Building Data Visualization with Power BI

Aysegul Dahi

Analytical Questions for the Project

1. Sales & Profitability Analysis

Question: Does a higher discount necessarily lead to higher sales? Identify any products or categories where this trend doesn't hold.

Why this question? Helps evaluate pricing strategies and their impact on profitability. Determines whether higher discounts effectively drive sales or if they lead to losses.

2. Regional & Customer Analysis

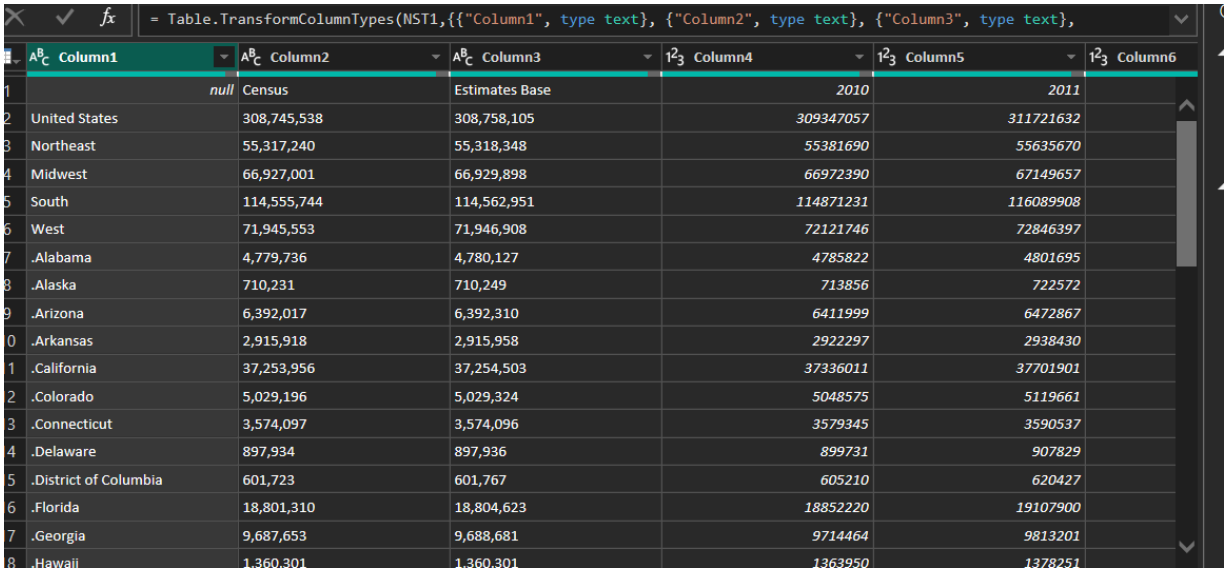
Question: Which regions demonstrate a significant impact of shipping costs on profit margins? What strategies could optimize these costs?

Why this question? Identifies regions where high shipping costs reduce profit margins. Helps in optimizing supply chain strategies for cost efficiency.

Data Loading and Transformation

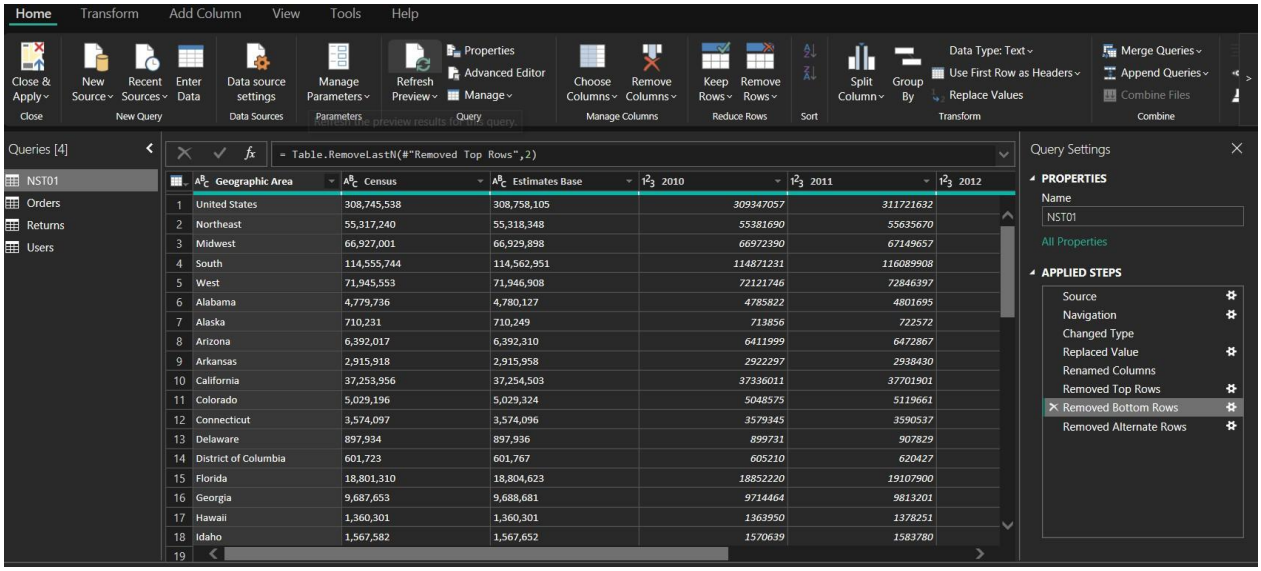
1. Removing Dots from Column Names

Dots were removed from column names to ensure consistency and better readability. This was done using the 'Replace Values' function in Power Query.



2. Renaming Columns

Columns were renamed to meaningful names to enhance data interpretation and usability.



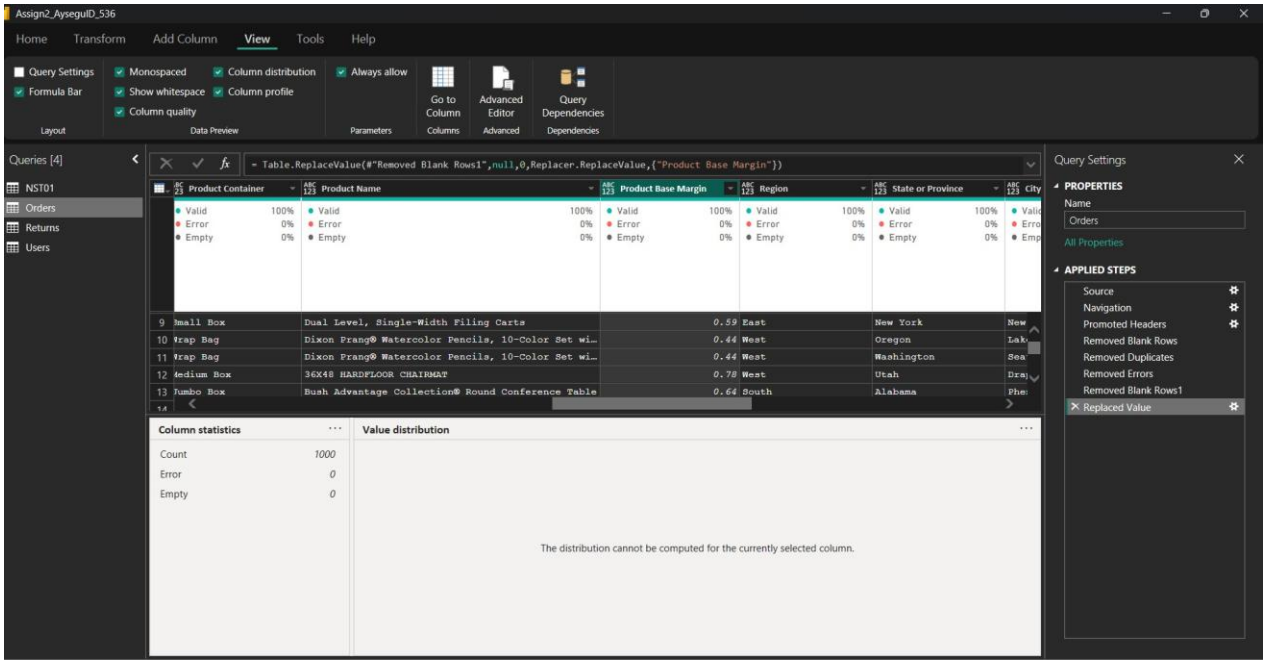
3. Removing Blank Rows

Blank rows were removed to ensure a clean dataset. This was done using the 'Remove Blank Rows' option in Power Query.

