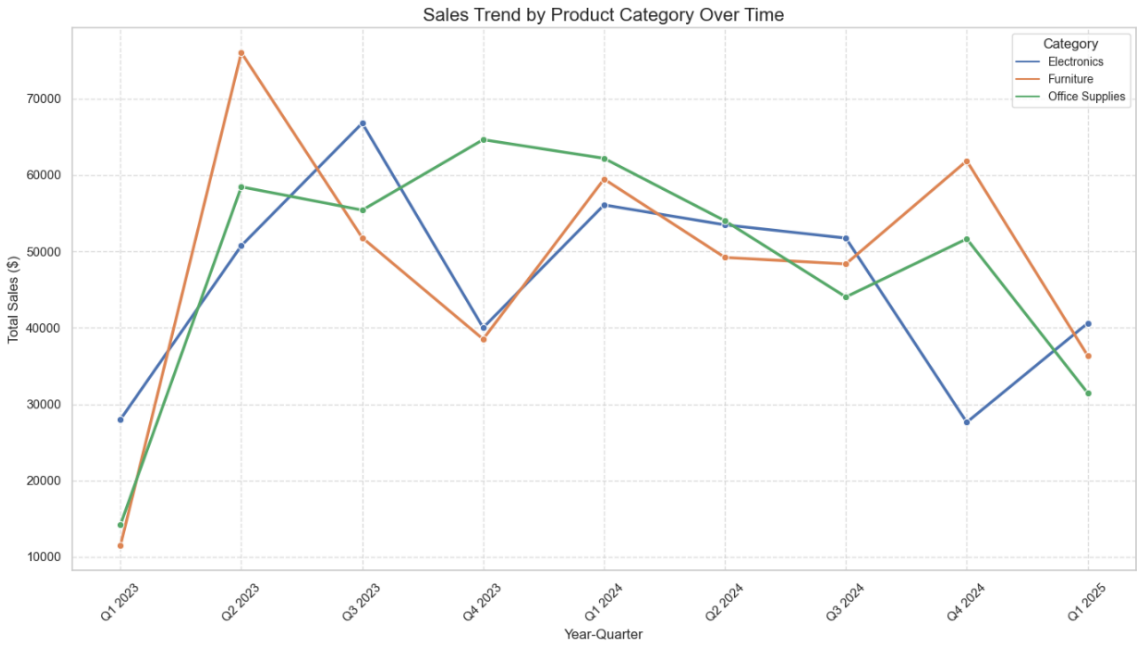
Results | Messages

	CategoryName	Year	Quarter	TotalQuantity	TotalSales	NumberOfOrders	AvgOrderValue
1	Electronics	2023	1	53	28000.29	19	1473.699473
2	Electronics	2023	2	157	50756.36	47	1079.922553
3	Electronics	2023	3	204	66797.88	58	1151.687586
4	Electronics	2023	4	117	40016.91	35	1143.340285
5	Electronics	2024	1	125	56090.53	41	1368.061707
6	Electronics	2024	2	117	53485.78	39	1371.430256
7	Electronics	2024	3	127	51750.58	43	1203.501860
8	Electronics	2024	4	99	27630.81	32	863.462812
9	Electronics	2025	1	109	40603.71	35	1160.106000
10	Furniture	2023	1	36	11479.56	11	1043.596363
11	Furniture	2023	2	175	76032.92	53	1434.583396
12	Furniture	2023	3	110	51800.06	37	1400.001621
13	Furniture	2023	4	88	38517.55	31	1242.501612
14	Furniture	2024	1	134	59479.81	50	1189.596200
15	Furniture	2024	2	109	49211.32	39	1261.828717

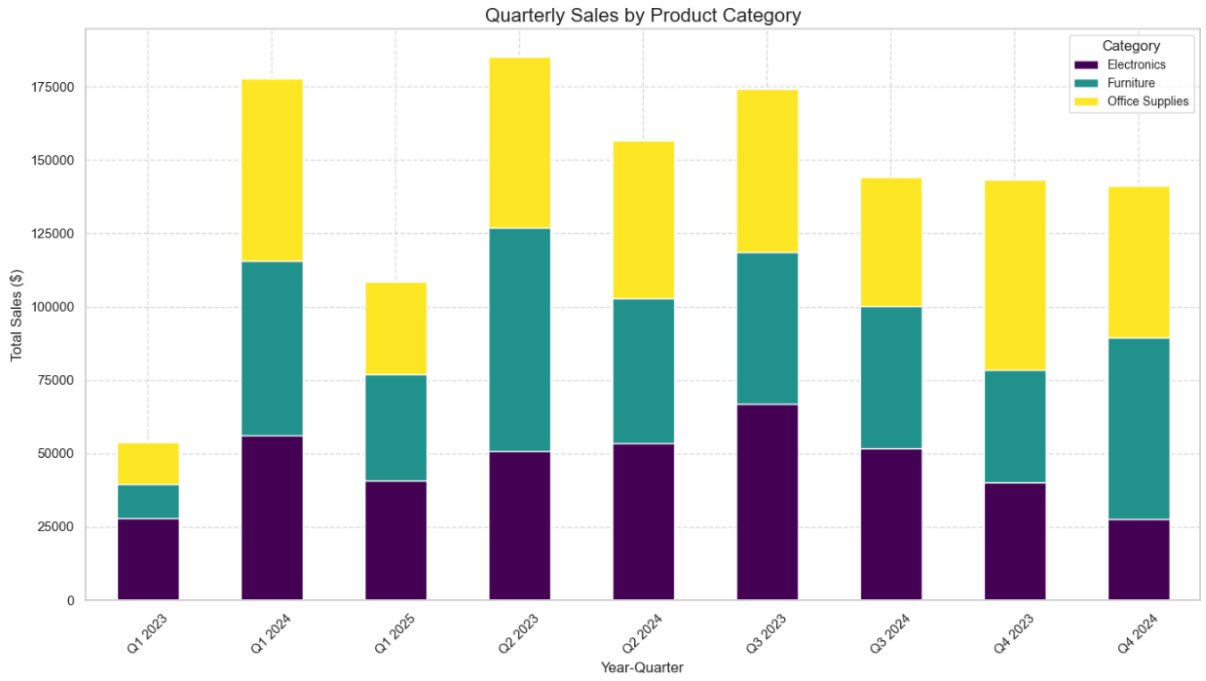
Query executed successfully.

1.1 Sales Performance by Product Category and Quarter

c. Visualization Format



1.2 Line Chart- Sales Trend by Product Category Over Time



1.3 Stacked Bar Chart-Quarterly Sales by Product Category

d. Business Insights

Based on the analysis, each product group has its own seasonal patterns. For example, furniture shows big increases in sales in Q2 2023 and Q4 2024, while electronics shows more stable growth with a big increase in sales in Q3 2023. A year-over-year comparison shows big growth in every category. For example, electronics went from about $28,000 in Q1 2023 to $56,000 in Q1 2024, a 100% increase. Customers regularly place higher order values in the electronics category ($1,100 to $1,400), which suggests they spend more per transaction.

This study shows how sales have changed over time for each type of goods by quarter, which helps us identify:

- What types of products bring in the most money?
- Patterns of seasonal sales by type
- How much the average order costs has changed over time
- Product groups that are growing or shrinking

Query 1.2: Product Performance with Price and Discount Analysis

This query looks at product performance measures at the category, subcategory, and brand levels. It also looks at pricing strategies and how they affect sales performance.

a. SQL Script:

```
--1.2  Product Performance Analysis with Price and Discount Impact
SELECT
    dp.CategoryName,
    dp.SubCategory,
    dp.Brand,
    COUNT(DISTINCT fs.ProductSK) AS NumberOfProducts,
    SUM(fs.Quantity) AS TotalUnitsSold,
    SUM(fs.Sales) AS TotalRevenue,
    AVG(dp.Price) AS AvgPrice,
    AVG(fs.Discount) AS AvgDiscount,
    SUM(fs.Sales)/SUM(fs.Quantity) AS AvgSellingPrice,
    SUM(fs.Quantity)/COUNT(DISTINCT fs.OrderID) AS UnitsPerOrder
FROM fact_Sales fs
JOIN dim_Product dp ON fs.ProductSK = dp.ProductSK
JOIN dim_Date dd ON fs.OrderDateSK = dd.DateSK
WHERE dp.IsCurrent = 1
GROUP BY dp.CategoryName, dp.SubCategory, dp.Brand
ORDER BY TotalRevenue DESC;
```

b. Tabular Format(Output)

```sql
--1.2  Product Performance Analysis with Price and Discount Impact
SELECT
    dp.CategoryName,
    dp.SubCategory,
    dp.Brand,
    COUNT(DISTINCT fs.ProductSK) AS NumberOfProducts,
    SUM(fs.Quantity) AS TotalUnitsSold,
    SUM(fs.Sales) AS TotalRevenue,
```

