blinkit

Power BI

Internship Under:

Classr ∞ m

Phone:

8981838547

Website:

https://classroomtech.in

LinkedIn:

https://www.linkedin.com/company/classroom-tech/

Duration:

3 months 1st April 2025 to 30th June 2025

📊 Project Title: Blinkit Sales Data Analysis

Name: Srimayee Bhattacharyya
College Name: St. Thomas' College of Engineering and Technology
Department: Computer Science and Engineering
 Semester: 7th semester
Internship Under Classroom and add the below details
📞 Contact Us
📱 Phone: 8981838547
🌐 Website: https://classroomtech.in
🔗 LinkedIn:
https://www.linkedin.com/company/classroom-tech/
📅 Duration: 3 months — 1st April 2025 to 30th June 2025

Acknowledgement

I would like to express my heartfelt gratitude to Classroom Tech for providing me with the opportunity to work on the Blinkit Sales Data Analysis project as part of my internship. This experience helped me gain valuable insights into data visualization, business intelligence, and real-world decision-making using Power BI.

I am especially thankful to Satyaki Sir for his constant support, constructive feedback, and encouragement throughout the internship.

I also want to thank my family and close friends for their unwavering support and motivation, which played a vital role in the successful completion of this project.

Table of Contents

1 Objective
Goal of the project and key analytical questions

2 Dataset Overview
Data source, size, and key fields used in analysis

3 Tools Used
Power BI, GitHub

4 Methodology
Data-cleaning steps and the rationale behind chosen visualizations

5 Visualizations & Insights
Individual charts with purpose and key take-aways

6 Key Findings
3–5 headline insights distilled from the analysis

7 Slide Headline
Lorem ipsum pulviar proin gravia hendrerit ctus.

8 Conclusion
Lessons learned and how the insights support decision-making

9 GitHub Link
URL to code, Power BI file, and datasets

10 Thank You

Objective / Problem Statement

- Goal: To analyze Blinkit's sales data to uncover business insights related to customer behavior, delivery trends, campaign effectiveness, and product performance.
- Key Questions:
1. Which products are most frequently purchased?
2. How do customer segments impact order volume?
3. What campaigns generate the highest revenue and ROAS?
4. What are common customer feedback categories?

Dataset Overview

Dataset Name: Blinkit Sales Dataset

Source: Kaggle : https://www.kaggle.com/datasets/akxiit/blinkit-sales-dataset

blinkit_orders.csv
- 26,229 rows × 10 columns
- Key columns: order_id, customer_id, order_date, product_id, quantity, order_total, delivery_status, delivery_time_minutes, rating, pincode_new

blinkit_products.csv
- 40 rows × 7 columns
- Key columns: product_id, product_name, category, price, stock_received, damaged_stock, discount_pct

blinkit_customer_feedback.csv
- ~5,000 rows × 5 columns
- Key columns: order_id, customer_id, rating, feedback_category, feedback_text

blinkit_campaigns.csv
- 7 rows × 4 columns
- Key columns: campaign_id, campaign_name, spend, revenue_generated

blinkit_stock_forecast.csv
- 36 rows × 3 columns
- Key columns: month_year, stock_received, product_count

blinkit_customer_details.csv
- ~5,000 rows × 4 columns
- Key columns: customer_id, customer_name, email, segment

blinkit_delivery_analysis.csv
- ~26,000 rows × 4 columns
- Key columns: order_id, delivery_status, delivery_time_minutes, distance_km

blinkit_rating_trends.csv
- ~10,000 rows × 6 columns
- Key columns: rating, feedback_text, year, quarter, month, day

Tools Used

• Power BI – For data visualization, dashboard creation, and insights extraction
 • GitHub – For version control and storing the project files, datasets, and report

Methodology

• Data Preprocessing:
 The dataset from Kaggle was already well-structured and cleaned. No major missing values or outliers were found during initial inspection. Basic validation was done to ensure consistency across files (e.g., matching product_id, order_id, and customer_id across tables).

• Visualization Process:
 Power BI was used to create interactive dashboards and visualizations including:
 – Bar Charts for order counts, product sales, and feedback distribution
 – Line Charts and Time Series for monthly trends and stock forecasts
 – Donut Charts for customer segments and feedback categories
 – KPI Cards to highlight revenue, customer retention, and order value
 – Funnel Charts and Decomposition Trees to track campaign performance and segment insights

Detailed Visualizations

1. Number of orders placed per customer

   - Type: Horizontal Bar Chart
   - Purpose: To identify which customers are placing the most orders.
   - Insight:
     - Top customers like Vedant, Bhanumati, and Falan Mastani are responsible for higher order counts.
     - Most customers place 1–4 orders.