4. Removing Duplicates: Duplicate records were removed from the dataset to avoid redundancy and incorrect analysis. It was achieved using the 'Remove Duplicates' function in Power Query.

5. Replacing Null Values in 'Product Base Margin' Column

There were 9 null values in the 'Product Base Margin' column. These were replaced with 0 to maintain consistency in numerical analysis. This was done using the 'Replace Values' option in Power Query.



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

Note: In my first assignment, I performed data cleaning using Jupyter Notebook, where I handled null values and outliers using Python libraries like Pandas and NumPy. This approach was easier for me compared to Tableau. I still think cleaning data in Python is faster than both Tableau and PowerBI. For removing outliers, I looked at column distribution and quality in PowerBI, but I had difficulties detecting them with built-in functions.

Similarities and Differences Between Power BI and Tableau

Similarities: Both tools provide GUI-based data cleaning with transformation functions. They offer data profiling features such as column statistics and distributions. In addition to these, they allow merging and appending datasets from multiple sources.

Differences: Power BI uses Power Query with M language for data transformation. On the other hand tableau uses Tableau Prep for data cleaning, offering a more visual workflow. Power BI

allows step-by-step transformations with applied steps, while Tableau Prep allows drag-and-drop transformations with flow diagrams.

Handling Missing Values, Outliers, and Errors: PowerBI, offers 'Replace Values' and 'Remove Rows' options in Power Query and allows filtering out outliers using visual filters.

Handling Outliers in Power BI

To ensure data accuracy and avoid misleading insights, we applied the **Interquartile Range (IQR) method** to detect and remove outliers in the **Profit** column of the Orders table. The IQR method is effective in identifying extreme values that fall significantly outside the normal range.

Date Structure

Relationships Created in This Project

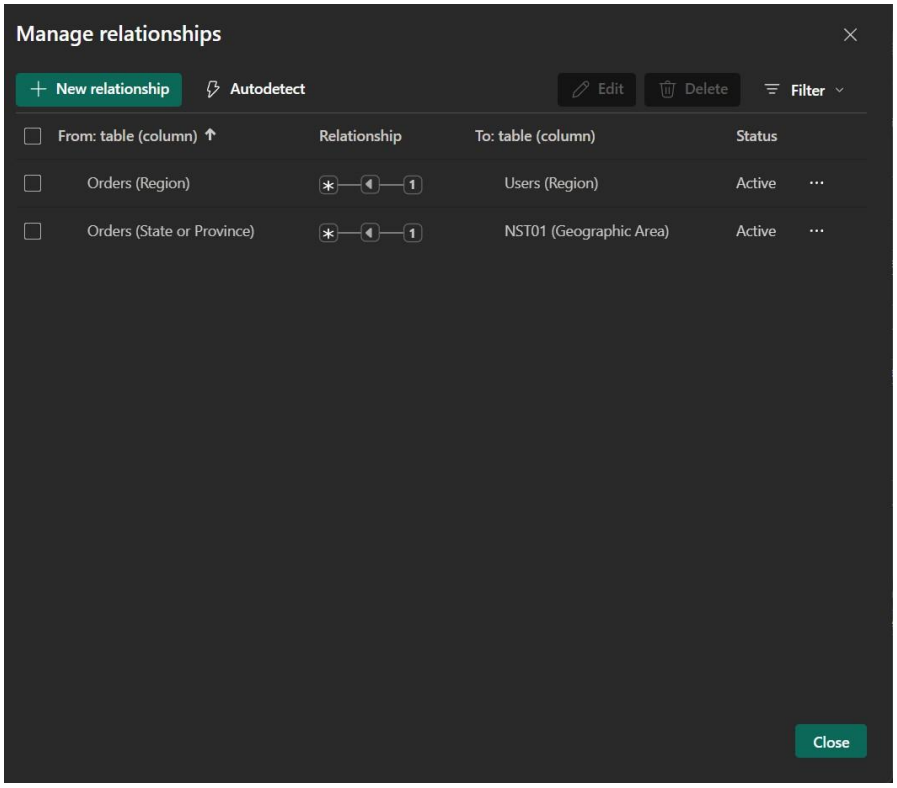
In this project, the following relationships were established to ensure proper data integration:

1. Orders (State or Province) to NST01 (Geographic Area) - Type: One-to-Many (1:*) -
Purpose: To link sales data from the 'Orders' table with census data from 'NST01' based on States.
2.
3. Orders (Region) to Users (Region) - Type: One-to-Many (1:*) -
Purpose: To associate sales orders with regional managers in the 'Users' table for better tracking of sales performance by region. These relationships ensure that geographic insights and regional responsibilities are properly reflected in Power BI dashboards and reports.

1. How Power BI Data Modeling Compares to Tableau Power BI and Tableau have different approaches to data modeling?

Power BI: Uses a structured data model with relationships similar to databases. Tables are linked using primary keys and foreign keys, enabling a well-defined data structure.

Tableau: Offers flexible data connections through relationships, joins, and data blending. Users can create joins between tables dynamically or use relationships that adapt based on the visualization.
In my first assignment, I used a Left Join in Tableau, which allowed me to merge data from multiple tables while keeping all records from the ‘Orders’ table. Power BI relationships, on the other hand, work dynamically across the model without requiring joins in every query.



2. Pros and Cons of Power BI vs Tableau Data Modeling

Power BI Data Modeling Pros: Uses a structured relationship model, similar to relational databases. Automatically detects relationships and enforces referential integrity. Supports advanced DAX calculations for deeper analysis.

Power BI Data Modeling Cons: Requires structured data before importing into Power BI. - Many-to-Many relationships can be complex and impact performance.

Tableau Data Modeling Pros: Offers flexibility with joins and data blending across multiple sources. - Relationships adjust dynamically based on visualization needs. Easier to integrate with real-time data sources.

Tableau Data Modeling Cons: Data blending does not create a persistent model, requiring repeated joins. Lacks built-in referential integrity, which can lead to inconsistent data joins.

In summary, Power BI is better for structured, relational data models, while Tableau offers more flexibility with dynamic joins and blending.

Field Creation & Categorization

To enhance my dataset and enable deeper analysis, I created **calculated fields (measures and columns)** in Power BI. Since Power BI does not support **Level of Detail (LOD) Expressions** like Tableau, we used **DAX measures** to achieve similar functionality.

Hierarchies, Groups, and Tableau Set Equivalent in Power BI

1. **Geographic Hierarchy**

Implementation Steps: Created a hierarchy in Power BI by **grouping Region, State, and City**. Allows users to **drill down** from **Regions** to **States** and **Cities**. Helps analyze **sales and profit trends across geographic levels**.

Outcome: Users can **explore geographic trends interactively** using maps and charts.

2. **Product Hierarchy**

Implementation Steps: Allows users to **drill down from broader product categories to individual products**. Helps analyze **sales and profit distribution at different product levels**.

3. **Grouping Low-Profit States**

Outcome: Allows filtering and segmentation of **underperforming regions**. Enables targeted analysis for **profitability improvements**.

4. **Replacing Tableau Sets in Power BI**

Since Power BI **does not have a direct equivalent to Tableau Sets**, we used:

DAX-Based Dynamic Grouping

Instead of manual grouping, I used a **calculated column e** to classify states dynamically:

```
State_Profitability =

VAR StateProfit = CALCULATE(SUM('Orders'[Profit]), ALLEXCEPT('Orders', 'Orders'[State or Province]))

VAR StateSales = CALCULATE(SUM('Orders'[Sales]), ALLEXCEPT('Orders', 'Orders'[State or Province]))

VAR ProfitMargin = IF(StateSales = 0, BLANK(), StateProfit / StateSales)

RETURN

   IF(ISBLANK(ProfitMargin), "No Sales Data",

      IF(ProfitMargin < 0, "High-Risk State",

         IF(ProfitMargin <= 0.1, "Moderate-Risk State",
```

"Low-Risk State’

Necessary Calculated Columns and Measures

1. Profit Margin Calculation (Measure)

To analyze profitability across products and regions, I used a measure for **Profit Margin**:

Profit_Margin = DIVIDE(SUM('Orders'[Profit]), SUM('Orders'[Sales]), 0)

This formula calculates the **profit percentage** relative to total sales. Helps in identifying which **products, regions, or categories** generate the highest profit margins.