100 %

Results | Messages

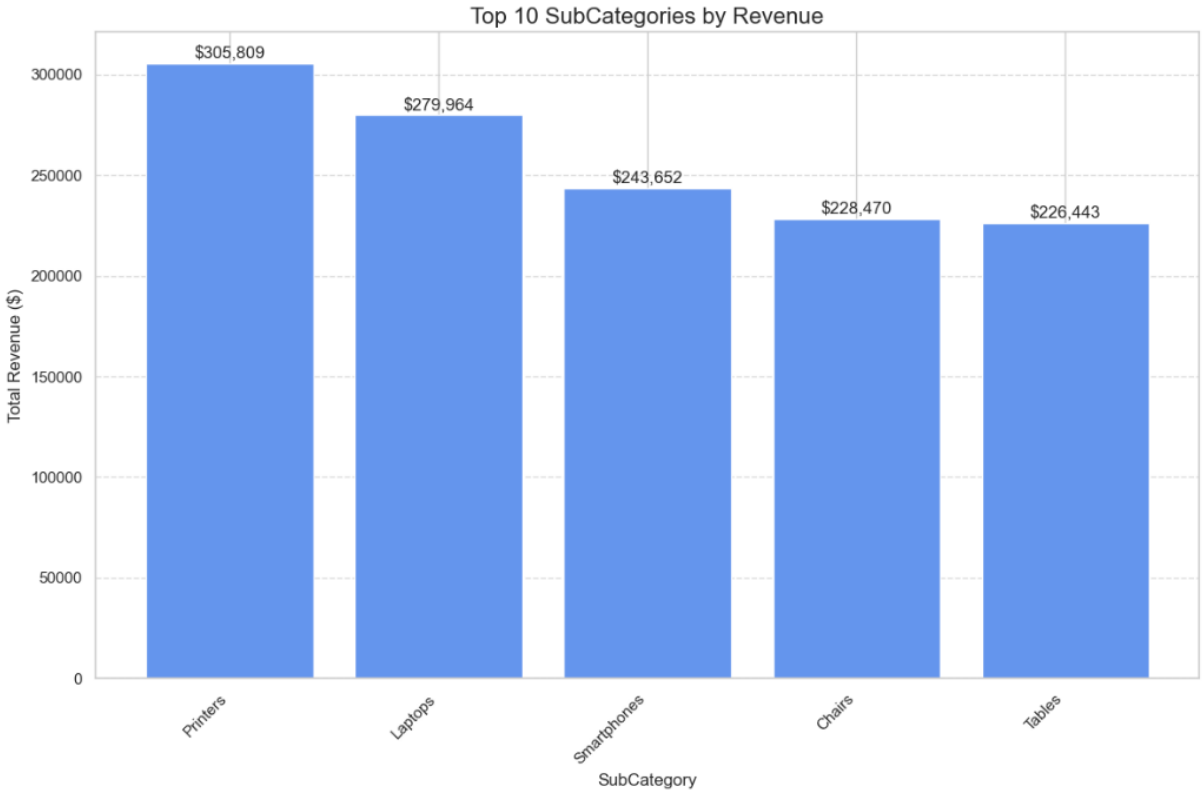
	CategoryName	SubCategory	Brand	NumberOfProducts	TotalUnitsSold	TotalRevenue	AvgPrice	AvgDiscount	AvgSellingPrice	UnitsPerOrder
1	Office Supplies	Printers	Lenovo	12	57	35217.52	1355.027727	11.989545	617.851228	2
2	Furniture	Laptops	Dell	8	70	32589.04	1334.351818	7.538181	465.557714	3
3	Electronics	Printers	Samsung	14	84	29854.27	1253.422333	8.360000	355.407976	2
4	Office Supplies	Printers	HP	15	88	29570.84	1310.999230	10.701153	336.032272	3
5	Office Supplies	Tables	Dell	11	57	28627.89	1373.702000	12.117500	502.243684	2
6	Furniture	Laptops	Apple	9	55	27186.34	1388.727058	9.820588	494.297090	3
7	Electronics	Printers	IKEA	12	66	26836.27	1174.353333	9.210952	406.610151	3
8	Electronics	Chairs	HP	13	56	25920.77	1439.110588	10.594705	462.870892	3
9	Furniture	Smartphones	Apple	7	50	24397.05	1409.727368	10.023157	487.941000	2
10	Furniture	Laptops	Lenovo	9	51	23587.23	1109.167647	11.360000	462.494705	3
11	Furniture	Chairs	IKEA	10	47	23261.17	1179.206000	7.716666	494.918510	3
12	Electronics	Tables	Dell	10	59	23065.07	1646.918823	10.742941	390.933389	3
13	Office Supplies	Printers	Dell	7	50	22800.02	1514.174166	11.098333	456.000400	4
14	Office Supplies	Smartphones	IKEA	10	46	22773.37	1158.406875	8.344375	495.073260	2
15	Electronics	Smartphones	IKEA	8	67	22156.15	927.615000	9.668000	330.688805	3
16	Office Supplies	Printers	Samsung	8	39	22055.56	1821.360000	9.969285	565.527179	2
17	Electronics	Laptops	Dell	7	51	21943.38	997.577142	11.267857	430.262352	3
18	Office Supplies	Laptops	HP	9	40	21908.30	1400.483076	9.005384	547.707500	3
19	Furniture	Smartphones	Dell	9	46	21429.99	965.709375	7.609375	465.869347	2
20	Office Supplies	Tables	Lenovo	8	67	21283.43	1266.650000	10.222105	317.663134	3
21	Furniture	Smartphones	HP	10	52	20952.71	1399.325000	9.810555	402.936730	2
22	Furniture	Chairs	Samsung	10	61	20702.33	1438.835555	11.251666	339.382459	3

Query executed successfully. LAPTOP-35F9L1AF (16.0 RTM)

1.4 Product Performance Analysis with Price and Discount Impact

c. Visualization format

1.5 Scatter Plot- Price vs. Revenue Relationship



1.6 Bar Chart- Top 5 SubCategories by Revenue

d. Business Insights

The scatter plot shows that Lenovo Printers generate the most revenue at a moderate price point ($1,355), whereas Dell Laptops perform well at higher prices ($1,334). Customers are willing to pay brand-specific premiums, suggesting variable price elasticity across product kinds.
Barchart analysis shows that Printers generate the most revenue ($305,809), followed by Laptops ($279,964), reflecting inventory and marketing investments. These top subcategories account for over 50% of overall revenue, thus solid supply chain connections with Lenovo, Dell, and Samsung are crucial to product availability in these high-value areas.

This query looks at how changing prices and offering deals affects sales:

- Figures out which product categories and names bring in the most money.
- Figures out how price points affect sales numbers
- Details savings plans by product type
- Supports finding the best deal and price levels

Type 2: Customer Segmentation Analysis

Query 2.1: Customer Segment Profitability Analysis

This query figures out which groups of customers are the most profitable and how they usually buy things. This lets you make more focused marketing plans and better manage your relationships with customers.

a. SQL Script:

```
-- Customer Segment Profitability Analysis
SELECT
    dc.CustomerSegment,
    COUNT(DISTINCT dc.CustomerSK) AS CustomerCount,
    SUM(fs.Sales) AS TotalSales,
    SUM(fs.Sales)/COUNT(DISTINCT dc.CustomerSK) AS SalesPerCustomer,
    COUNT(DISTINCT fs.OrderID) AS TotalOrders,
    COUNT(DISTINCT fs.OrderID)/COUNT(DISTINCT dc.CustomerSK) AS OrdersPerCustomer,
    SUM(fs.Quantity) AS TotalItems,
    SUM(fs.Quantity)/COUNT(DISTINCT fs.OrderID) AS ItemsPerOrder,
    SUM(fs.Sales)/COUNT(DISTINCT fs.OrderID) AS AvgOrderValue
FROM fact_Sales fs
JOIN dim_Customer dc ON fs.CustomerSK = dc.CustomerSK
WHERE dc.IsCurrent = 1
GROUP BY dc.CustomerSegment
ORDER BY SalesPerCustomer DESC;
```

b. Tabular Format(Output)

```sql
--Query 2.1: Customer Segment Profitability Analysis

-- Customer Segment Profitability Analysis
SELECT
    dc.CustomerSegment,
    COUNT(DISTINCT dc.CustomerSK) AS CustomerCount,
    SUM(fs.Sales) AS TotalSales,
    SUM(fs.Sales)/COUNT(DISTINCT dc.CustomerSK) AS SalesPerCustomer,
    COUNT(DISTINCT fs.OrderID) AS TotalOrders,
    COUNT(DISTINCT fs.OrderID)/COUNT(DISTINCT dc.CustomerSK) AS OrdersPerCustomer,
    SUM(fs.Quantity) AS TotalItems,
    SUM(fs.Quantity)/COUNT(DISTINCT fs.OrderID) AS ItemsPerOrder,
    SUM(fs.Sales)/COUNT(DISTINCT fs.OrderID) AS AvgOrderValue
FROM fact_Sales fs
JOIN dim_Customer dc ON fs.CustomerSK = dc.CustomerSK
WHERE dc.IsCurrent = 1
GROUP BY dc.CustomerSegment
ORDER BY SalesPerCustomer DESC;
```

