

Name: Aafrin Zahid Memon

ID: 300388614

Course: Data Visualization

Assignment 2 : PowerBI

Phase 1 : Problem Definition & Data Preparation

The following actions were made to guarantee clean and organized data in Power BI during the initial data preparation phase:

1. Superstore Sales Data

- Region Table Deleted: Because of missing data, the table was eliminated.
- User Table: To standardize column names, the first row served as a header.
- Order Table: To guarantee correct analysis, null values in the "Product Base Margin" column were eliminated.
- The formatting of numerical values (such as currency adjustments) was accurate.

2. Geographic & Census Data

- Power BI does not support the dataset's native Microsoft Excel 97-2003 Worksheet (.xls) format hence for compatibility, the file was transformed into a contemporary Excel workbook (.xlsx).
- Using the Split Column function and the First Row as Headers option, the dataset was divided into Region and State tables.

o What are the similarities and differences in how data cleaning is handled in both tools?

Similarities: Both allow removal of null values, data formatting, column renaming, and type conversion. They support splitting and restructuring tables.

Differences:

- Power BI uses Power Query Editor (built-in) for transformations, while Tableau relies on Tableau Prep.
- Power BI requires Excel (.xls) conversion, whereas Tableau supports legacy formats.
- Power BI offers Split Column by Delimiter, while Tableau allows data blending from multiple sources.

o How does each application approach missing values, outliers, or errors in the dataset?

Power BI: Employs "Replace Values" or eliminates nulls, uses Data Profiling to find outliers, and automatically fixes Power Query data type issues.

Tableau: Requires manual type corrections, using filters to exclude nulls, and uses box plots and clustering to identify outliers.

o What are the pros and cons of Power BI's data cleaning features compared to Tableau's?

Advantages of Power BI:

Power Query Editor integrated within the system (no additional tool required).
easy-to-follow, step-by-step changes.

automated data profiling for outliers and missing variables.

Strong M Language ETL (Extract, Transform, Load) capabilities.

Drawbacks of Power BI

Some formats require file conversion.

less adaptable than Tableau in terms of data merging.

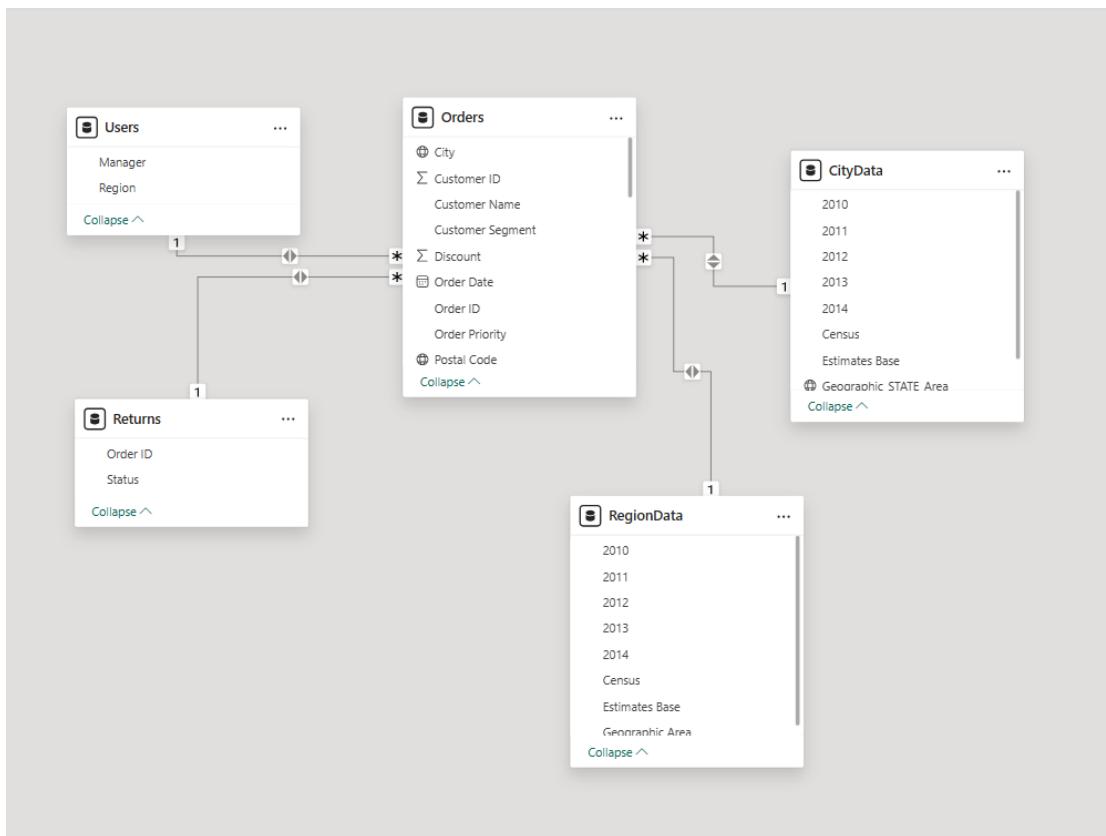
Limited detection of outliers (deeper analysis requires DAX or Python).

Phase 2:

Data Structure:

Relationships:

- Returns (Order ID) → Orders (Order ID): One-to-One (1:1) relationship, guaranteeing that there is only one return for every order ID in the Orders database.
- Region-by-Region Orders → Users: Multiple orders may be a part of a single user region due to a many-to-one (*:1) relationship.
- RegionData (Geographic Area) -> Orders (Region): Orders and geographical regions are connected by a many-to-one (*:1) connection.
- Geographic_STATE_Area → CityData → Orders (State or Province): Orders to a city-based dataset are mapped using a many-to-one (*:1) connection.



These connections facilitate the development of precise data models, which allow for cross-table analysis in dashboards and reports.

DateTable:

In summary, the **Date Table** was added to enable **time-based analysis, improve filtering, enhance performance, and ensure consistency** in handling dates across our Year on year analysis.