2. Order Duration Calculation (Calculated Column): To track the time between order placement and shipment, I created **Order Duration** as a calculated column:

Order_Duration = DATEDIFF('Orders'[Order Date], 'Orders'[Ship Date], DAY)

3. Shipping Cost Impact Score (Measure)

To understand how much shipping costs affect revenue, I created the **Shipping Cost Impact Score** measure:

Shipping_Cost_Impact = DIVIDE(SUM('Orders'[Shipping Cost]), SUM('Orders'[Sales]), 0)

4. State-Level Profit Margin (Measure)

This measure **fixes calculations at the State level**, ignoring other filters.

State_Profit_Margin =

VAR TotalProfit = CALCULATE(SUM('Orders'[Profit]), ALLEXCEPT('Orders', 'Orders'[State or Province]))
VAR TotalSales = CALCULATE(SUM('Orders'[Sales]), ALLEXCEPT('Orders', 'Orders'[State or Province]))
RETURN DIVIDE(TotalProfit, TotalSales, 0)

5. Average Discount by Product Category (Measure)

This ensures the **discount average is calculated per category**, even if filters are applied.

Avg_Discount_Per_Category =
CALCULATE(AVERAGE('Orders'[Discount]), ALLEXCEPT('Orders', 'Orders'[Product Category]))

6. Total Sales per Region (Measure)

This measure **aggregates sales per region**, independent of other filters.

```
Total_Sales_Per_Region =
CALCULATE(SUM('Orders'[Sales]), ALLEXCEPT('Orders', 'Orders'[Region]))
```

7. Discount Percentile Classification (Calculated Column)

Since Power BI does **not allow percentile measures with filters**, I use **PERCENTILE.INC** and store it as a **column**.

```
Discount_Percentile =
VAR DiscountValue = 'Orders'[Discount]
VAR P20 = PERCENTILE.INC('Orders'[Discount], 0.20)
RETURN
IF(DiscountValue > P20, "High Discount", "Low Discount")
```

8. **State Profitability Classification (Calculated Column)**

```
State_Profitability =  VAR StateProfit = CALCULATE(SUM('Orders'[Profit]),
ALLEXCEPT('Orders', 'Orders'[State or Province]))

VAR StateSales = CALCULATE(SUM('Orders'[Sales]), ALLEXCEPT('Orders', 'Orders'[State or
Province]))

VAR ProfitMargin = IF(StateSales = 0, BLANK(), StateProfit / StateSales)

RETURN

   IF(ISBLANK(ProfitMargin), "No Sales Data",

      IF(ProfitMargin < 0, "High-Risk State",

         IF(ProfitMargin <= 0.1, "Moderate-Risk State",

          "Low-Risk State’
```

9. Cumulative Sales (Measure)

This measure calculates the cumulative percentage of total sales across product sub-categories, enabling Pareto analysis and identifying top-performing categories.

```
Cumulative_Sales = VAR TotalSales = CALCULATE(SUM('Orders'[Sales]),
ALLSELECTED('Orders'[Product Sub-Category]))

 VAR RunningTotal = CALCULATE( SUM('Orders'[Sales]), FILTER(
ALLSELECTED('Orders'[Product Sub-Category]), 'Orders'[Product Sub-Category] <=
MAX('Orders'[Product Sub-Category]) ) )

RETURN DIVIDE(RunningTotal, TotalSales, 0) * 100
```

10. Profitability by Shipping Mode (Measure)

Measures are dynamically calculated for each **shipping mode**.

```
Profit_Margin_By_Shipping =

VAR TotalProfit = SUM('Orders'[Profit])

VAR TotalSales = SUM('Orders'[Sales])

RETURN

IF(TotalSales = 0, BLANK(), TotalProfit / TotalSales)
```

11. State Profit Rank

This measure ranks each state by total profit in descending order, providing a national profitability ranking for state-level analysis.

```
State_Profit_Rank = RANKX(

   ALL('Orders'[State or Province]),

   CALCULATE(SUM('Orders'[Profit])),

   ,

   DESC,

   DENSE

)
```

12. State Margin Profit Column (Calculated Column)

This measure calculates the profit margin for each state by dividing the total profit by total sales, allowing for profitability analysis across different states.

State_Profit_Margin_Column = VAR TotalProfit = CALCULATE(SUM('Orders'[Profit]), ALLEXCEPT('Orders', 'Orders'[State or Province])) VAR TotalSales = CALCULATE(SUM('Orders'[Sales]), ALLEXCEPT('Orders', 'Orders'[State or Province])) RETURN IF(TotalSales = 0, BLANK(), TotalProfit / TotalSales)

13. Selected State Profit Rank (Measure)

This measure is designed for the drill-through page, dynamically retrieving the profit rank of the selected state to provide insights into its profitability relative to other states.

Selected_State_Profit_Rank = VAR SelectedState = SELECTEDVALUE('Orders'[State or Province]) RETURN CALCULATE([State_Profit_Rank], 'Orders'[State or Province] = SelectedState)

14. Selected Measure for Pareto Chart (Profit or Sales)

This measure enables dynamic switching between Sales and Profit based on user selection, allowing for comparative analysis within visualizations.

Selected_Measure = VAR SelectedValue = SELECTEDVALUE('Measure_Selector'[Measure]) RETURN SWITCH(SelectedValue, "Sales", SUM('Orders'[Sales]), "Profit", SUM('Orders'[Profit]))

15. Cumulative Profit (Measure)

This measure calculates the cumulative percentage of total profit across product sub-categories, enabling Pareto analysis to identify which sub-categories contribute most to overall profitability.

Cumulative_Profit = VAR TotalProfit = CALCULATE(SUM('Orders'[Profit]), ALLSELECTED('Orders'[Product Sub-Category])) VAR RunningTotal = CALCULATE(SUM('Orders'[Profit]), FILTER(ALLSELECTED('Orders'[Product Sub-Category]), 'Orders'[Product Sub-Category] <= MAX('Orders'[Product Sub-Category]))) RETURN DIVIDE(RunningTotal, TotalProfit, 0) * 100

16. Average Delivery Time

It calculates the average delivery time in days for each state or province, allowing for regional analysis of shipping efficiency and potential delays.

Avg_Delivery_Time = CALCULATE(AVERAGE('Orders'[Delivery_Days]), ALLEXCEPT('Orders', 'Orders'[State or Province])

17. Average Discount Per Category (Measure)

This measure calculates the average discount applied within each product category, enabling analysis of discounting strategies across different categories and their potential impact on sales and profitability.

Avg_Discount_Per_Category = VAR CategoryDiscount = CALCULATE(AVERAGE('Orders'[Discount]), ALLEXCEPT('Orders', 'Orders'[Product Category])) RETURN IF(ISBLANK(CategoryDiscount), 0, CategoryDiscount)

18. Discount Percentile (Calculated Column)

This measure classifies discounts into "High Discount" and "Low Discount" based on the 20th percentile threshold, allowing for comparative analysis of discounting strategies across different products and categories.

Discount_Percentile = VAR DiscountValue = 'Orders'[Discount] VAR P20 = PERCENTILE.INC('Orders'[Discount], 0.20) RETURN IF(DiscountValue > P20, "High Discount", "Low Discount")

19. Profit Outliers Flag (Calculated Column)

It identifies profit outliers using the interquartile range (IQR) method, flagging transactions with unusually high or low profits to improve data accuracy and anomaly detection in profitability analysis.