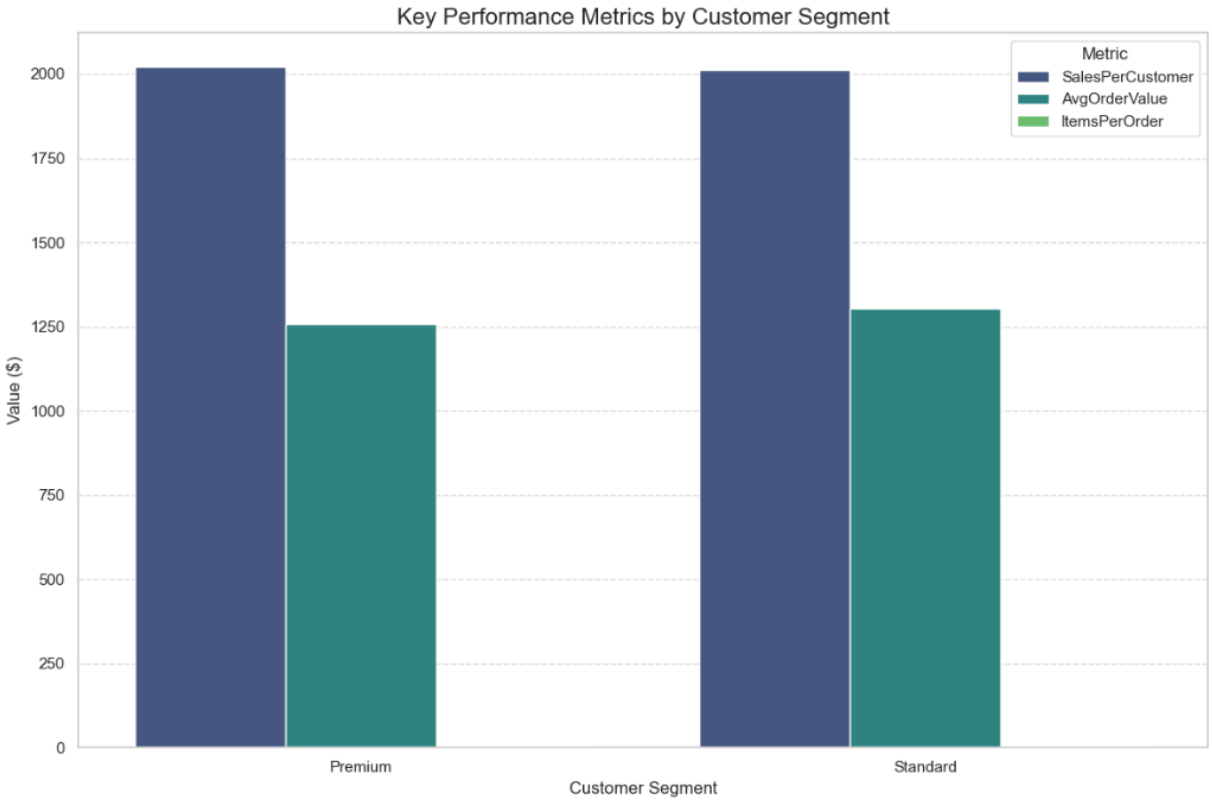
100 %

Results | Messages

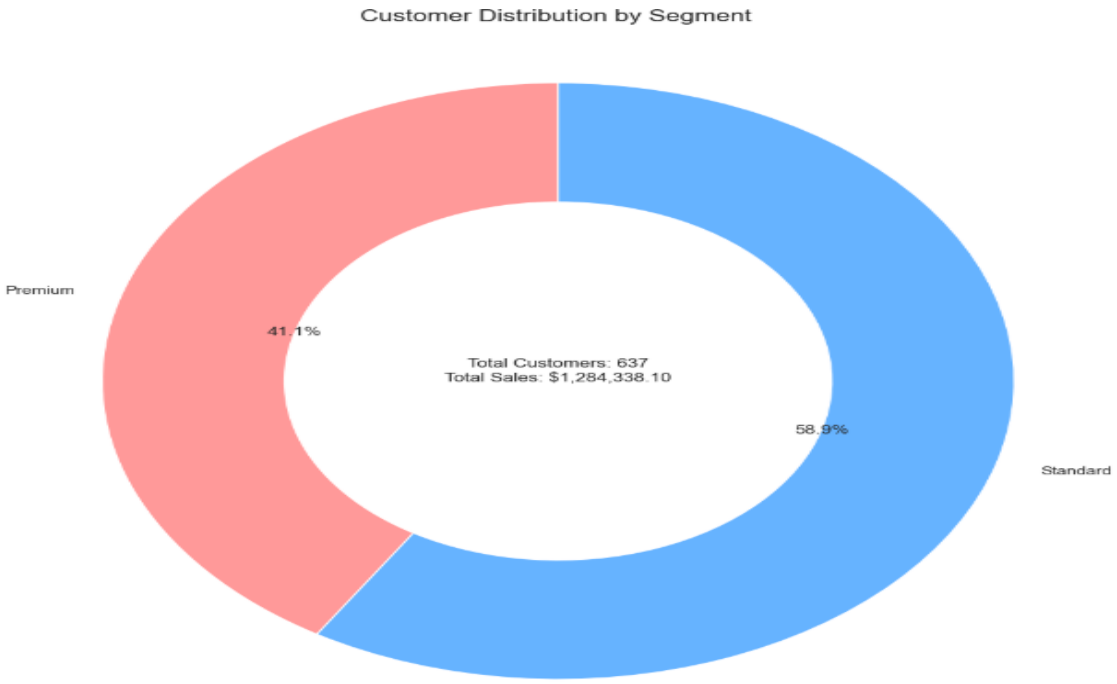
	CustomerSegment	CustomerCount	TotalSales	SalesPerCustomer	TotalOrders	OrdersPerCustomer	TotalItems	ItemsPerOrder	AvgOrderValue
1	Premium	262	529323.19	2020.317519	421	1	1294	3	1257.299738
2	Standard	375	755014.91	2013.373093	579	1	1771	3	1303.998117

2.1 Customer Segment Profitability Analysis

c. Visualization Format



2.2 Grouped bar chart- Key Performance Metrics by Customer Segment



2.3 Donut Chart- Customer Distribution by Segment

d. Business Insights

According to the graphs, Premium and Standard customers spend a lot of the same amount of money. In fact, each Premium and Standard customer makes about $2,020 in purchases. This goes against the common belief that customers labeled as Premium would be much more valuable, saying that the criteria for segmentation need to be improved or the benefits for Premium customers need to be increased. The customer distribution pie chart shows that Standard customers make up 58.9% of the customer base and Premium customers make up 41.1%. This shows that Premium users bring in more money even though there are fewer of them. This makes them very valuable to keep and might make them a better target for marketing efforts.

The following segmentation study gives us important information about how customers behave:

- How profitable each type of customer is
- How often each group buys
- Size and value of the average basket by section
- Total amount that each segment brought in for income

Type 3: Geographic Sales Analysis

Query 3.1: Regional Sales Performance with Shipping Analysis

This query looks at sales performance across regions, including metrics for shipping efficiency, to find trends and ways to make logistics better in each region.