Relationships of DateTable

Orders (Order Date) → DateTable (Date)

- *One-to-Many (1: Relationship)**
- The **DataTable** serves as the **primary date reference** for the **Orders** table.
- Each order has a single **Order Date**, while the **DataTable** contains a continuous range of dates.

○ How do the data modeling features in Power BI compare to Tableau's ?

Power BI's data modeling works like a database, using a drag-and-drop "Manage Relationships" tool to create one-to-one (1:1) and many-to-one (*:1) relationships, making it easier for structured reports. In my Power BI project, this helped me link Orders, Users, and Geographic Data efficiently. Tableau, however, offers Joins, Relationships, and Data Blending, which allows combining data sources without full integration. While joins merge data at the row level, relationships connect tables dynamically. Data blending was useful when working with separate Excel files, but it required more manual adjustments

○ What are the pros and cons of these data modeling features in Power BI compared to Tableau's?

In Power BI, working with multiple datasets like Orders, Users, and Geographic Data. Power BI's relationship manager helps to easily link tables using one-to-one (1:1) and many-to-one (*:1) relationships, making structured data analysis smooth. It automatically detects relationships, which is useful as a beginner, but I had to manually adjust joins for better accuracy. Tableau, on the other hand, allows dynamic data blending without merging tables, which would have been helpful when working with separate Excel files. However, blending slows down with large datasets. Power BI is great for structured business reports, while Tableau is better for quick ad-hoc analysis across different sources.

Field Creation & Categorization:

Purpose of Using Calculated Fields (Measures): We have used the below calculated measures to help us throughout our analysis. And for efficiency all the calculated field are stored in a folder name “Sales Measure”

- **Total Orders** → Counts the number of orders placed.
- **Running Total Sales** → Computes cumulative sales over time.
- **Average Order Value** → Helps understand the average revenue per order.
- **Profit Margin** → Provides profitability insights per order or category.
- **YoY Sales Growth** → Year-over-Year sales growth calculation.
- **YoY Profit Growth** → Year-over-Year profit increase.
- **Running Total Sales** → Tracks progressive sales trends over time.
- **Conditional Formatting & Data Visualization**
- **Heading Gradient Color** → Used for formatting visual titles dynamically.
- **State Profit Gradient / Shipping Profit Gradient** → Conditional color formatting for profit insights.
- **High Profit Orders** → Identifies orders with higher-than-average profits.
- **Profit Category** → Groups orders into profit ranges.
- **CustomerRetentionRate** → Helps analyze returning customers over a period.
- **Discount Percentage** → Monitors discount impact on sales.

✓ Sales Measures

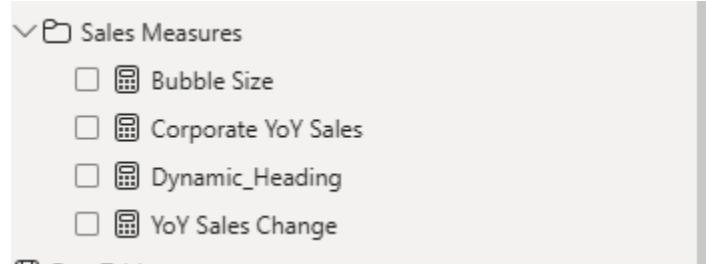
- Average Order Value
- CustomerRetentionRate
- Discount Percentage
- Heading Gradient Color
- High Profit Orders
- KPI Metrics
- Profit Category
- Profit Margin
- Running Total Sales
- Shipping Profit Gradient
- State Profit Gradient
- Total Orders
- Total Profit by State
- YoY Profit Growth
- YoY Sales Growth

Why do we use these?

Measures perform calculations on-demand, instead of storing redundant data in the model.

DAX calculations are faster compared to using calculated columns.

A few more fields were created in the CityData Table for dynamic heading, their purpose was only for formatting the dashboard



- **Bubble Size** → Determines the size of data points in a bubble chart based on sales, profit, or orders.
- **Corporate YoY Sales** → Calculates year-over-year sales for corporate-level data to analyze growth trends.
- **Dynamic_Heading** → Generates dynamic report titles based on user-selected filters like year or region.
- **YoY Sales Change** → Computes the percentage change in sales compared to the previous year for trend analysis.

Hierarchy

Geographic Hierarchy: Users can drill from **Region** → **State** → **City** to see more detailed insights, used for region wise orders, profit and sales analysis.



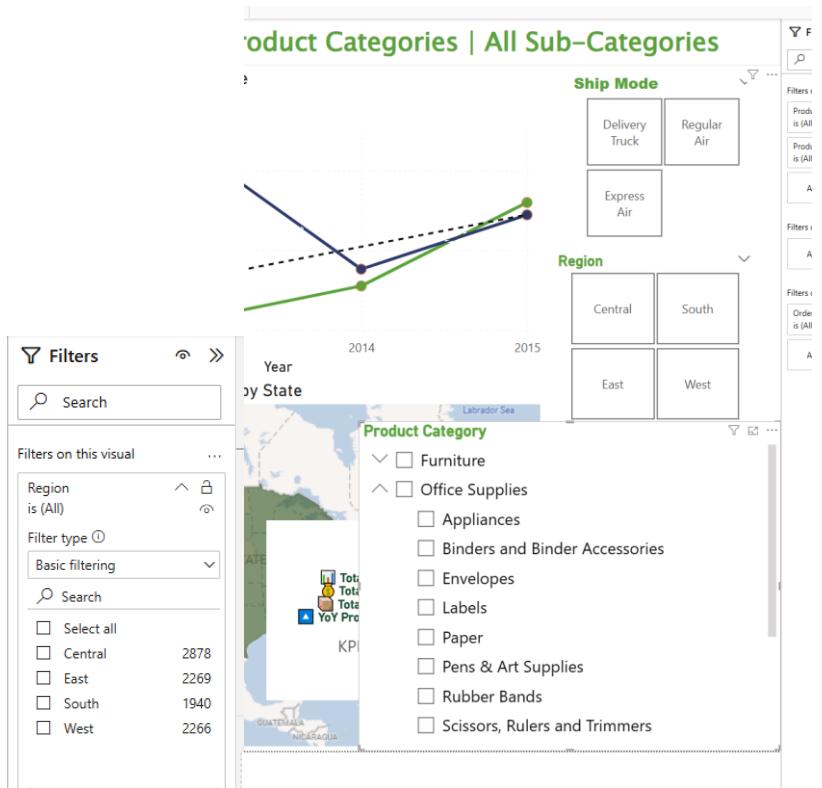
Power BI	Tableau Approach
In Power BI, field creation relies on DAX (Data Analysis Expressions) for calculations, such as Average Order Value, Profit Margin, and YoY Sales Growth, as seen in the Sales Measures section. Hierarchies like Region → State or Province → City in the	Tableau offers LOD (Level of Detail) expressions, Sets, and Quick Table Calculations, making it easier to create dynamic aggregations and groupings without writing complex formulas. Unlike Power BI, Sets in Tableau allow more flexible filtering and segmentation.