Profit_Outlier_Flag = VAR Q1_Profit = PERCENTILE.INC('Orders'[Profit], 0.25) VAR Q3_Profit = PERCENTILE.INC('Orders'[Profit], 0.75) VAR IQR_Profit = Q3_Profit - Q1_Profit VAR Lower_Bound_Profit = Q1_Profit - (1.5 * IQR_Profit) VAR Upper_Bound_Profit = Q3_Profit + (1.5 * IQR_Profit) VAR Profit_Value = 'Orders'[Profit] RETURN IF(Profit_Value < Lower_Bound_Profit || Profit_Value > Upper_Bound_Profit, "Outlier", "Normal")

20. Sales Outlier Flag (Calculated Column)

This calculate column detects sales outliers using the interquartile range (IQR) method, flagging unusually high or low sales transactions to enhance data accuracy and trend analysis.

Sales_Outlier_Flag = VAR Q1_Sales = PERCENTILE.INC('Orders'[Sales], 0.25) VAR Q3_Sales = PERCENTILE.INC('Orders'[Sales], 0.75) VAR IQR_Sales = Q3_Sales - Q1_Sales VAR Lower_Bound_Sales = Q1_Sales - (1.5 * IQR_Sales) VAR Upper_Bound_Sales = Q3_Sales + (1.5 * IQR_Sales) VAR Sales_Value = 'Orders'[Sales] RETURN IF(Sales_Value < Lower_Bound_Sales || Sales_Value > Upper_Bound_Sales, "Outlier", "Normal")

21. Measure Selector (Table)

This table provides a selection option for users to toggle between "Sales" and "Profit" in visualizations.

Measure_Selector = DATATABLE("Measure", STRING, { {"Sales"}, {"Profit"} }

22. Time Series Table (Table)

This summarized table aggregates total sales over time, enabling time-series analysis of sales trends.

TimeSeriesTable = SUMMARIZE('Orders', 'Orders'[Order Date], "Total Sales", SUM('Orders'[Sales]))

Differences in Field Creation & Categorization in Power BI vs. Tableau

Field creation and categorization play a crucial role in structuring datasets for analysis. Power BI and Tableau take **different approaches** to calculated fields, categorization, and data modeling.

1. Power BI Does Not Have LOD Expressions, Sets, or Quick Table Calculations Like Tableau. How to Reproduce These in Power BI?

Tableau Feature	Power BI Equivalent	Implementation
LOD Expressions ({ FIXED })	**DAX Measures with CALCULATE & ALLEXCEPT**	Instead of { FIXED [State] : SUM([Profit]) }, use: DAX State_Profit_Margin = VAR TotalProfit = CALCULATE(SUM('Orders'[Profit]), ALLEXCEPT('Orders', 'Orders'[State])) VAR TotalSales = CALCULATE(SUM('Orders'[Sales]), ALLEXCEPT('Orders', 'Orders'[State])) RETURN DIVIDE(TotalProfit, TotalSales, 0)
Sets (Dynamic Segments)	**Groups or DAX Measures**	Instead of Sets, Power BI uses **Grouping** or a custom **DAX Measure**. Example: Create a custom category for **high-discount products**.
Quick Table Calculations	**DAX Measures**	Instead of dragging Quick Calculations like **Running Total**, we use DAX: DAX Running_Total = CALCULATE(SUM('Orders'[Sales]),

		FILTER(ALL('Orders'), 'Orders'[Order Date] <= MAX('Orders'[Order Date]))

2. Pros & Cons of Power BI's Approach Compared to Tableau's
3.

Feature	Power BI Approach	Pros	Cons
LOD Expressions	Uses CALCULATE(ALLEXCEPT(...)) to fix aggregation levels.	More control over filters & calculations.	More complex to write compared to { FIXED }.
Sets	Uses **Groups or DAX Calculations**.	Groups are easy to create and modify.	**Not as dynamic** as Tableau's **Sets**, where users can interactively filter within visualizations.
Quick Table Calculations	Requires **DAX Measures** instead of drag-and-drop.	More flexibility and customizability.	**Harder to learn** than Tableau's built-in Quick Calculations.

In conclusion, Power BI provides **powerful alternatives** to LODs, Sets, and Quick Table Calculations using **DAX**, but they require **a steeper learning curve** compared to Tableau's **drag-and-drop functionality**.

Implementing Filters in Power BI

In Power BI, filters can be applied at **three levels**:

1. **Visual-Level Filters** – Affects only a single chart or table.
2. **Page-Level Filters** – Applies to all visuals on a specific report page.
3. **Report-Level Filters** – Applies to all pages in the report.

Reflection on Filters in Power BI vs. Tableau

Similarities: Both **Power BI and Tableau** allow filtering at **different levels** (Visual, Page, Report in Power BI vs. Sheet, Dashboard, Story in Tableau). Both support **interactive filtering** (Power BI Slicers vs. Tableau Quick Filters). Both enable **dynamic filtering** using parameters.

Differences

Feature	Power BI	Tableau
Filter Types	Uses **Visual-Level, Page-Level, and Report-Level Filters**.	Uses **Sheet, Dashboard, and Story Filters**.
Interactivity	Uses **Slicers** for interactive selection.	Uses **Quick Filters & Dashboard Actions**.
Drillthrough	Allows users to **filter into deeper details**.	Uses **Dashboard Actions** to jump to details.
Global Filters	Report-Level Filters apply across **all pages**.	Filters in one dashboard may not affect another.

Pros & Cons of Power BI Filters

Pros: More control over filter levels (Report, Page, Visual). Slicers allow easy user interaction. Drillthrough enables in-depth analysis.

Cons: Lacks "Context Filters", which improve Tableau's filter efficiency. Some Drillthrough actions require setup.

Filters Used In This Project

Report-Level Outlier Filtering Using Outlier_Flag

1. Profit Outlier Detection

A **calculated column** named Profit_Outlier_Flag was created to classify profit values as **"Normal"** or **"Outlier"** using the **IQR method**:

Profit_Outlier_Flag = VAR Q1_Profit = PERCENTILE.INC('Orders'[Profit], 0.25) VAR Q3_Profit = PERCENTILE.INC('Orders'[Profit], 0.75) VAR IQR_Profit = Q3_Profit - Q1_Profit VAR Lower_Bound_Profit = Q1_Profit - (1.5 * IQR_Profit) VAR Upper_Bound_Profit = Q3_Profit + (1.5 * IQR_Profit) VAR Profit_Value = 'Orders'[Profit] RETURN IF(Profit_Value < Lower_Bound_Profit || Profit_Value > Upper_Bound_Profit, "Outlier", "Normal")

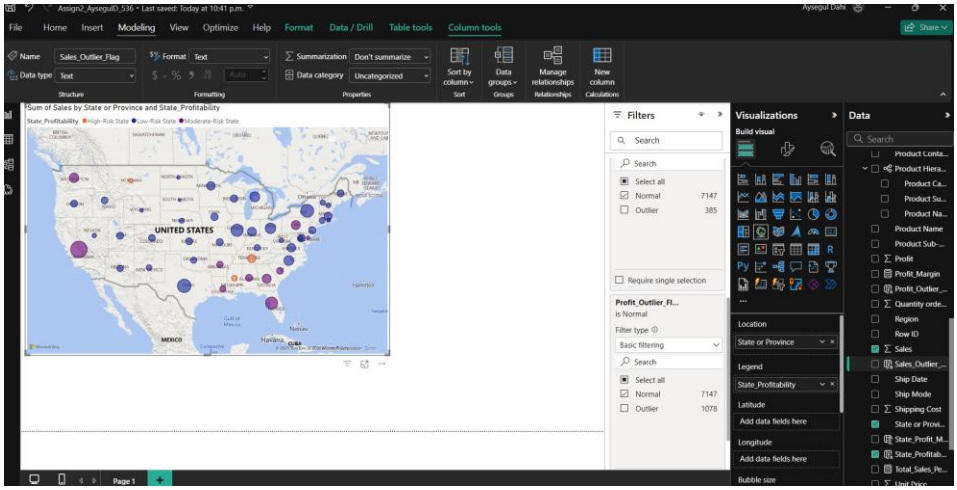
2. Sales Outlier Detection

Similarly, a **calculated column** named Sales_Outlier_Flag was created to classify sales values as **"Normal"** or **"Outlier"**:

Sales_Outlier_Flag = VAR Q1_Sales = PERCENTILE.INC('Orders'[Sales], 0.25) VAR Q3_Sales = PERCENTILE.INC('Orders'[Sales], 0.75) VAR IQR_Sales = Q3_Sales - Q1_Sales VAR Lower_Bound_Sales = Q1_Sales - (1.5 * IQR_Sales) VAR Upper_Bound_Sales = Q3_Sales + (1.5 * IQR_Sales) VAR Sales_Value = 'Orders'[Sales] RETURN IF(Sales_Value < Lower_Bound_Sales || Sales_Value > Upper_Bound_Sales, "Outlier", "Normal")

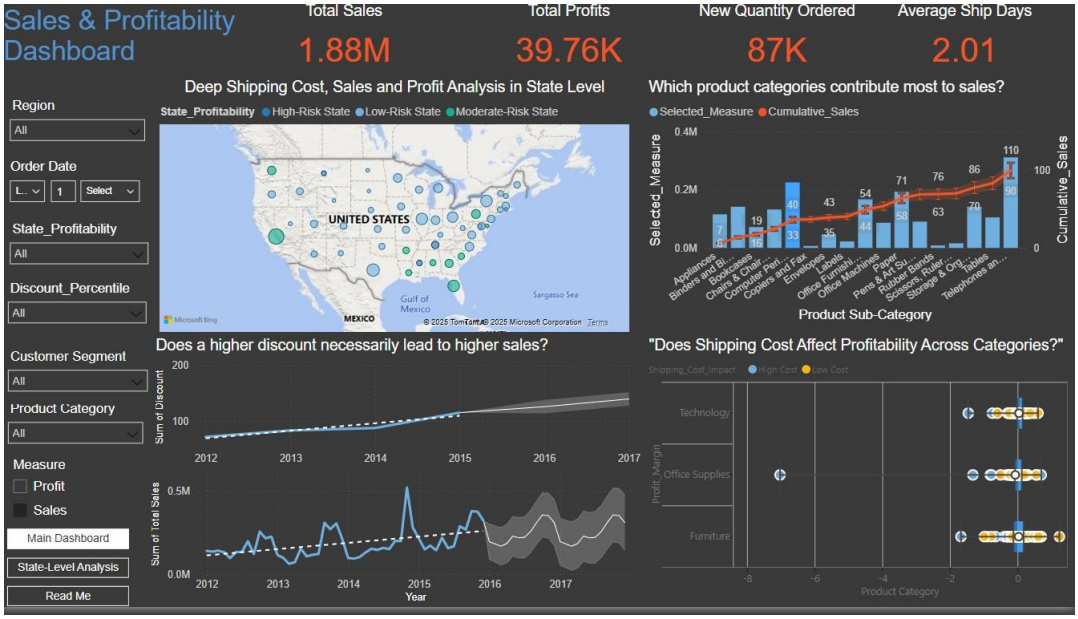
3. Applying Report-Level Filters

To prevent outliers from **skewing analysis**, both **Profit_Outlier_Flag and Sales_Outlier_Flag** were applied as **Report-Level Filters (All Pages)** in Power BI.



Page-Level Filters (Affecting the Entire Report Page)
Region Filter (Slicer): Located on the left panel, this slicer allows users to select specific regions (Central, East, South).
Effect: It applies to all charts on the page, ensuring that only data from the selected region is displayed.
Order Date Filter: Allows users to select a specific time frame for analysis.
Effect: Filters all visuals on the page to reflect data within the chosen date range.
State Profitability Filter (Dropdown): Enables users to filter based on profitability levels: High-Risk State, Low-Risk State, Moderate-Risk State.
Effect: This filter applies to all visuals, including the map and profit-related visuals.
Discount Percentile Filter (Dropdown): This filter segments the data based on High Discount vs. Low Discount.
Effect: Applies across all visuals where discount plays a role, ensuring consistency.
Customer Segment Filter (Dropdown): Users can choose a specific customer segment (e.g., Corporate, Consumer, Home Office).
Effect: Impacts all visuals related to sales and profitability based on customer segmentation.
Product Category Filter (Dropdown): Filters data to show specific product categories.

Effect: Ensures only relevant sales and profit data for selected product categories are displayed in all visuals.
Measure Selector (Profit vs. Sales) (Dropdown): Allows toggling between Sales and Profit for relevant visuals.
Effect: Updates multiple charts to dynamically switch between profit and sales data.



Visual-Level Filters (Affecting Only a Single Chart/Table)

State Profitability in the Map: The State_Profitability legend (High-Risk, Low-Risk, Moderate-Risk) is a visual-level filter applied only to the map.

Effect: The map updates based on profitability categories, but other visuals remain unchanged.

Cumulative Sales vs. Selected Measure Chart: The dropdown Selected Measure (Profit or Sales) impacts only this chart.

Effect: The Pareto chart updates, but other charts do not.

Profit Margin Box Plot (Shipping Cost Impact): The Shipping_Cost_Impact filter (High Cost, Low Cost) applies only to this specific visual.

Effect: It does not affect other charts or tables.

Charts Used in This Project

1. **Map: Deep Shipping Cost, Sales, and Profit Analysis in State Level**

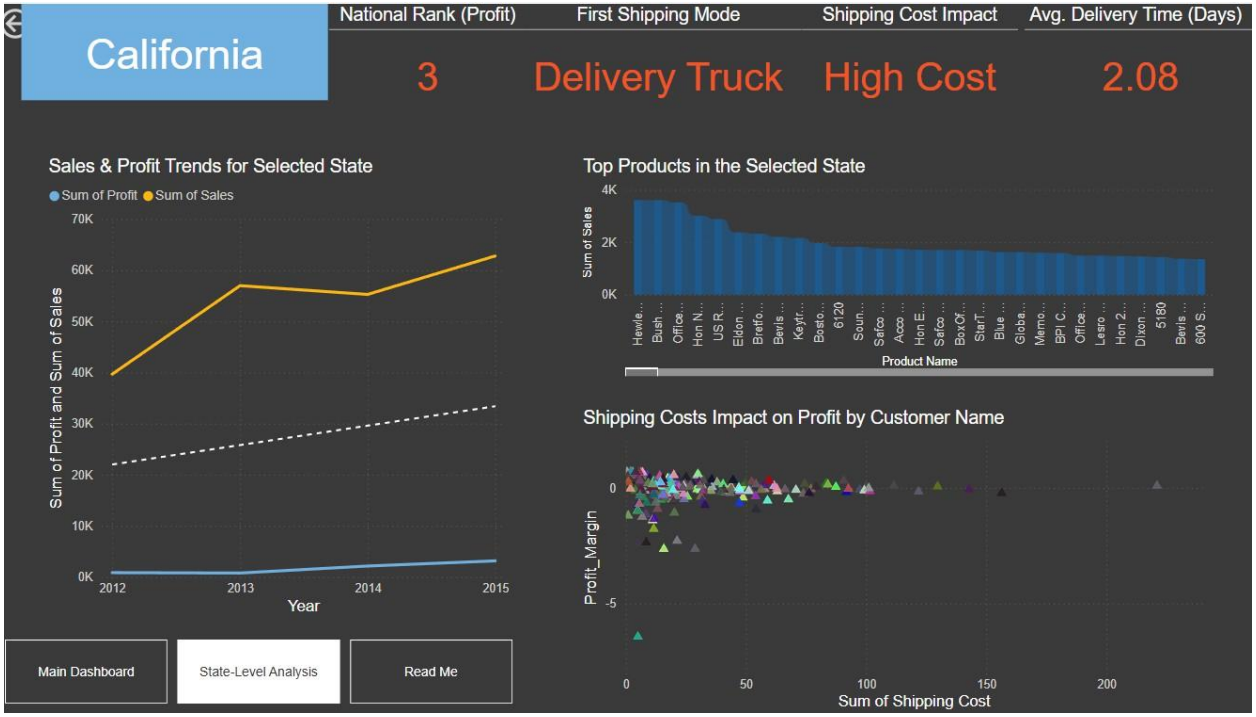
This map visualization provides a comprehensive analysis of **state-level shipping costs, sales, and profitability** across the United States. It helps answer the question: Which Regions Have High Shipping Costs Impacting Profits? The geographic visualization allows to analyze regional variations in sales, profit margins, and the impact of shipping costs.