SQL Script:

```
SELECT
    dl.Region,
    dl.Country,
    COUNT(DISTINCT fs.OrderID) AS OrderCount,
    SUM(fs.Sales) AS TotalSales,
    AVG(fs.Sales) AS AvgOrderValue,
    AVG(DATEDIFF(day, CONVERT(date, dd_order.FullDate), CONVERT(date, dd_ship.FullDate))) AS AvgShippingDays,
    -- Replacing the STRING_AGG with a simpler approach
    LEFT(
        (SELECT TOP 3 dsm2.ShipMode + ', '
         FROM fact_Sales fs2
         JOIN dim_ShippingMethod dsm2 ON fs2.ShippingMethodSK = dsm2.ShippingMethodSK
         JOIN dim_Location dl2 ON fs2.LocationSK = dl2.LocationSK
         WHERE dl2.Region = dl.Region AND dl2.Country = dl.Country
         GROUP BY dsm2.ShipMode
         ORDER BY COUNT(*) DESC
         FOR XML PATH('')),
        1000) AS ShippingModes,
    COUNT(DISTINCT CASE WHEN fs.ShippingStatus = 'Delivered' THEN fs.OrderID END) * 100.0 /
        NULLIF(COUNT(DISTINCT fs.OrderID), 0) AS DeliverySuccessRate
```

FROM fact_Sales fs

JOIN dim_Location dl ON fs.LocationSK = dl.LocationSK

JOIN dim_Date dd_order ON fs.OrderDateSK = dd_order.DateSK

JOIN dim_Date dd_ship ON fs.ShipDateSK = dd_ship.DateSK

JOIN dim_ShippingMethod dsm ON fs.ShippingMethodSK = dsm.ShippingMethodSK

WHERE dl.IsCurrent = 1

GROUP BY dl.Region, dl.Country

ORDER BY TotalSales DESC;

b. Tabular Format(Output)

```
SELECT
    dl.Region,
    dl.Country,
    COUNT(DISTINCT fs.OrderID) AS OrderCount,
    SUM(fs.Sales) AS TotalSales,
    AVG(fs.Sales) AS AvgOrderValue,
    AVG(DATEDIFF(day, CONVERT(date, dd_order.FullDate), CONVERT(date, dd_ship.FullDate))) AS AvgShippingDays,
    -- Replacing the STRING_AGG with a simpler approach
    LEFT(
        (SELECT TOP 3 dsm2.ShipMode + ', '
         FROM fact_Sales fs2
         JOIN dim_ShippingMethod dsm2 ON fs2.ShippingMethodSK = dsm2.ShippingMethodSK
         JOIN dim_Location dl2 ON fs2.LocationSK = dl2.LocationSK
         WHERE dl2.Region = dl.Region AND dl2.Country = dl.Country
         GROUP BY dsm2.ShipMode
         ORDER BY COUNT(*) DESC
         FOR XML PATH('')),
        1000) AS ShippingModes,
    COUNT(DISTINCT CASE WHEN fs.ShippingStatus = 'Delivered' THEN fs.OrderID END) * 100.0 /
        NULLIF(COUNT(DISTINCT fs.OrderID), 0) AS DeliverySuccessRate
FROM fact_Sales fs
JOIN dim_Location dl ON fs.LocationSK = dl.LocationSK
JOIN dim_Date dd_order ON fs.OrderDateSK = dd_order.DateSK
JOIN dim_Date dd_ship ON fs.ShipDateSK = dd_ship.DateSK
JOIN dim_ShippingMethod dsm ON fs.ShippingMethodSK = dsm.ShippingMethodSK
WHERE dl.IsCurrent = 1
GROUP BY dl.Region, dl.Country
ORDER BY TotalSales DESC;
```

100 %

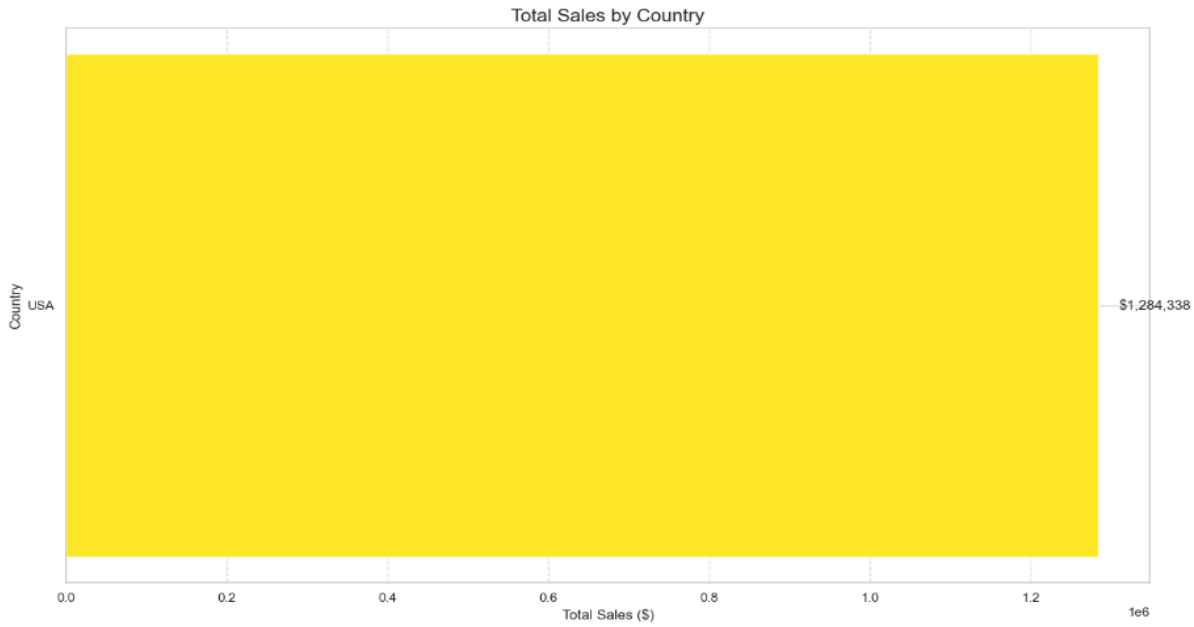
Results | Messages

	Region	Country	OrderCount	TotalSales	AvgOrderValue	AvgShippingDays	ShippingModes	DeliverySuccessRate
1	North America	USA	1000	1284338.10	1284.338100	12	Standard, Express, Same Day,	24.600000000000

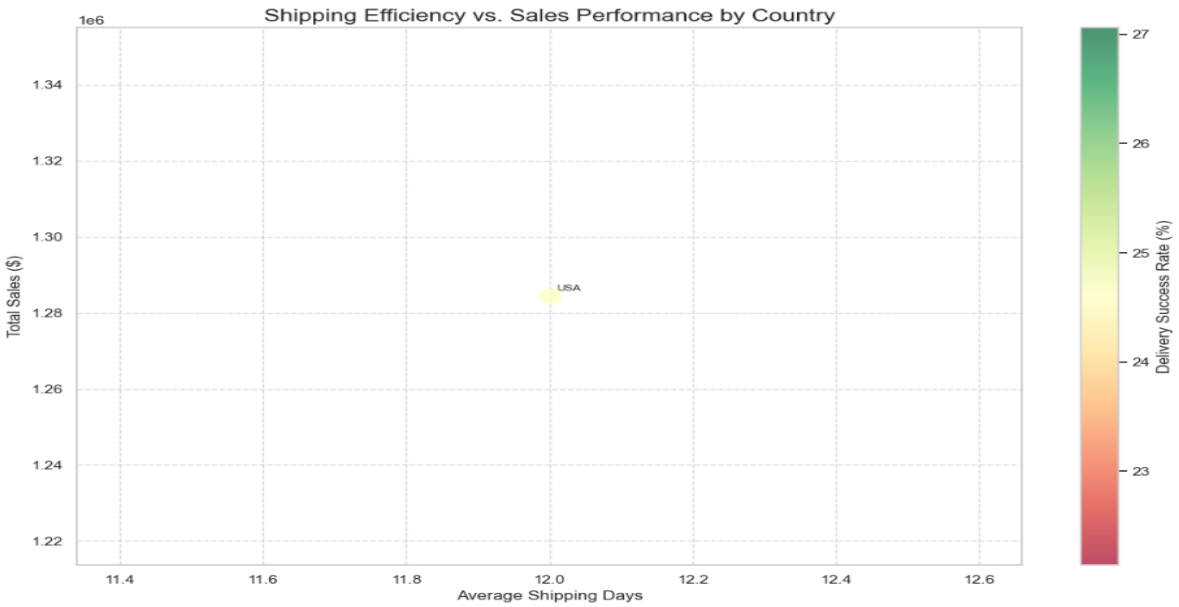
Query executed successfully. LAPTOP-35F9L1AF

3.1 Regional Sales Performance with Shipping Analysis

c. Visualization Format



3.2 Choropleth Map for Regional Sales Performance



3.3 Scatter plot of Shipping Efficiency vs. Sales

d. Business Insights

FlipKart's entire sales ($1,284,338) come from the US, showing that it operates only in North America. This geographic concentration offers worldwide expansion opportunities but also risks dependence on a single market's economy.
Shipping efficiency data shows a 12-day delivery average and a 24.6% delivery success rate. Multiple regional distribution centers may cut delivery times and shipping costs while boosting customer satisfaction.