Geographic Hierarchy allow structured drill-down navigation.	
<ul style="list-style-type: none"> • DAX provides high performance for large datasets. • More control over aggregations using explicit measures. • Built-in relationships eliminate the need for data blending. • Less dynamic 	<ul style="list-style-type: none"> • there's no direct equivalent to LOD expressions. • Tableau has Quick Table Calculations, requiring users to manually create running totals or percentages.

Power BI's structured approach is best for relational data models, while Tableau provides faster, flexible insights with LOD expressions, Sets, and Table Calculations.

Filters Used:

- **Product Category** allows analysis and comparison of sales trends across different product categories like Furniture, Office Supplies, and Technology.
- **Region** filters sales and profit data by geographical region, helping track regional performance differences.
- **Sum of Sales** ensures proper aggregation of sales data within visuals, enabling comparison of total revenue across different filters.
- **Date - Year (is not 2010, 2011, ...)** excludes older data to focus on recent sales trends and improve year-over-year (YoY) growth analysis.
- **YoY Sales Growth / YoY Profit Growth** filters track year-over-year changes in sales and profit, ensuring only relevant trends are displayed.
- **Bubble Size (All)** affects bubble chart visuals by determining the relative size of each data point, ensuring proportional representation of sales and profit.
- **Sum of Profit & Sum of Shipping Cost** helps in analyzing profitability and tracking profit margins concerning shipping expenses.
- **State or Province** is used for geographical analysis, allowing users to drill down into state-level performance through heatmap visuals.
- **State Profit Gradient** applies to gradient map visuals, highlighting states with higher or lower profit to facilitate quick profitability comparisons.
- **Total Profit by State** displays total profit figures per state, working alongside the profit gradient filter for clearer regional performance analysis.
- **Order ID (Report-Level Filter)** ensures data consistency across all pages, allowing uniform selection, accurate analysis, and drill-down into individual orders for detailed insights.



- o What are the similarities and differences between how filters work in Power BI and Tableau?

In my Power BI project, I used visual-level filters to refine individual charts, such as filtering Order Date - Quarter, Product Sub-Category, Sum of Profit, and Sum of Sales. This helped focus each visual on relevant data without affecting the entire report. Power BI allows multi-level filtering, making it easy to isolate data while keeping the rest of the report dynamic. Unlike Tableau, which applies filters at the worksheet level, Power BI's visual filters provide more control over individual charts without impacting other report elements.

- o How do filters in Power BI differ from Tableau's filters?

To manage filtering across an entire report page, I applied page-level filters, which allow filtering all visuals on a single page without affecting others. Additionally, I used Advanced Filtering at the report level, filtering Order ID for values ≤ 2012 , ensuring consistency across multiple pages. Power BI's approach is similar to

Tableau's Data Source Filters, but Tableau also has Context Filters, which improve performance by pre-filtering data before calculations.

- o What are the pros and cons of the filters in Power BI compared to Tableau's?

Power BI's Top N filtering, as seen in Order Date - Month filtered by Running Total Sales, helps focus on key metrics. However, Power BI lacks Tableau's quick Context Filters and requires DAX for advanced hierarchy filtering. Power BI's structured approach makes filtering easy for reports, while Tableau provides more dynamic filtering options for deeper exploration.

a few observations on filters:

- Power BI's filters are more **intuitive and user-friendly**, while Tableau offers **more control** over filter application.
- If working with **dynamic, interactive dashboards**, Power BI is easier to use.
- If working with **large datasets and need query efficiency**, Tableau might be more performant.

Phase 3 : Understanding Our Reports & Dashboard

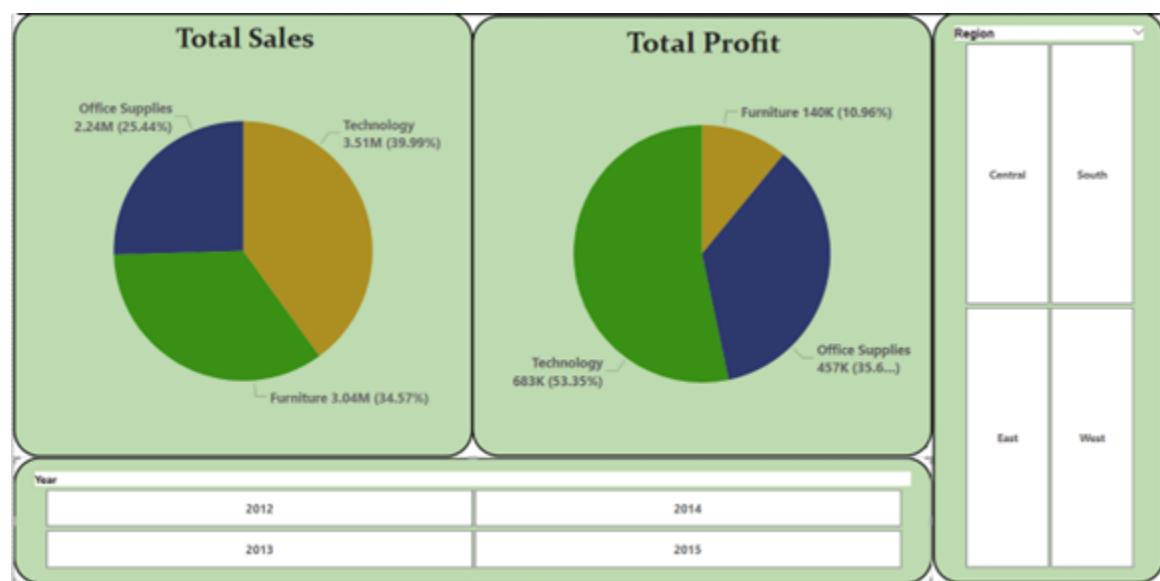
To understand our reports and dashboard, we begin by defining our key business questions.