The **state profitability classification** is represented using a **color-coded system**:

High-Risk States (Navy): States with **negative profit margins**, indicating that high shipping costs are leading to overall losses.
Low-Risk States (Blue): States with **stable or positive profit margins**, where shipping costs are not significantly impacting profitability.
Moderate-Risk States (Green): States that have **moderate profitability**, suggesting a balance between sales revenue and shipping costs.
Key Insights:

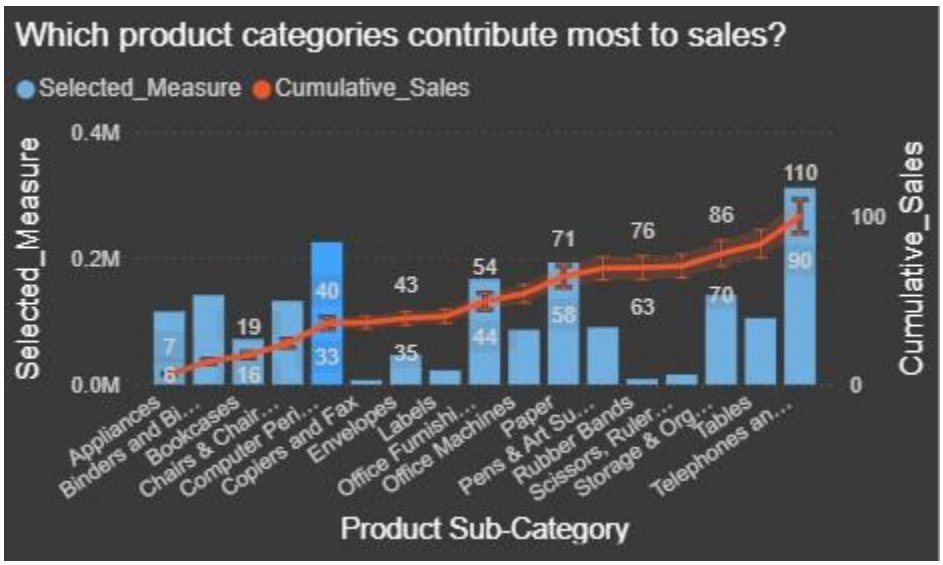
Profitability Trends by Region: The **Eastern and Southern** states tend to have **more moderate to high-risk** classifications, suggesting higher shipping costs relative to sales.
High Shipping Cost Regions: Larger circle sizes indicate states with **higher shipping costs**, implying a direct relationship between **logistics expenses and profitability.**
Identifying Strategic Adjustments: Businesses operating in high-risk states should **optimize shipping strategies**, consider **regional fulfillment centers**, or adjust **pricing models** to maintain profitability.

This map allows users to **drill through** by **right-clicking on a state** to access **state-level profitability insights**, including sales trends, profit margins, and shipping cost impacts.



2. Pareto Chart: Identifying Key Product Sub-Categories

This Pareto Chart provides insights into how different **Product Sub-Categories** contribute to overall **Sales** and follows the **80/20 Rule** (Pareto Principle). It answers the following key business questions: Which product sub-categories contribute the most to total sales? Do a few product categories drive the majority of revenue? Are there sub-categories where sales and profit trends differ significantly?



Insights from the Pareto Analysis:

Pareto Principle Confirmation (80/20 Rule): A few top-selling product categories contribute the majority of total sales, while many others generate minimal revenue. The chart visually confirms the Pareto Principle by showing how approximately 80% of total sales come from the top-performing sub-categories.

Identifying Underperforming Sub-Categories: Sub-categories contributing **less than 5%** to total sales are potential areas for review in terms of profitability, marketing, or inventory management.Categories such as **Phones, Chairs, and Binders** appear to be high performers, while others may require strategic adjustments.

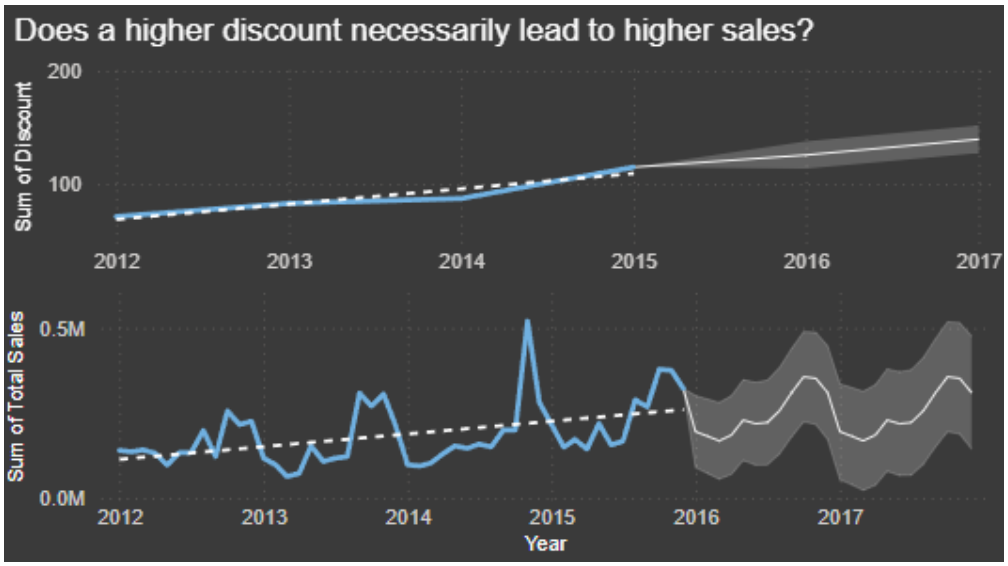
Cumulative Sales with Confidence Intervals: The **error bars** (as shown in the settings) provide a **range of expected variation** in cumulative sales, adding an extra layer of **statistical confidence** to decision-making.

Selected Measure Parameter: Users can toggle between **Sales or Profit** allowing dynamic analysis of high-sales but low-profit products.

Cumulative Sales Calculation: The cumulative percentage helps in identifying breakpoints where additional product categories **no longer significantly impact total sales**.

3. **Discount vs. Sales Forecast Chart**:

This time-series analysis explores whether **increasing discounts** results in a **significant boost in sales** over time. It helps answer critical business questions: Does offering higher discounts directly drive more sales? What are the long-term trends in sales and discounting strategies? Does the forecast suggest a continued correlation between discounts and sales growth?



Discount Forecast (Top Chart)

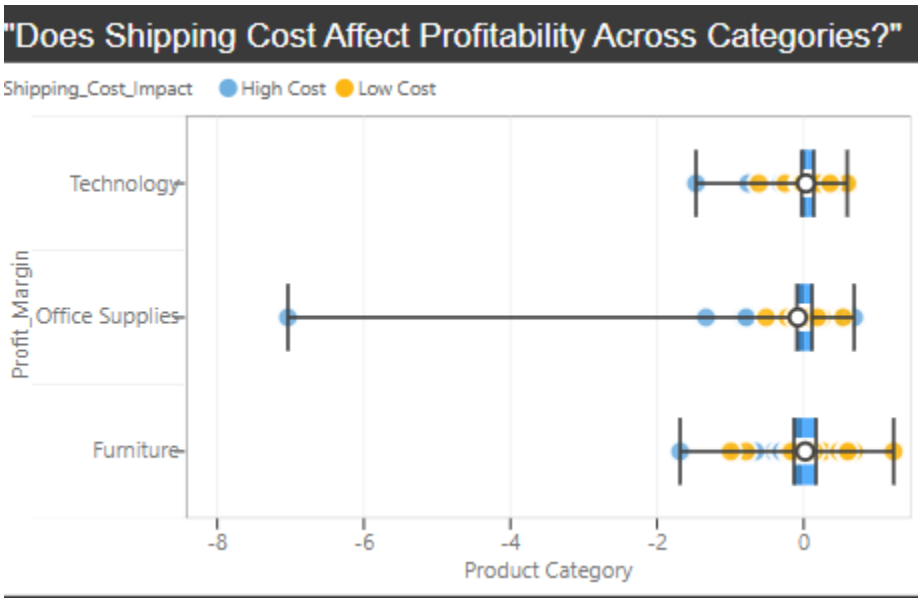
The forecast suggests that discount values will continue rising steadily over the next two years. The projected growth aligns with previous trends, implying businesses are increasing discounts to drive sales.This may indicate a competitive pricing strategy where higher discounts are used to maintain or expand market share.