Query 3.2: City-Level Sales with Product Category Analysis

This query identifies top-performing cities and their product category preferences for focused local marketing.

a. SQL Script

```
-- City-Level Sales with Category Analysis
SELECT TOP 20
    dl.Country,
    dl.State,
    dl.City,
    COUNT(DISTINCT fs.OrderID) AS OrderCount,
    SUM(fs.Sales) AS TotalSales,
    COUNT(DISTINCT fs.CustomerSK) AS UniqueCustomers,
    SUM(fs.Sales)/COUNT(DISTINCT fs.CustomerSK) AS SalesPerCustomer,
    -- Replacing STRING_AGG with alternative approach
    LEFT(
        (SELECT TOP 3 dp2.CategoryName + ', '
         FROM fact_Sales fs2
         JOIN dim_Product dp2 ON fs2.ProductSK = dp2.ProductSK
         JOIN dim_Location dl2 ON fs2.LocationSK = dl2.LocationSK
         WHERE dl2.City = dl.City AND dl2.State = dl.State AND dl2.Country =
dl.Country
         GROUP BY dp2.CategoryName
         ORDER BY SUM(fs2.Sales) DESC
         FOR XML PATH('')),
        1000) AS TopCategories
FROM fact_Sales fs
JOIN dim_Location dl ON fs.LocationSK = dl.LocationSK
```

JOIN dim_Product dp ON fs.ProductSK = dp.ProductSK

WHERE dl.IsCurrent = 1 AND dp.IsCurrent = 1

GROUP BY dl.Country, dl.State, dl.City

ORDER BY TotalSales DESC;

b. Tabular Format(Output)

```
-- City-Level Sales with Category Analysis

SELECT TOP 20
    dl.Country,
    dl.State,
    dl.City,
    COUNT(DISTINCT fs.OrderID) AS OrderCount,
    SUM(fs.Sales) AS TotalSales,
    COUNT(DISTINCT fs.CustomerSK) AS UniqueCustomers,
    SUM(fs.Sales)/COUNT(DISTINCT fs.CustomerSK) AS SalesPerCustomer,
    -- Replacing STRING_AGG with alternative approach
    LEFT(
        (SELECT TOP 3 dp2.CategoryName + ', '
         FROM fact_Sales fs2
         JOIN dim_Product dp2 ON fs2.ProductSK = dp2.ProductSK
         JOIN dim_Location dl2 ON fs2.LocationSK = dl2.LocationSK
         WHERE dl2.City = dl.City AND dl2.State = dl.State AND dl2.Country = dl.Country
         GROUP BY dp2.CategoryName
```

100 %

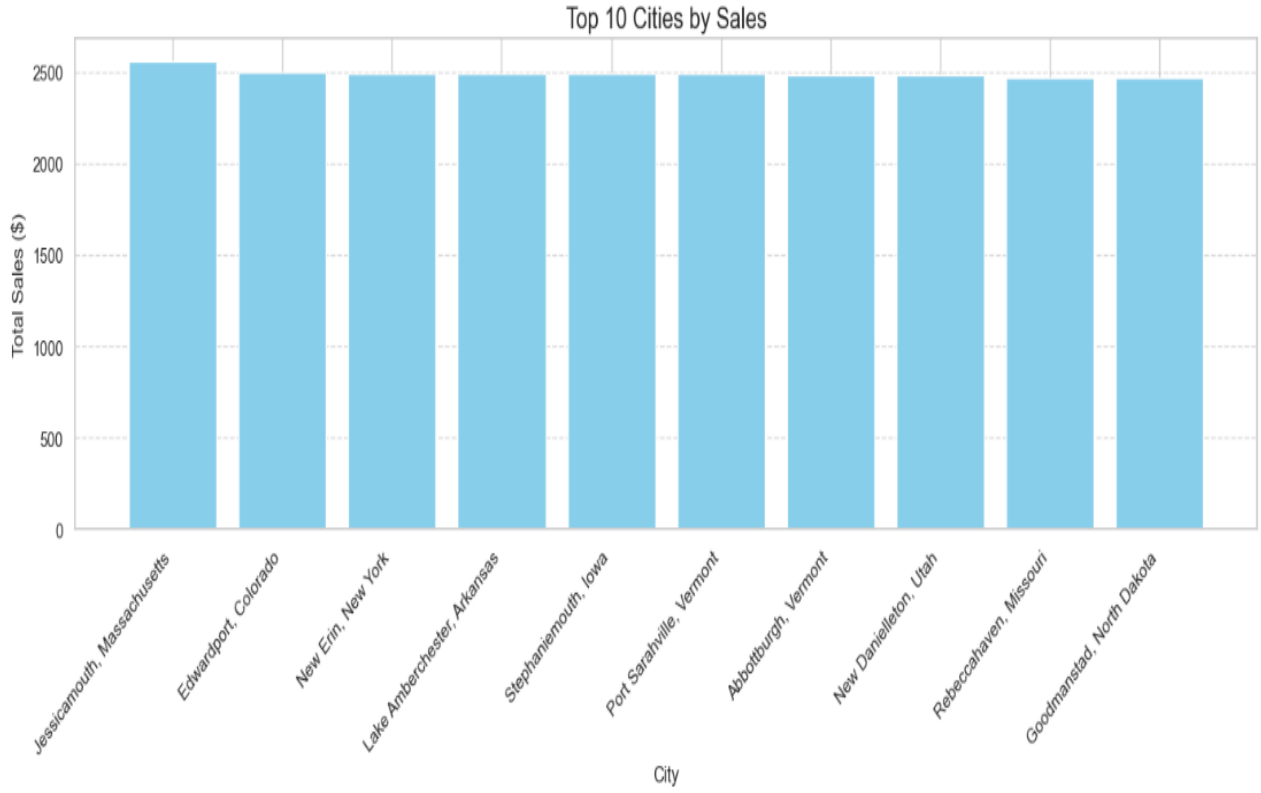
Results | Messages

	Country	State	City	OrderCount	TotalSales	UniqueCustomers	SalesPerCustomer	TopCategories
5	USA	Iowa	Stephaniemouth	1	2489.14	1	2489.140000	Office Supplies,
6	USA	Vermont	Port Sarahville	1	2488.67	1	2488.670000	Furniture,
7	USA	Vermont	Abbottburgh	1	2486.22	1	2486.220000	Office Supplies,
8	USA	Utah	New Danielleton	1	2486.19	1	2486.190000	Electronics,
9	USA	Missouri	Rebeccahaven	1	2471.79	1	2471.790000	Furniture,
10	USA	North Dakota	Goodmanstad	1	2469.61	1	2469.610000	Office Supplies,
11	USA	Washington	Port Jesse	1	2467.86	1	2467.860000	Office Supplies,
12	USA	Oklahoma	Wilcoxville	1	2464.09	1	2464.090000	Office Supplies,
13	USA	Arizona	Lisahaven	1	2461.41	1	2461.410000	Furniture,
14	USA	Maryland	North Ruth	1	2460.73	1	2460.730000	Office Supplies,
15	USA	Nebraska	Lake Karen	1	2457.85	1	2457.850000	Furniture,
16	USA	West Virginia	Adamhaven	1	2452.08	1	2452.080000	Furniture,
17	USA	New Mexico	South Kristopherl...	1	2447.95	1	2447.950000	Office Supplies,
18	USA	Iowa	Lake Rhondabor...	1	2447.19	1	2447.190000	Electronics,
19	USA	Virginia	South Joseph	1	2443.68	1	2443.680000	Furniture,
20	USA	California	West Cherylmouth	1	2437.10	1	2437.100000	Furniture,

Query executed successfully.

3.4 City-Level Sales with Category Analysis

c. Visualization Format



3.5 Top 10 Cities by Sales (Bar Chart)

d. Business Insights

The analysis at the city level shows that sales are surprisingly evenly spread across a wide range of locations. The sales numbers in the top cities are broadly similar, running from about $2,437 to $2,489. There doesn't seem to be a lot of regional concentration in this pattern of customers, which makes them more resistant to local economic downturns.