1. How do sales and profit vary across different geographic regions and product categories?

To answer this, we analyze the following sub-questions:

Which product categories contribute the most to total sales and profit?

→ Visual: Pie Charts for Total Sales & Total Profit

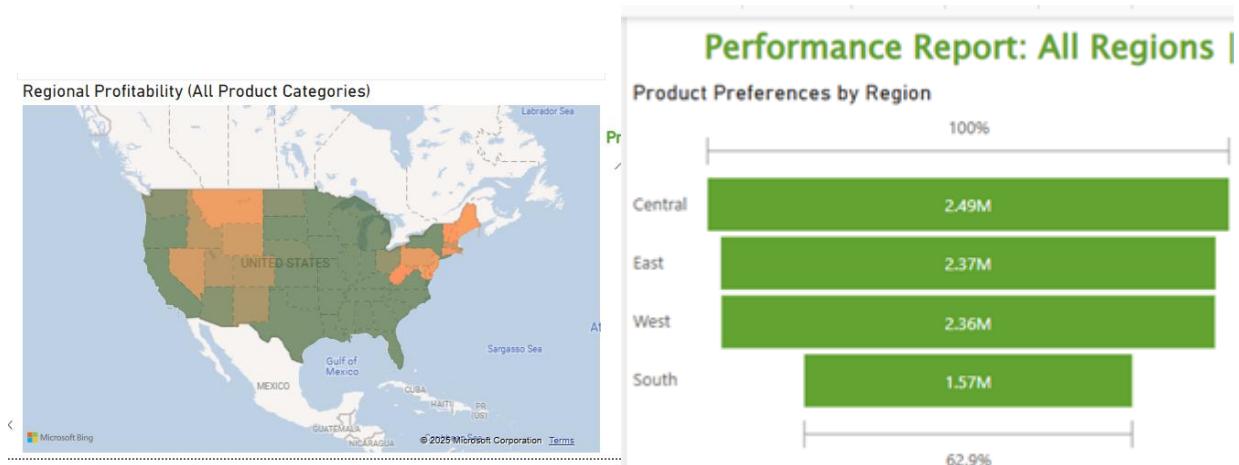


Why was this visual used?

- Pie charts effectively display proportions of a whole.
- They allow quick comparisons of product categories' contributions to total sales and profit.
- They help identify dominant categories driving revenue and profitability.

Which regions generate the highest orders and revenue?

→ Visuals: Regional Profitability Map with Gradient & Funnel Chart by Region



Regional Profitability (Filled Map with Gradient)

- A filled map (choropleth) with a profitability gradient was chosen to intuitively visualize geographic performance.
- Highlights profitable vs. underperforming states, identifying strong and weak profitability zones.
- Color gradients help distinguish between high-profit (green) and low-profit (orange) states.

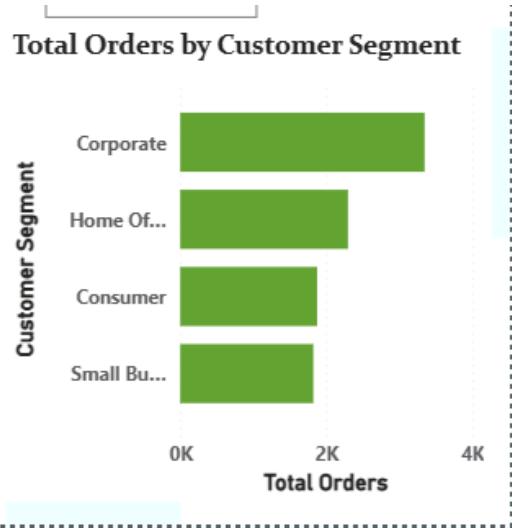
Product Preferences by Region (Funnel Chart)

- A Funnel chart is ideal for comparing categorical data, making regional sales differences easy to analyze.
- Represents total sales per region, highlighting the top contributors to revenue.
- The horizontal format improves readability and facilitates quick comparisons.

How does Order distribution vary by customer segment?

→ Visual: Bar Chart - Orders by Customer Segment

- This visual enables comparison of different customer segments' contributions to total Orders.



Which regions and product categories are underperforming?

→ Visuals: YoY Growth Chart & Profit Gradient Map

Year-over-Year (YoY) Sales & Profit Growth (Line Chart)

- A line chart is best suited for visualizing trends over time.
- Displays whether sales and profit growth are improving or declining over the years.
- Compares YoY sales growth (green) vs. YoY profit growth (blue) to determine if profit aligns with sales trends.
- We have also used an average line here for swift analysis.



2. How do shipping costs impact profitability across different regions and order types?

- ◆ How does shipping cost correlate with profit?

Visual: Bubble Chart - Shipping Cost vs. Profit

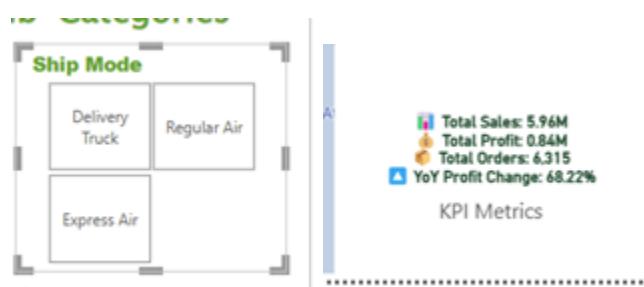
Why was this visual used?