Sales Forecast (Bottom Chart): While sales are projected to grow, fluctuations remain due to external factors such as seasonality, economic conditions, or market demand. The shaded forecast area represents confidence intervals, showing high variability in future sales projections. This suggests that **discounting alone does not directly guarantee higher sales**—other elements like customer demand and promotional strategies play a role.

Key Insights & Business Implications: Discounts are expected to increase, but their impact on sales remains inconsistent. The sales trend is upward, yet volatile, indicating that discounts may not be the only driver of revenue. Businesses should analyze customer behavior, product demand, and external economic factors alongside discounting strategies for sustainable sales growth.

4. **Box Plot: Does Shipping Cost Affect Profitability Across Categories?**

This box plot visualizes the relationship between shipping costs (High vs. Low) and profitability across different product categories, addressing the question: *How does shipping cost impact profitability across different product categories?*



Profit Margin Variability Across Categories: The Office Supplies category exhibits high variability, with some products experiencing negative profit margins despite lower shipping costs. Technology and Furniture show more consistent profit margins, but still, products with higher shipping costs tend to have slightly lower profitability.

Impact of High vs. Low Shipping Costs: High-cost shipments (blue circles) tend to have lower profit margins compared to low-cost shipments (yellow circles). The spread of data points suggests that profitability fluctuates significantly for Office Supplies, potentially due to lower-margin products or discounts.

Furniture and Technology Are Less Affected by Shipping Costs: Technology products show less impact from shipping cost fluctuations, suggesting that higher product value offsets shipping costs more effectively. Furniture products appear to have a stable profitability range, possibly due to bulk pricing strategies or efficient logistics management.

Key Business Takeaways: Office Supplies may require an optimized shipping strategy to minimize cost-related profit reductions. High-value product categories (Technology, Furniture) are less sensitive to shipping costs, reinforcing the need for pricing strategies that accommodate logistics expenses. The company should assess shipping cost structures per category to identify opportunities for cost reductions or alternative shipping methods to maximize profitability.

Comparison of Filters and Interactivity Features in Power BI and Tableau

Similarities and Differences Between Interactivity Features in Power BI and Tableau

Similarities: Both Power BI and Tableau offer interactive filtering options that allow users to dynamically explore data. Power BI's Slicers and Drill Through and Tableau's Parameters and Actions serve a similar purpose in enabling user-driven analysis.Both tools support cross-filtering and highlighting, meaning selecting one filter can dynamically update all related visualizations. Both platforms allow hierarchical filtering, where selecting a high-level category (e.g., Region) updates subcategories (e.g., States).

Differences:

Power BI:

Slicer & Drill Through: Slicers are interactive visual filters that apply selections across the report. Drill Through allows users to right-click on a data point to navigate to a detailed page focused on that selection. Interactivity is controlled using Edit Interactions, which allows specific

visuals to be affected by filters.

Tableau: Filters & Parameters: Filters in Tableau are more flexible and can be applied at the worksheet, dashboard, or data source level. Parameters allow users to create customizable input fields, which can dynamically change visualizations based on user inputs. Actions (Filter, Highlight, URL, Set) offer interactive ways to navigate between dashboards and highlight data relationships.

Differences Between Power BI's Features and Tableau's Filters:

Feature	Power BI	Tableau
Basic Filtering	Uses **Slicers**, which are interactive filter visuals placed in reports.	Uses **Filters** at the worksheet, dashboard, or data source level.
Interactivity	Uses **Drill Through** for deeper insights and custom navigation.	Uses **Filter and Set Actions** for more advanced interactivity.
Parameters	No direct equivalent to Tableau's **Parameter** feature. Uses **Measures and DAX** for dynamic selections.	Allows users to input values and dynamically update multiple visuals.
Hierarchical Filtering	Uses Drill Down in visualizations and **Sync Slicers** across pages.	Uses Filter Hierarchies and Field Actions.
Navigation Between Pages	**Drill Through & Page Navigation Buttons** enable moving to detail pages.	**Dashboard Navigation & Filter Actions** allow smooth transitions.

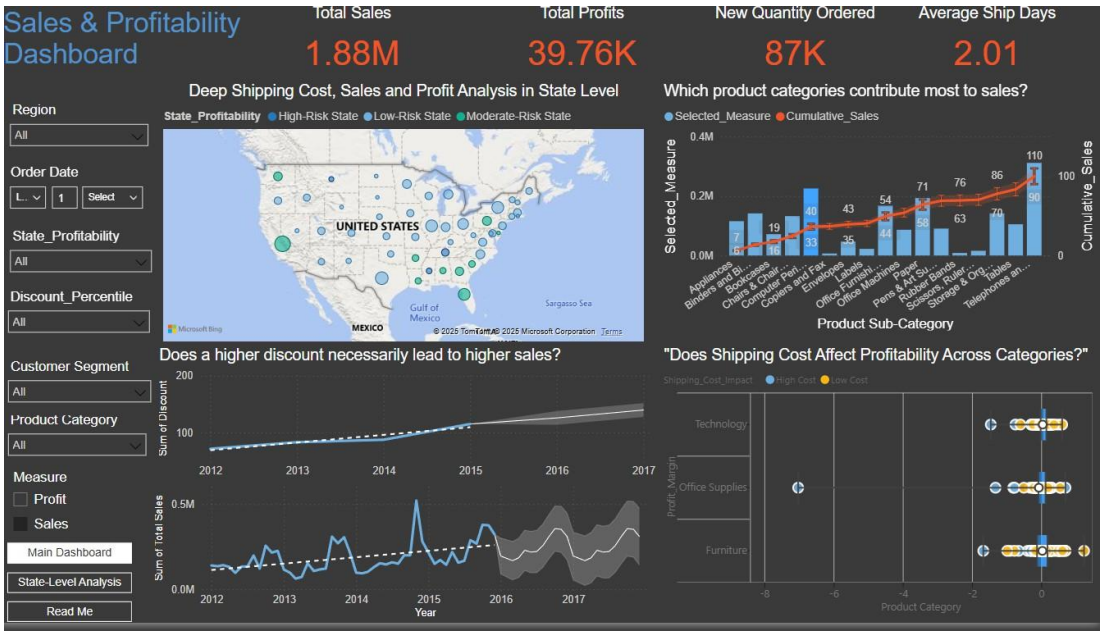
Pros and Cons of Power BI's Features Compared to Tableau's

Feature	Power BI: Pros	Power BI: Cons	Tableau: Pros	Tableau: Cons
Filters	Easy-to-use Slicers, supports hierarchical filtering, and persistent filtering.	Limited customization for advanced filtering.	Highly customizable, supports global and local filters.	Can be complex to set up multiple filters.
Drill Through & Navigation	Drill Through provides **clear navigation** between summary and detail pages.	Drill Through **requires setup** and is not as flexible as Tableau's actions.	Filter Actions allow more dynamic drill-downs without predefined pages.	Navigation across dashboards **requires extra setup**.
Parameters & Dynamic Selections	Uses DAX and measures for dynamic analysis.	Lacks direct **parameter input fields** like Tableau.	Parameters allow users to input custom values and dynamically update visuals.	Requires more effort to set up calculations.
Interactivity	**Sync Slicers and Edit Interactions** allow good control.	Drill Through requires separate **dedicated pages**.	**More intuitive interactivity** through Filter and Highlight actions.	Filter Actions can sometimes be too **sensitive** or difficult to control.

Both Power BI and Tableau provide strong filtering and interactivity features, but Power BI focuses on structured navigation and ease of use, whereas Tableau offers more dynamic interactions with customizable filters and parameters.