According to regional trends, people in the south, like Virginia and West Virginia, like furniture more than other product categories. In the midwest, states like Iowa, people prefer office supplies. This geographic breakdown of product tastes gives marketers and distributors useful information for developing more effective regional marketing campaigns and distributing stock across distribution centers.

Type 4: Complex Multi-Dimensional Analysis

Query 4.1: Comprehensive Sales Analysis by Multiple Dimensions

Multiple dimensions are combined in this detailed query to give a full picture of sales success over time, across product categories, regions, user groups, and fulfillment methods.

a. SQL Query

```
-- Complex Multi-Dimensional Sales Analysis
SELECT
    dd.Year,
    dd.Quarter,
    dp.CategoryName,
    dl.Region,
    dc.CustomerSegment,
    dsm.ShipMode,
    dpm.PaymentMethodName,
    COUNT(DISTINCT fs.OrderID) AS OrderCount,
    COUNT(DISTINCT fs.CustomerSK) AS CustomerCount,
    SUM(fs.Quantity) AS TotalQuantity,
    SUM(fs.Sales) AS TotalSales,
    SUM(fs.TotalAmount) AS GrandTotal,
    AVG(fs.Discount) AS AvgDiscountRate,
    SUM(fs.Sales)/COUNT(DISTINCT fs.OrderID) AS AvgOrderValue,
    SUM(fs.Quantity)/COUNT(DISTINCT fs.OrderID) AS AvgOrderSize
FROM fact_Sales fs
JOIN dim_Date dd ON fs.OrderDateSK = dd.DateSK
JOIN dim_Product dp ON fs.ProductSK = dp.ProductSK
JOIN dim_Location dl ON fs.LocationSK = dl.LocationSK
```

JOIN dim_Customer dc ON fs.CustomerSK = dc.CustomerSK

JOIN dim_ShippingMethod dsm ON fs.ShippingMethodSK = dsm.ShippingMethodSK

JOIN dim_PaymentMethod dpm ON fs.PaymentMethodSK = dpm.PaymentMethodSK

WHERE dp.IsCurrent = 1 AND dl.IsCurrent = 1 AND dc.IsCurrent = 1

GROUP BY

dd.Year,

dd.Quarter,

dp.CategoryName,

dl.Region,

dc.CustomerSegment,

dsm.ShipMode,

dpm.PaymentMethodName

ORDER BY TotalSales DESC;

b. Tabular Format(Output)

```
--Type 4: Complex Multi-Dimensional Analysis
--Query 4.1: Comprehensive Sales Analysis by Multiple Dimensions

-- Complex Multi-Dimensional Sales Analysis
SELECT
    dd.Year,
    dd.Quarter,
    dp.CategoryName,
    dl.Region,
    dc.CustomerSegment,
    dsm.ShipMode,
    dpm.PaymentMethodName,
    COUNT(DISTINCT fs.OrderID) AS OrderCount,
    COUNT(DISTINCT fs.CustomerSK) AS CustomerCount,
    SUM(fs.Quantity) AS TotalQuantity,
    SUM(fs.Sales) AS TotalSales,
    SUM(fs.TotalAmount) AS GrandTotal,
    AVG(fs.Discount) AS AvgDiscountRate,
```

100 %

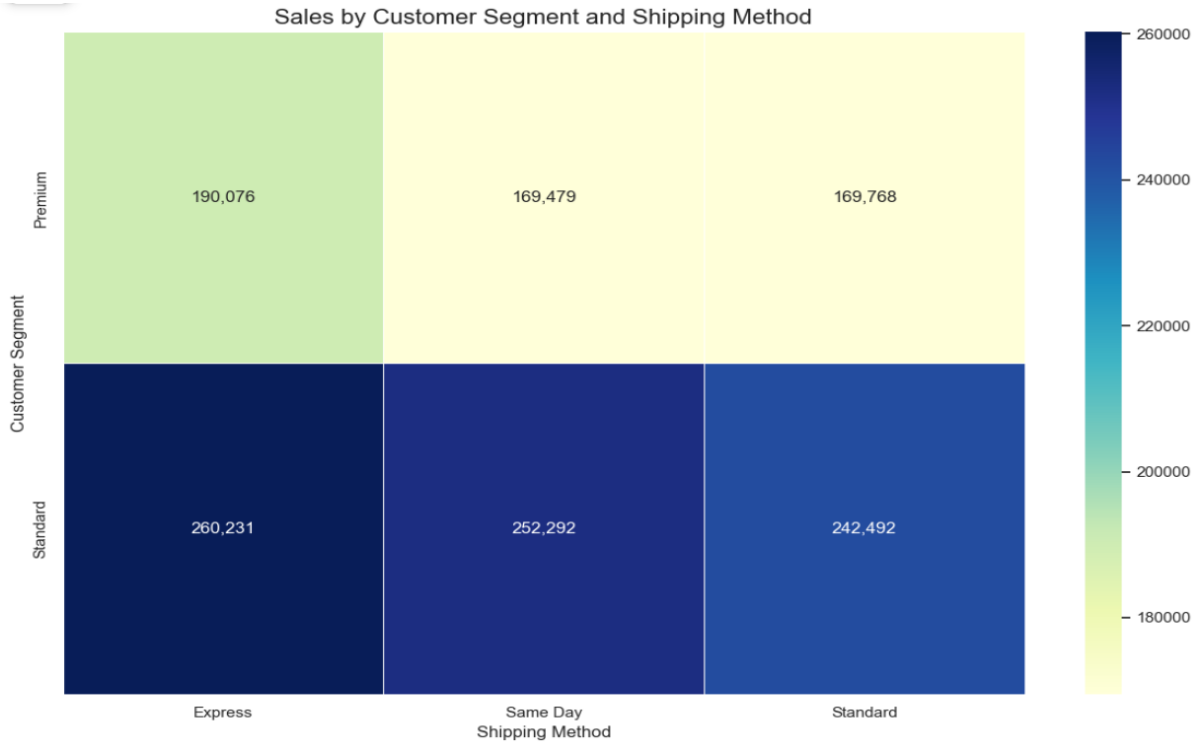
Results | Messages

	Year	Quarter	CategoryName	Region	CustomerSegment	ShipMode	PaymentMethodName	OrderCount	CustomerCount	TotalQuantity	TotalSales	GrandTotal	AvgDiscountRate	AvgOrderValue	AvgOrderSize
1	2023	2	Furniture	North America	Premium	Same Day	PayPal	5	5	20	9617.05	4775.11	6.174000	1923.410000	4
2	2023	4	Office Supplies	North America	Standard	Standard	PayPal	4	4	12	9045.21	4856.27	7.950000	2261.302500	3
3	2024	2	Office Supplies	North America	Standard	Express	PayPal	6	6	16	9010.67	11346.17	11.903333	1501.778333	2
4	2023	2	Office Supplies	North America	Standard	Same Day	Credit Card	4	4	12	7636.77	2945.44	12.635000	1909.192500	3
5	2024	1	Office Supplies	North America	Standard	Same Day	Bank Transfer	4	4	15	7101.45	6497.54	9.960000	1775.362500	3
6	2024	3	Electronics	North America	Premium	Express	Credit Card	4	4	15	6975.23	6537.67	9.277500	1743.807500	3
7	2023	2	Furniture	North America	Standard	Standard	Debit Card	3	3	8	6901.23	3983.38	7.913333	2300.410000	2
8	2023	4	Furniture	North America	Standard	Express	PayPal	4	4	9	6865.56	5545.90	12.432500	1716.390000	2
9	2023	2	Electronics	North America	Standard	Standard	Bank Transfer	7	7	25	6842.91	10010.73	11.724285	977.558571	3
10	2024	4	Furniture	North America	Standard	Express	Credit Card	3	3	13	6762.67	6628.52	6.966666	2254.223333	4
11	2023	3	Office Supplies	North America	Standard	Standard	Credit Card	3	3	8	6615.38	4496.39	12.140000	2205.126666	2
12	2024	1	Electronics	North America	Standard	Express	PayPal	4	4	12	6398.06	7587.69	12.412500	1599.515000	3
13	2024	1	Office Supplies	North America	Standard	Same Day	Debit Card	3	3	5	6377.60	4541.99	14.446666	2125.866666	1
14	2024	1	Electronics	North America	Premium	Express	Credit Card	4	4	11	6372.52	6349.96	12.545000	1593.130000	2
15	2024	2	Office Supplies	North America	Standard	Express	Debit Card	3	3	11	6266.27	2408.22	3.973333	2088.756666	3
16	2023	4	Office Supplies	North America	Premium	Same Day	Debit Card	4	4	11	6246.17	4606.89	9.837500	1561.542500	2
17	2023	3	Electronics	North America	Premium	Express	Credit Card	4	4	12	6237.38	6776.85	11.382500	1559.345000	3