- A bubble chart effectively correlates shipping costs with profitability.
- Bubble size represents the number of orders, while color intensity indicates profitability.



Which ship mode is the most cost-effective & profitable?

→ Visuals: Ship Mode Slicer + KPI Cards



Ship Mode Slicer (Delivery Truck, Regular Air, Express Air)

- Allows filtering of performance metrics based on selected shipping modes.
- Helps compare the impact of different shipping methods on total sales, profit, and order volume.

Findings:

- Selecting different ship modes dynamically updates KPI metrics.
- Some shipping modes may be cheaper but less profitable, while others have higher costs but drive better margins.

KPI Cards (Total Sales, Total Profit, Total Orders, YoY Profit Growth)

- Provides a quick summary of key metrics related to sales, profit, and order volume.
- Helps users instantly compare performance across different shipping modes.

Do certain regions have higher shipping costs that reduce profit margins?

→ Visual: Regional Profitability Map with Gradient - Shipping Cost Impact

What pricing or logistics strategies can optimize shipping cost impact?

→ Visual: Trend Charts & Filters on Shipping Mode & Region

How Combining These Visuals Strengthens Insights

1. Geographic and Product Performance Connection

- Funnel Chart (Product Preferences by Region) → Shows which regions drive the most sales.
- Choropleth Map (Regional Profitability) → Highlights which states are actually profitable, not just generating high sales.
-  **Combined Insight:** A region with high sales but low profitability might have high discounts or high shipping costs.

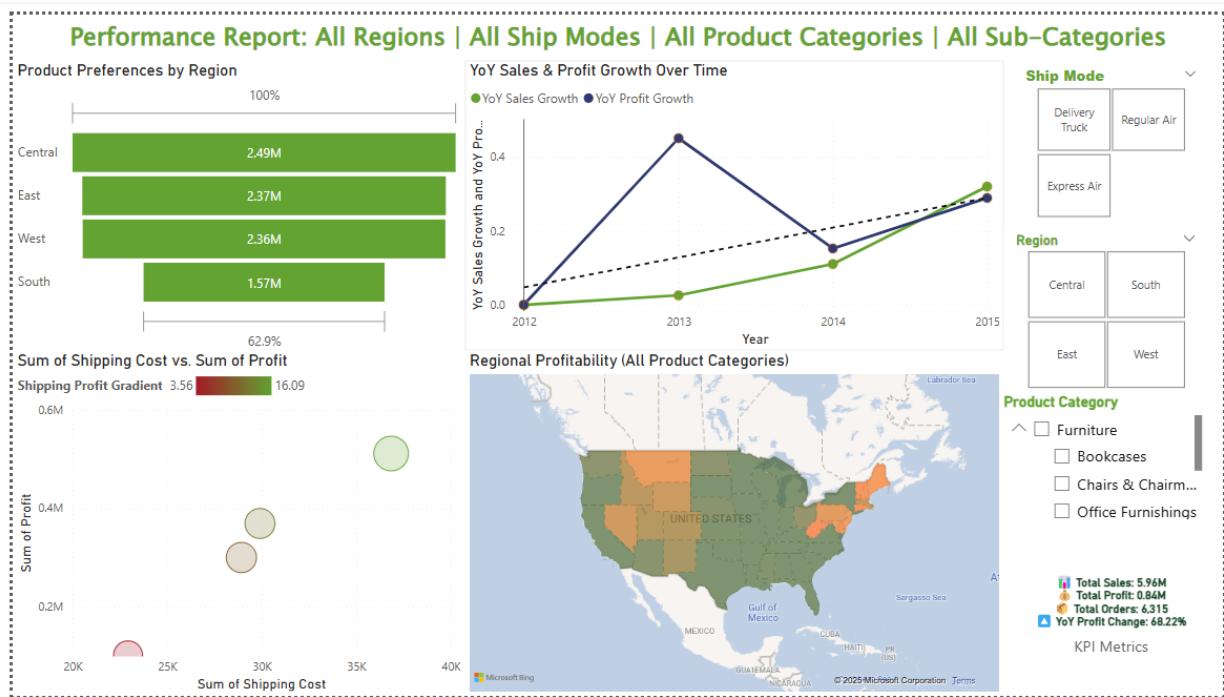
2. Time-Based Performance Trends

- Line Chart (YoY Sales & Profit Growth) → Shows the overall trend of sales and profit over time.
- Bubble Chart (Shipping Cost vs. Profit) → Explains why profit might fluctuate due to increasing costs, inefficient logistics, or seasonal trends.

- 💡 **Combined Insight:** If profit dips despite increasing sales, logistics costs or discounting strategies might be the issue.

3. Logistics Cost and Revenue Optimization

- Bubble Chart (Shipping Cost vs. Profit)** → Shows if high shipping costs reduce profit margins.
- Regional Profitability Map** → Pinpoints where high logistics costs hurt profit the most.
- 💡 **Combined Insight:** Helps businesses optimize shipping modes or reconsider warehouse placements for cost efficiency.



Why This Approach Works?