Sales & Profitability Dashboard Interpretation

Purpose: This dashboard provides deep insights into sales, profitability, shipping costs, and regional trends across different product categories and states. It enables users to interact with filters, drill through state-level analysis, and identify patterns affecting business performance.



1. **Regional & Customer Analysis**

   A. Which regions have high shipping costs impacting profit margins?

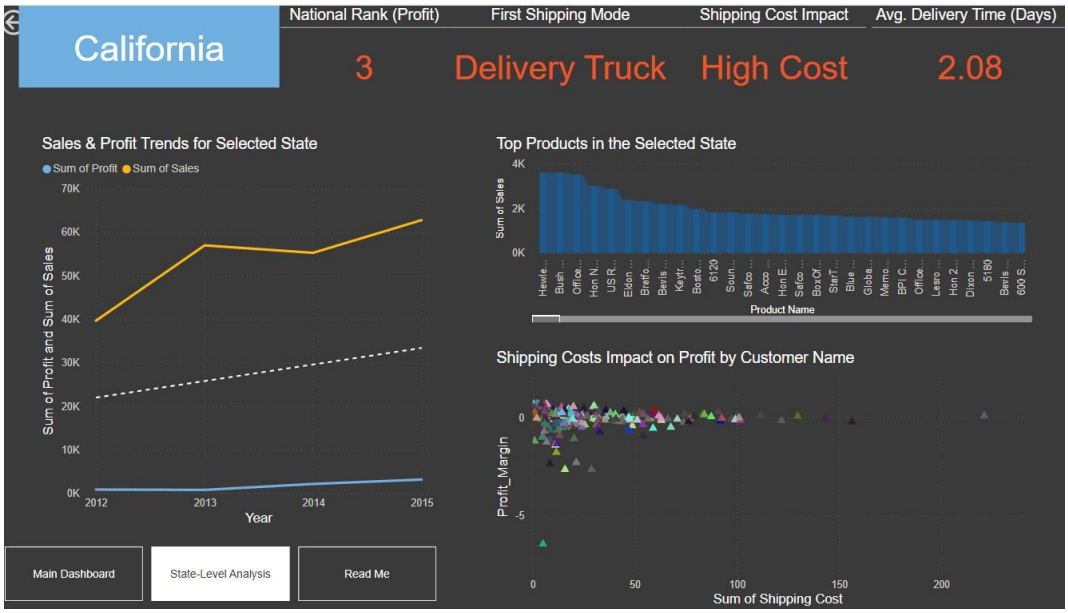
The **state profitability map** categorizes states into high-risk, moderate-risk, and low-risk based on their profit margins. Larger blue bubbles indicate low-risk states with stable profitability, while smaller or different-colored bubbles represent regions with fluctuating profits.
Key Insight: Shipping costs appear to impact southeastern states more heavily, contributing to lower profit margins.

   B. Are certain states consistently low-profit due to logistics expenses?

The **drill-through state analysis page** provides a deeper view of individual states. California, for instance, ranks **third** in profitability but has **high shipping costs**, indicating logistics could be affecting profits. The **scatter plot on shipping cost vs. profit margin** reveals how different customers and regions experience profitability variations.

   C. How does shipping mode affect cost and profitability?

First shipping mode in the drill-through view indicates the dominant shipping method for each state. Regular delivery trucks appear to be the most common mode, but it's also associated with high costs in some regions.



2. Sales & Profitability Trends

A. Does a higher discount necessarily lead to higher sales?

The **discount vs. sales trend line** on the bottom-left chart tracks how discounts change over time. While discounts show an **upward trend**, sales do not always correspond with higher discount percentages. This suggests that discounting strategies should be optimized based on product categories rather than blanket reductions.

B. Which product sub-categories contribute most to total sales?

The **Pareto Chart (Product Sales Distribution)** ranks sub-categories by their total sales contribution. The **cumulative sales line (orange line)** follows the **80/20 rule**, confirming that **a few top categories drive the majority of revenue**. Phones, chairs, and binders contribute significantly to total sales, making them high-priority for inventory and marketing strategies.

C. Are there sub-categories where sales and profit trends differ significantly?

By switching between **profit and sales views using the selector**, users can analyze discrepancies. Some categories may generate high sales but low profit due to high costs or discounts.

Drill-Through Analysis Features

How the Drill-Through Works?

Right-clicking on a state or product category allows the user to navigate to the state-level analysis page. This drill-through feature allows a **focused deep dive** into specific locations, analyzing: Profitability Rank, Top Selling Products, Sales & Profit Trends Over Time

And Shipping Costs and Their Impact on Profitability in Customer Name level.

Visual and Functional Enhancements

Interactive Filters

Region, State Profitability, Customer Segment, and Discount Percentile filters allow dynamic exploration. Users can toggle between different **profitability tiers** to isolate high-risk states.

Buttons & Navigation

Navigaton buttons between Main Dashboard, State-Level Analysis and Read Me pages improve usability. Ensures users can seamlessly move between overall insights and specific state-focused investigations.

Comparison: Power BI vs. Tableau Dashboard Structure

Power BI's Dashboard-Report-Page Structure

Dashboards: A Power BI dashboard is a **single-page canvas** where pinned visuals from different reports are displayed. It provides a **high-level view** of KPIs and insights.
Reports: A report is a **multi-page** collection of interactive visualizations **built from one dataset**. Reports allow **detailed exploration and filtering**.
Pages: Inside a report, multiple **pages** can be created to **organize content** by category, making it easier to navigate and explore data.
Interactivity: Reports allow **Drill Through**, **Drill Down**, and **Cross-Filtering**, but dashboards **only allow visual interactions** (no direct drill-through).

Tableau's Workbook-Sheet-Dashboard-Story Structure

Workbooks: Similar to an Excel file, a Tableau workbook contains **multiple sheets and dashboards**.
Sheets: Each sheet contains an individual **visualization**, which can be combined into dashboards.
Dashboards: A **collection of visualizations from different sheets**, allowing **dynamic interactivity** through actions and filters.
Stories: Tableau offers a **story feature**, which is a **narrative** of data insights, guiding users through an analysis step by step.

Pros and Cons: Power BI vs. Tableau's Dashboarding Approach

Feature	Power BI (Dashboard-Report-Page)	Tableau (Workbook-Sheet-Dashboard-Story)
Data Model	Works best with **structured datasets** (SQL, Azure, Excel).	More flexible for **multiple data sources** (SQL, Google Sheets, cloud data).
Dashboard Structure	Dashboards are **single-page summaries** with pinned visuals from reports.	Dashboards are **multi-sheet interactive pages** with dynamic filtering.
Report Organization	Reports have **multi-page layouts** for structured insights.	Workbooks allow **multiple sheets** per dashboard.
Drill-Through & Actions	Uses **Drill-Through & Slicers** for deep analysis.	Uses **Filter Actions & Set Actions** for dynamic exploration.
Storytelling	No dedicated **Story** feature (relies on report pages).	**Story Feature** allows **narrative-driven insights**.
Ease of Use	Easier for **business users** and integrates well with **Excel & Microsoft tools**.	More **analyst-friendly**, requires understanding of **Tableau Calculations**.
Customizability	**Less customizable** in dashboard layout compared to Tableau.	Highly **customizable layout and interactivity**.

Justification of Power BI Dashboard Design

The Sales & Profitability Dashboard was designed to provide comprehensive insights into sales trends, profitability, and shipping costs across different states, product categories, and customer segments.

Why This Structure?

Main Dashboard (High-Level Overview)

The main dashboard provides summary metrics (Total Sales, Profit, Average Delivery Time). Regional profit distribution (via map) allows quick identification of high and low-performing states. Trend analysis charts track sales performance over time. Pareto analysis helps in understanding which product categories contribute most to revenue.

State-Level Analysis (Drill Through for Detailed Insights)

The Drill-Through Page allows users to select a specific state and view profitability trends, shipping costs, and product performance. Ranking of states helps in comparing performance across the US.

Conclusion

Power BI's structured report-page format was chosen because it ensures a clear, hierarchical flow of insights, making it easier for stakeholders to explore high-level data and then drill down into specific details. While Tableau provides greater flexibility with dashboard layouts and story-driven narratives, Power BI's structure ensures business users can access structured insights with minimal effort.