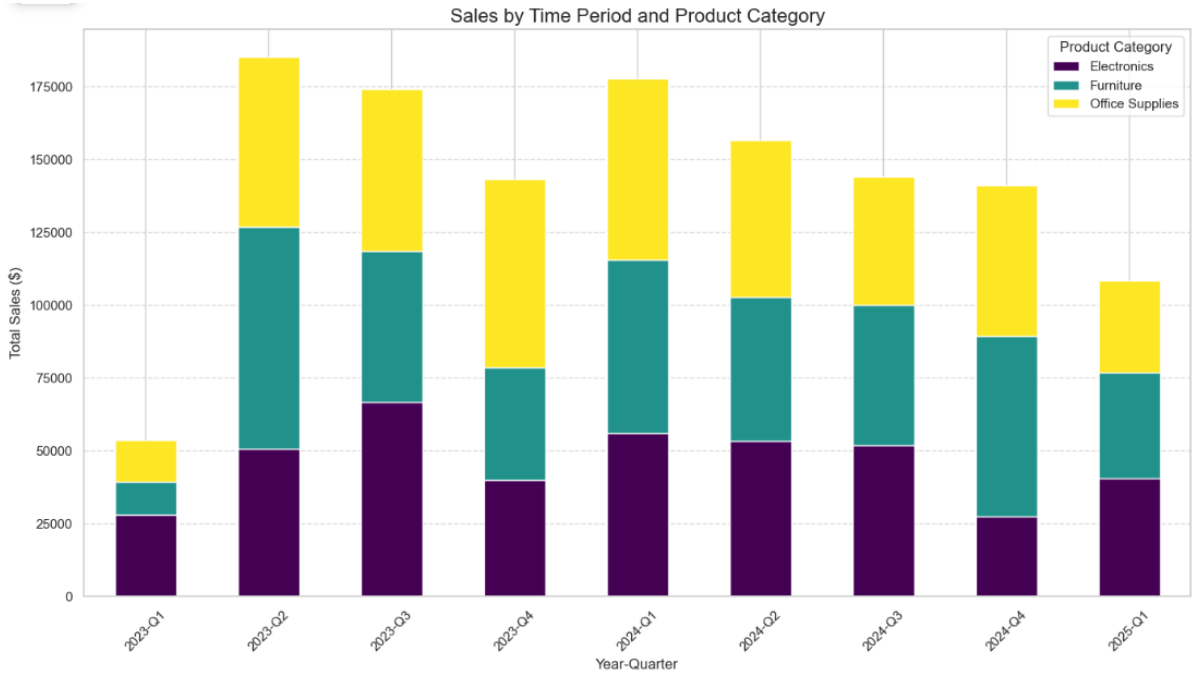
Query executed successfully. LAPTOP-35F9L1AF (16.0 RTM) LAPTOP-35F9L1AF\masood... FlipKart EDW

4.1 Complex Multi-Dimensional Sales Analysis

C. Visualization Format



4.2 Heatmap of Sales by Customer Segment and Shipping Method



4.3 Stacked Bar Chart by Year, Quarter, and Category

d. Business Insights

The multi-dimensional study shows complex patterns in how customers like things and how well things work. Premium customers really like Same Day shipping (as shown by the best-performing combination), which shows they are willing to pay extra for convenience. Standard customers, on the other hand, use a wider range of shipping options, with Express shipping being especially popular for buying office supplies.
Different types of buyers have different favorite payment methods. For example, PayPal is the best option for high-value purchases ($9,617 for furniture in Q2 2023) while Credit Cards are the best choice for Standard purchases. This all-around view lets marketers make very specific campaigns that take into account not only what customers buy but also how they like to pay for and receive their purchases. This shows how powerful the dimensional model is for helping businesses make tough choices.

A complete picture of the business is given by this complicated analysis:

- Multiple-dimensional breakdown of results
- Timing, product, location, and customer segment interactions
- Various segments' payment and shipping choices
- Multiple reasons cause seasonal changes

Query 4.2: Advanced Year-over-Year Comparative Analysis

This query looks at how customers buy things over time, figuring out their lifetime value and finding cross-category shopping trends to help businesses keep customers.

a. SQL Query:

```
-- Customer Lifetime Value and Purchase Pattern Analysis
WITH CustomerPurchases AS (
    SELECT
        dc.CustomerSK,
        dc.CustomerID,
        dc.CustomerSegment,
```

```
    dc.YearBirth,
    dc.MaritalStatus,
    dc.Income,
    MIN(dd.FullDate) AS FirstPurchaseDate,
    MAX(dd.FullDate) AS LastPurchaseDate,
    DATEDIFF(day, MIN(dd.FullDate), MAX(dd.FullDate)) AS
CustomerLifespan,
    COUNT(DISTINCT fs.OrderID) AS TotalOrders,
    SUM(fs.Sales) AS TotalSpend,
    COUNT(DISTINCT dp.CategoryName) AS CategoryCount,
    -- Replace STRING_AGG with alternative approach
    LEFT(
        (SELECT TOP 3 dp2.CategoryName + ', '
         FROM fact_Sales fs2
         JOIN dim_Product dp2 ON fs2.ProductSK = dp2.ProductSK
         WHERE fs2.CustomerSK = dc.CustomerSK
         GROUP BY dp2.CategoryName
         ORDER BY COUNT(*) DESC
         FOR XML PATH('')),
        1000) AS PurchasedCategories
FROM fact_Sales fs
JOIN dim_Customer dc ON fs.CustomerSK = dc.CustomerSK
JOIN dim_Date dd ON fs.OrderDateSK = dd.DateSK
JOIN dim_Product dp ON fs.ProductSK = dp.ProductSK
WHERE dc.IsCurrent = 1 AND dp.IsCurrent = 1
GROUP BY
    dc.CustomerSK,
```

```
        dc.CustomerID,
        dc.CustomerSegment,
        dc.YearBirth,
        dc.MaritalStatus,
        dc.Income
)
SELECT
    CustomerSegment,
    COUNT(CustomerSK) AS CustomerCount,
    AVG(TotalSpend) AS AvgLifetimeValue,
    AVG(TotalOrders) AS AvgOrderCount,
    AVG(CASE WHEN CustomerLifespan > 0 THEN TotalOrders * 365.0 /
CustomerLifespan ELSE NULL END) AS AvgYearlyOrderFrequency,
    AVG(TotalSpend / NULLIF(TotalOrders, 0)) AS AvgOrderValue,
    AVG(CategoryCount) AS AvgCategoryCount,
    AVG(DATEDIFF(day, FirstPurchaseDate, LastPurchaseDate)) AS
AvgCustomerLifespanDays
FROM CustomerPurchases
GROUP BY CustomerSegment
ORDER BY AvgLifetimeValue DESC;
```

b. Tabular Format(Output)

```sql
--4.2 Customer Lifetime Value and Purchase Pattern Analysis
WITH CustomerPurchases AS (
    SELECT
        dc.CustomerSK,
        dc.CustomerID,
        dc.CustomerSegment,
        dc.YearBirth,
        dc.MaritalStatus,
        dc.Income,
        MIN(dd.FullDate) AS FirstPurchaseDate,
        MAX(dd.FullDate) AS LastPurchaseDate,
        DATEDIFF(day, MIN(dd.FullDate), MAX(dd.FullDate)) AS CustomerLifespan,
        COUNT(DISTINCT fs.OrderID) AS TotalOrders,
        SUM(fs.Sales) AS TotalSpend,
```