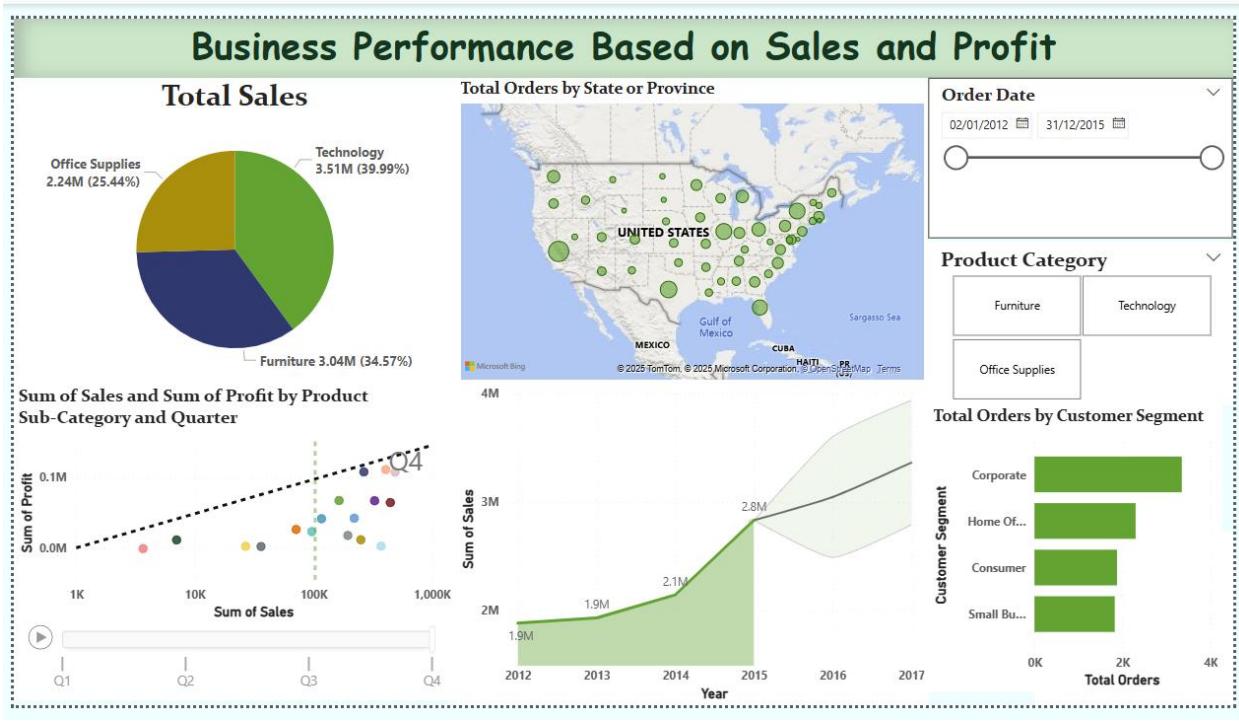
✓ **Big Picture + Granular Details** → Sales alone isn't enough; profitability and cost efficiency matter. ✓ **Detecting Patterns & Trends** → Line charts show long-term trends, while maps and bar charts show current regional performance. ✓ **Data-Driven Decision Making** → Helps companies adjust pricing, inventory, and logistics strategies based on interconnected insights.

Comprehensive Sales and Profit Analysis Across Categories and Regions

This dashboard provides a **holistic view of business performance** by analyzing total sales, order distribution, and profit trends across product categories, customer segments, and geographic regions.

- The **pie chart** highlights the contribution of different product categories to total sales, helping identify dominant and underperforming categories.
- The **map visualization** shows total orders by state, identifying high-demand regions.
- The **scatter plot and trend line** analyze sales-to-profit correlation, helping detect inefficiencies.
- The **line chart** presents year-over-year sales growth, revealing long-term business trends.
- A **forecast has been added to the line chart** to predict future sales trends based on historical data. This helps businesses anticipate demand fluctuations, adjust inventory, and make strategic pricing decisions.
- The **bar chart** compares total orders by customer segment, providing insights into **customer preferences**.

By combining these visuals, businesses can **track key trends, optimize inventory and pricing strategies, and enhance sales growth while improving profitability** across different regions and customer segments.

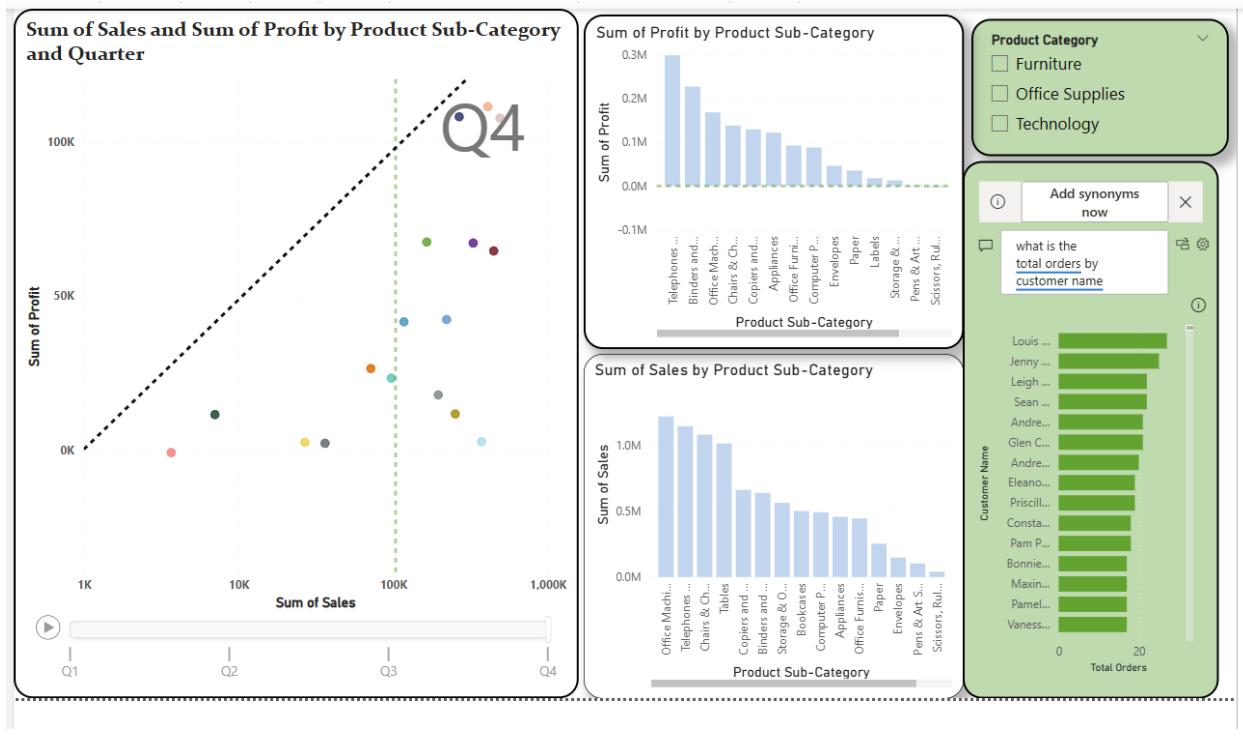


Sales and Profit Correlation Across Product Sub-Categories

This visualization was created to analyze the relationship between **sales and profit across product sub-categories** over different quarters. The scatter plot helps identify **high-sales, low-profit categories** that may require pricing adjustments, promotional strategies, or cost optimization.

