**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)

A screenshot of a computer

AI-generated content may be incorrect.

* 1. Sales Performance by Product Category and Quarter

c. Visualization Format

A graph of a chart

AI-generated content may be incorrect.

* 1. Line Chart- Sales Trend by Product Category Over Time

A graph of a bar chart

AI-generated content may be incorrect.

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)

A screenshot of a computer

AI-generated content may be incorrect.

1.4 Product Performance Analysis with Price and Discount Impact

c. Visualization format

A graph showing different colored circles

AI-generated content may be incorrect.

1.5 Scatter Plot- Price vs. Revenue Relationship

A graph of blue rectangular objects

AI-generated content may be incorrect.

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)

A screenshot of a computer

AI-generated content may be incorrect.2.1 Customer Segment Profitability Analysis

c. Visualization Format

**A graph showing a bar chart

AI-generated content may be incorrect.**

2.2 Grouped bar chart- Key Performance Metrics by Customer Segment

**A pie chart with numbers and a red circle

AI-generated content may be incorrect.**

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)

A screenshot of a computer

AI-generated content may be incorrect.

3.1 Regional Sales Performance with Shipping Analysis

c. Visualization Format

**A yellow rectangular object with numbers

AI-generated content may be incorrect.**

3.2 Choropleth Map for Regional Sales Performance

**A graph with numbers and a yellow dot

AI-generated content may be incorrect.**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)

A screenshot of a computer

AI-generated content may be incorrect.

3.4 City-Level Sales with Category Analysis

c. Visualization Format

A graph of blue bars with black text

AI-generated content may be incorrect.

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)

A screenshot of a computer

AI-generated content may be incorrect.

4.1 Complex Multi-Dimensional Sales Analysis

C. Visualization Format

A chart of sales by customer segment and shipping method

AI-generated content may be incorrect.

4.2 Heatmap of Sales by Customer Segment and Shipping Method

A graph of a bar chart

AI-generated content may be incorrect.

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)

A screenshot of a computer

AI-generated content may be incorrect.

4.4 Customer Lifetime Value and Purchase Pattern Analysis

c. Visualization Format

A diagram of a customer segment comparison

AI-generated content may be incorrect.

4.5 Radar Chart for Customer Segment Comparison

A graph with different colored bars

AI-generated content may be incorrect.

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.