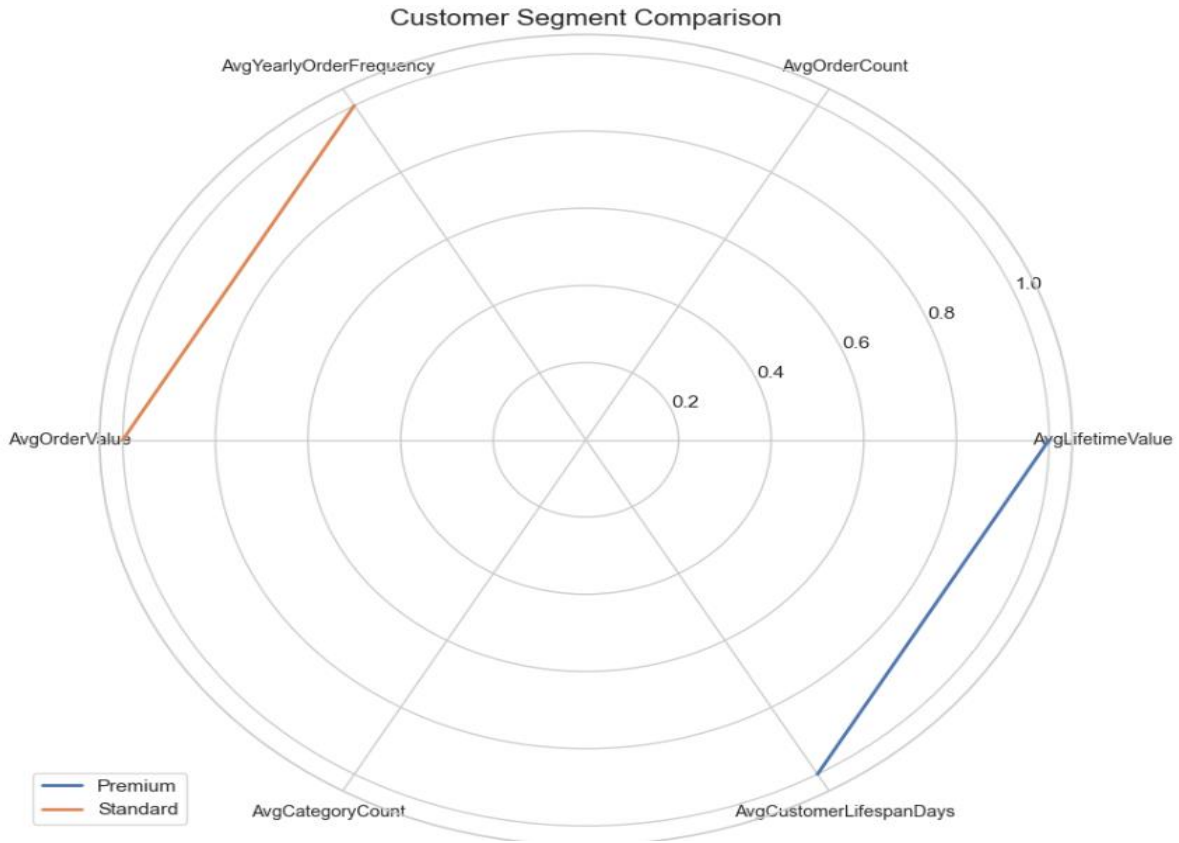
100 %

Results | Messages

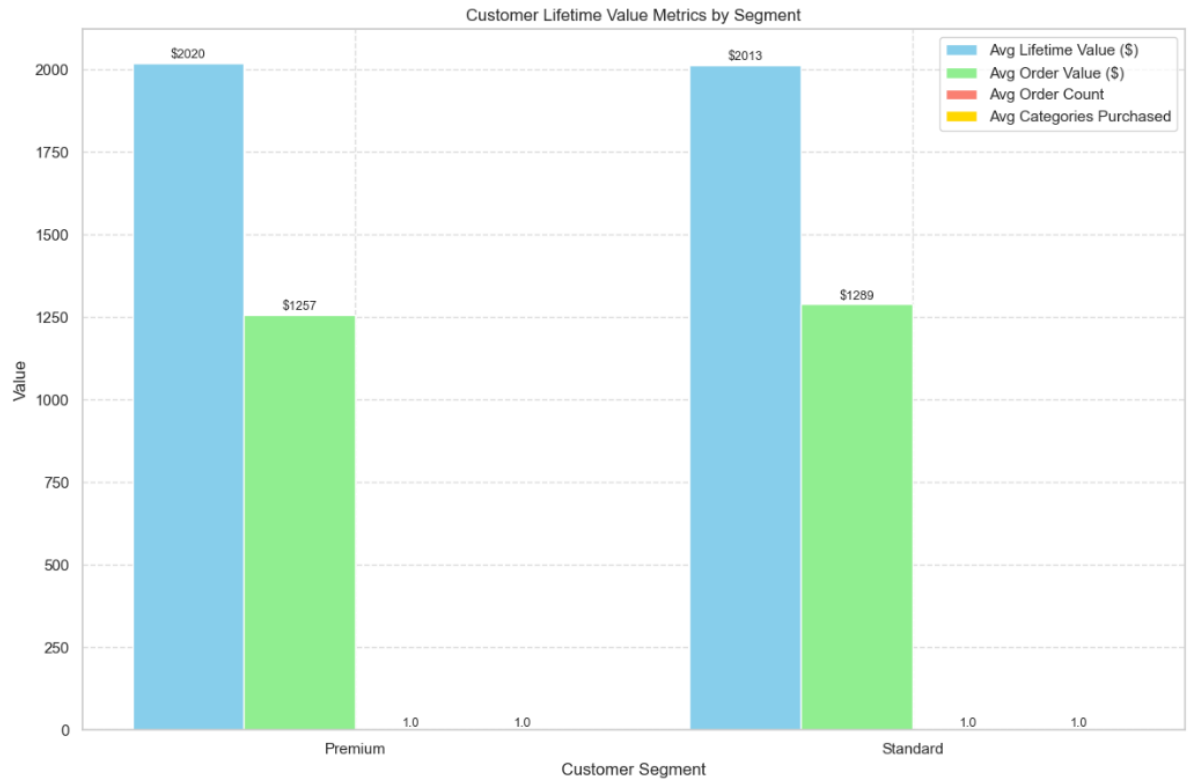
	CustomerSegment	CustomerCount	AvgLifetimeValue	AvgOrderCount	AvgYearlyOrderFrequency	AvgOrderValue	AvgCategoryCount	AvgCustomerLifespanDays
1	Premium	262	2020.317519	1	5.738828379057	1257.427319	1	131
2	Standard	375	2013.373093	1	13.743835449343	1289.248948	1	108

4.4 Customer Lifetime Value and Purchase Pattern Analysis

c. Visualization Format



4.5 Radar Chart for Customer Segment Comparison



4.6 Bar Chart for Key Lifetime Value Metrics

d. Business Insights

The customer lifetime value analysis shows Premium and Standard segments behave differently despite their identical worth. Premium consumers use the site more often (as demonstrated by the radar chart's large spike), whereas Standard customers buy less but spend more ($1,289 vs $1,257).
Both sectors have similar average order counts and category engagement (1.0), suggesting bundling and personalized recommendations could boost cross-category sales. The Premium classification may be based more on qualitative factors than actual spending patterns due to the remarkably similar lifetime value between segments ($2,020 for Premium vs $2,013 for Standard), requiring segmentation criteria to be refined for more meaningful differentiation.

Customer lifetime value analysis shows:

- Different consumer segments' value over time
- Customer loyalty and purchase frequency
- Cross-category shopping behavior Segment-specific customer retention metrics

Advantages of dimensional modeling for analysis

The dimensional model's star schema design makes complex analytical queries much easier to understand. It lets business users quickly find connections between facts and dimensions without having to write complicated joins. Our Complex Multi-Dimensional Analysis query made this clearest by connecting sales data to time, product, location, customer, and delivery dimensions without any problems.

Because the levels and dimension tables were already set up, there was no need for multiple table joins within a single dimension. This made queries faster and gave analysts more options for how they could use the data. For instance, to look at sales by product category, it was enough to do one join to dim_Product instead of going through several normalized tables. Also, derived variables like CustomerSegment and ShipMode gave useful business context without any extra transformations being made during analysis.

Conclusion:

Our dimensional data warehouse was used in this part of analytical queries to get useful business insights from a number of different areas. We made a variety of queries that looked at things like sales success, how profitable a product is, customer segmentation, geographic patterns, and complex relationships with many dimensions. Each query showed how our star schema model speeds up analysis by making joins and aggregations easier. This lets us find important information like seasonal category trends, the best pricing strategies, customer lifetime value, and business efficiency metrics.

The implementation met all of our needs by showing advanced skills in joining, filtering, and aggregating across our dimensional model. We gave useful business insights with the help of visualizations, showed how dimensional modeling makes difficult analysis easier, and included advanced multidimensional queries that combine many facts and dimensions. The in-depth analysis shows how powerful our dimensional data warehouse is as an analytical tool, giving us actionable information to help us make business decisions.