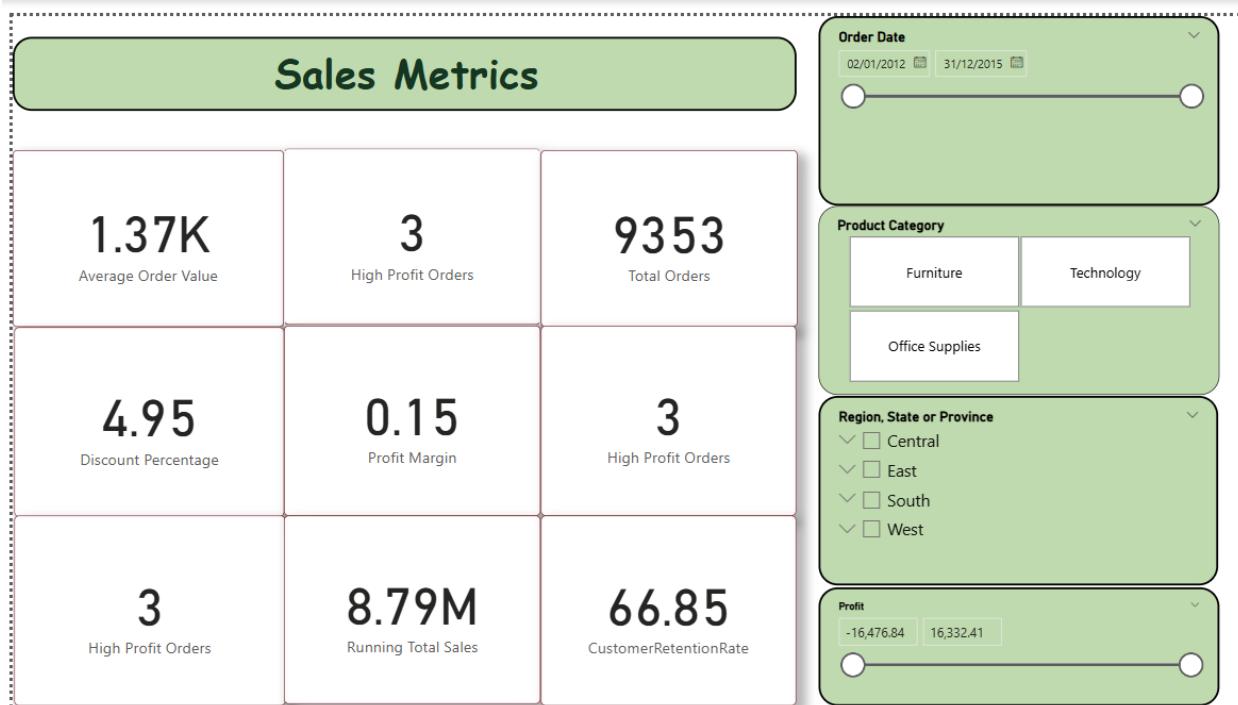
- The **dotted trend line** indicates the expected profit-to-sales ratio.
- The **bubble colors and sizes** help detect anomalies—such as products generating high sales but low profitability.
- The bar charts on the right further break down **sales and profit per sub-category**, allowing for a **detailed comparison**.

By combining these visuals, businesses can **pinpoint which product sub-categories are underperforming**, enabling data-driven **inventory, marketing, and pricing strategies** for increased profitability.



All Sale Metrics

Additionally, a visual was created combining our metrics on tile cards. The Sales Metrics dashboard provides a quick summary of key business performance indicators related to sales, profitability, and customer retention. It helps users track sales trends, profit margins, and order values immediately.



Final Thought💡

By combining these visuals, businesses move beyond surface-level insights and uncover the real factors driving success or inefficiency—leading to more profitable, strategic decisions.

How filters are implemented in Power BI and Tableau.

In Power BI, I used Slicers and Drill Through to enhance interactivity, similar to Parameters in Tableau. The slicers allow users to filter by year, quarter, region, and product category, making analysis dynamic. Power BI's Drill Through enables deeper insights into specific categories, whereas Tableau's Parameters provide more flexibility by allowing dynamic calculations.

Power BI's filters differ from Tableau's in hierarchical navigation and global filtering. While Power BI supports multi-level filters (visual, page, and report-wide), Tableau uses Context Filters for performance optimization. Tableau's filters can be applied across multiple dashboards, while Power BI requires relationships between tables to achieve the same.

Power BI Pros:

- Easier to use with drag-and-drop filtering.
- Drill Through improves category-specific analysis.
- Slicers provide a user-friendly filtering experience.

Power BI Cons:

- Lacks dynamic parameter-driven calculations like Tableau.
- Filters are more rigid, requiring DAX for advanced scenarios.

Overall, Power BI's filters are better for structured reporting, while Tableau's filters offer more flexibility for exploratory analysis

Power BI vs. Tableau Dashboard Structure

Power BI follows a report-page structure, where reports consist of multiple pages, and each page acts as a dashboard. In contrast, Tableau uses a workbook containing multiple sheets, dashboards, and stories. Tableau's dashboards combine multiple sheets, while Power BI dashboards are self-contained pages with interactive elements like filters, slicers, and drill-throughs.

Power BI Pros:

- Better report organization through structured pages.
- Integrated filters across all visuals without separate setup.
- Drill-through functionality for detailed insights.

Power BI Cons:

- Less flexible than Tableau's free-form dashboard layout.
- Lacks a direct "Story" feature for sequential storytelling.