Number of orders placed per customer
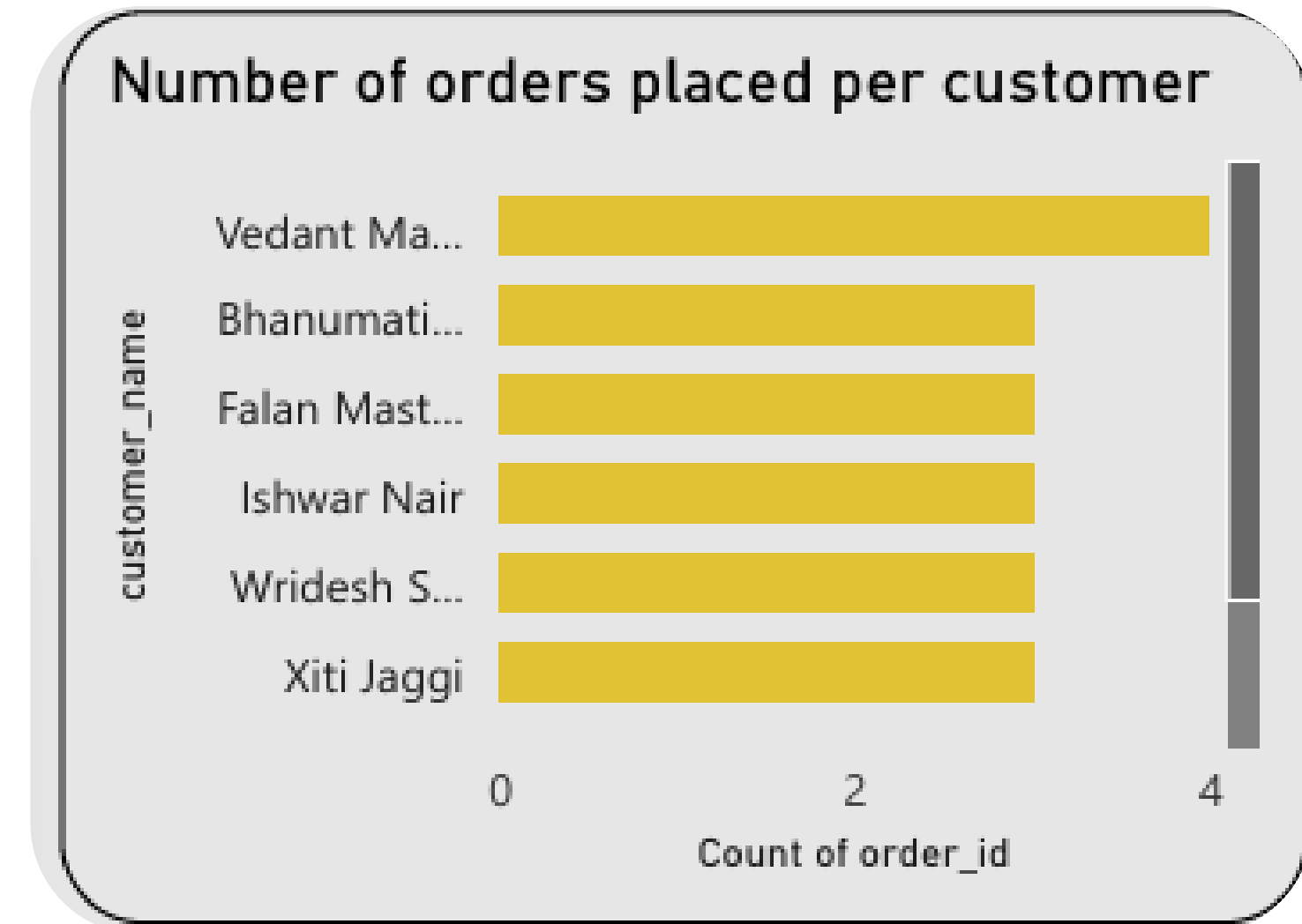
2. Table visualization showing customer details (customer_id, name, email)

- Type: Data Table
- Purpose: To show raw customer data (ID, name, email).
- Insight:
  - Useful for mapping orders to users or conducting targeted campaigns.