Phase 5: A Readme page

README Page

How to Use This Report

This Power BI report provides interactive visualizations for exploring sales and profit trends across different regions, time periods, and product categories. Users can interact with filters, slicers, and drill-through options to gain deeper insights.

Overview of Report Pages

Performance Report Dashboard

This dashboard provides insights into regional sales performance, profitability trends, shipping cost impact, and product preferences. Users can interact with filters to analyze different metrics and uncover key business insights.

How to Use This Dashboard

Product Preferences by Region (Bar Chart)

- Displays total sales volume per region (Central, East, West, South).
- Helps identify the best-performing regions.

YoY Sales & Profit Growth (Line Chart with Trendline)

- Compares Year-over-Year (YoY) Sales Growth (green) and Profit Growth (blue).
- A trendline is added to forecast growth trends.

Sum of Shipping Cost vs. Profit (Bubble Chart)

- Shows the relationship between shipping costs and profitability.
- Bubble size represents order volume, and color gradient (red-green) indicates profit impact.

Regional Profitability (Choropleth Map)

- Highlights the most and least profitable states with color-coded gradients.
- Green states are high-profit, while orange states indicate lower profitability.

Filters & Interactive Features

- **Ship Mode Filter:** Analyze the impact of different shipping methods (Delivery Truck, Regular Air, Express Air).
- **Region & Product Category Filters:** Compare sales and profit metrics across different regions and product types.
- **Hover Over Visuals:** Gain quick insights into specific data points.

Key Insights to Look For

- Identify top-performing regions and product categories.
- Compare how profitability trends align with sales growth.
- Determine which shipping costs negatively impact profit margins.
- Use filters to explore how different variables impact business performance.

This dashboard enables data-driven decision-making for regional expansion, cost optimization, and sales strategies. ↗

Introduction & Analysis Questions

This Business Performance Dashboard provides insights into sales and profit trends across product categories, customer segments, and geographic locations. It highlights total sales distribution, order trends by state, sales and profit by quarter, and customer segment order volume. Users can filter by quarter and product category to explore trends and identify key business opportunities.

Key Business Questions

1 How do sales and profitability vary across geographic regions and product categories?

• Why this question?
Helps identify which regions and product categories contribute most to total sales and profit.
Supports strategic decision-making on product allocation and marketing.

Sub-questions (Supported by Specific Visuals)

- Which product categories contribute the most to total sales and profit?
→ (Pie Charts - Total Sales & Total Profit Breakdown in Dashboard 1 and 2)
- Which regions generate the highest orders and revenue?
→ (Regional Profitability Map with Gradient & Bar Chart by Region in Dashboard 1 and 2)
- How does sales and profit distribution vary by customer segment?
→ (Stacked Bar Chart - Sales & Profit by Customer Segment in Dashboard 2)
- Which regions and product categories are underperforming?
→ (YoY Growth Chart & Profit Gradient Map in Dashboard 1 & 2)
- What strategies can improve underperforming regions and product categories?

2 How do shipping costs impact profitability across different regions and order types?

• Why this question?
Helps in cost optimization by evaluating the relationship between shipping cost and profit margins across different shipping modes & regions.
Supports logistics and fulfillment strategy improvements.

Sub-questions (Supported by Specific Visuals)

- How does shipping cost correlate with profit?
→ (Bubble Chart - Shipping Cost vs. Profit in Dashboard 1)
- Which ship mode is the most cost-effective & profitable?
→ (Ship Mode Slicer + KPI Cards in Dashboard 1 & 2)
- Do certain regions have higher shipping costs that reduce profit margins?
→ (Regional Profitability Map with Gradient - Shipping Cost Impact in Dashboard 1)
- What pricing or logistics strategies can optimize shipping cost impact?
→ (Trend Charts & Filters on Shipping Mode & Region in Dashboard 1 & 2)

Conclusion

The Business Performance Dashboard provides a comprehensive analysis of sales, profitability, and operational efficiency across different regions, product categories, and shipping modes. The dashboard offers key insights into geographic performance, profit distribution, and cost-impacting factors.

Key Findings & Observations

Sales & Profit Variations Across Geographic Regions & Product Categories

- Why is this important?
 - Identifies top-performing regions & product categories to enhance profitability and strategize product allocation.
- Technology contributes the highest to total sales, followed by Furniture and Office Supplies.
- The Central & East regions generate the highest total orders, while the South region has the lowest sales volume.
- Certain product categories have high sales but lower profit margins, indicating room for pricing optimization.
- YoY Profit Growth trends indicate an increase in revenue, but some regions show stagnation, requiring targeted improvements.

Observations & Recommendations

- High-performing regions & categories present opportunities for expansion.
- Underperforming product categories need pricing or promotional strategies to boost profitability.

Impact of Shipping Costs on Profitability Across Regions & Order Types

- Why is this important?
 - Helps businesses optimize shipping cost strategies and increase overall profit margins.
- Shipping costs significantly impact profitability, especially in regions with high logistics expenses.
- Regular Air & Express Air shipping modes show higher shipping costs, potentially affecting profit margins.
- South & West regions have higher logistics costs, reducing profitability compared to Central & East.
- The relationship between shipping costs & profit suggests a need for cost optimization.

Observations & Recommendations

- Optimizing shipping modes based on region-specific profit trends can enhance overall margins.
- Regions with higher logistics costs should explore partnerships with cost-efficient carriers.
- Customer behavior in different shipping modes should be analyzed to reduce overall fulfillment costs.

Final Recommendations & Next Steps

What should businesses do with these insights?

- Expand high-performing regions & product categories to sustain revenue growth.

- Optimize logistics strategies to reduce high shipping costs in specific regions.

- Targeted marketing & pricing strategies for underperforming categories can enhance revenue & profitability.

- Further analyze customer behavior in high-cost shipping regions to identify alternative fulfillment strategies.

PowerBI shared link: https://app.powerbi.com/links/ZBkjJEHCsK?ctid=3af48838-cd53-4507-9e7f-fc6dac355e33&pbi_source=linkShare&bookmarkGuid=e7c7b3ee-bbed-448f-af8b-2252930143fc