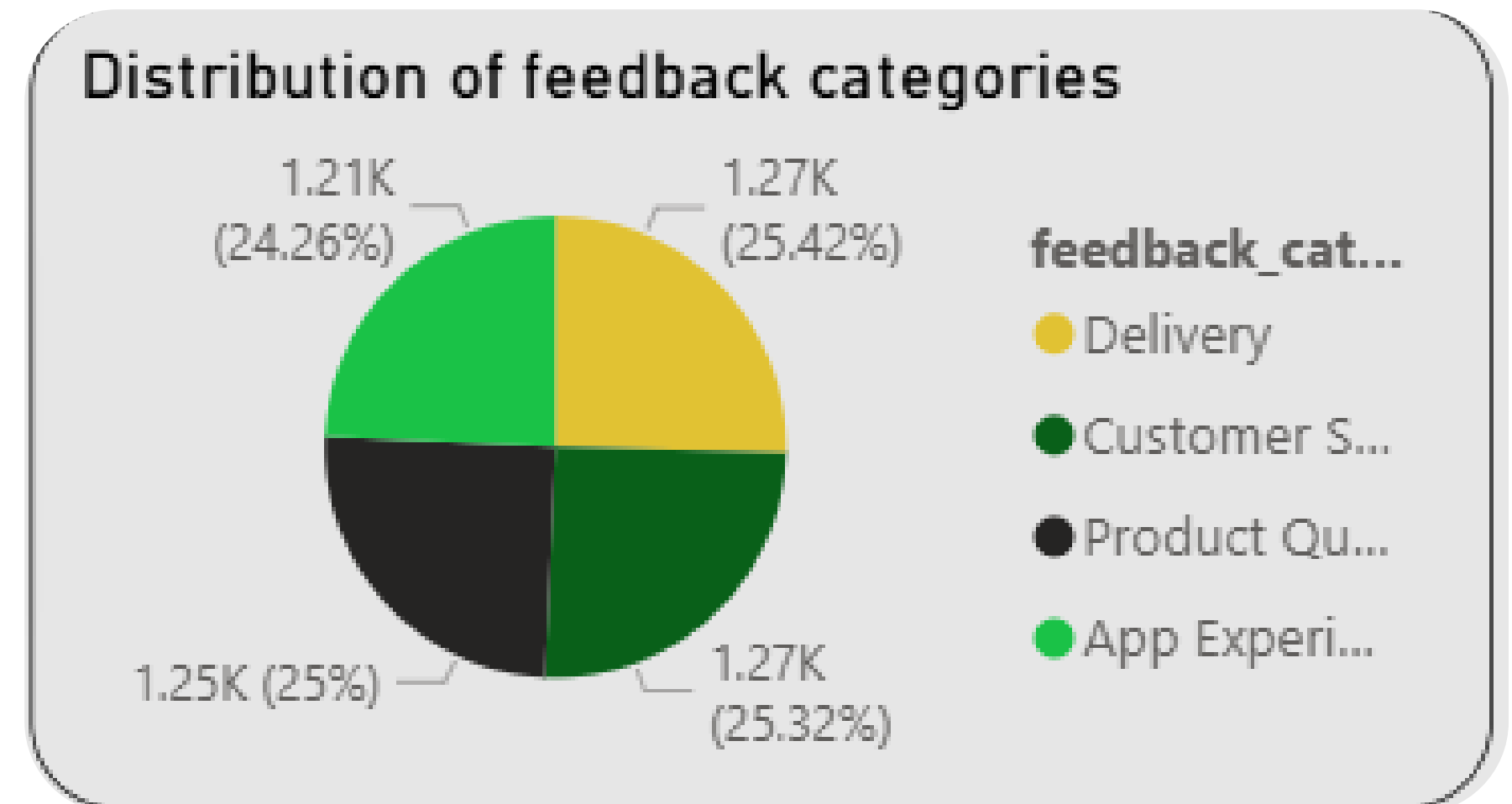| customer_id	customer_name	email
80315623	Varenya Banik	aachalsarma@example.com
70230254	Hamsini Sandhu	aachalsheth@example.net
14161586	Aachal Nazareth	aadhya71@example.org
61226563	Chameli Kala	aadhya91@example.com
74735607	Naksh Amble	aadhyabhalla@example.com
71615827	Yash Contractor	aadi81@example.com
27955158	Abeer Kamdar	aaggarwal@example.net
49915498	Harini Mangal	aahana49@example.com
43459807	Charan Brahmbhatt	aahanatailor@example.net

3. Total number of orders placed in a given month

- Type: Column Chart
- Purpose: To analyze order trends over months.
- Insight:
  - Orders steadily increased from Jan to August, peaking mid-year.
  - Slight drop toward year-end—might be seasonal or due to campaign gaps.

Total number of orders placed in a given month
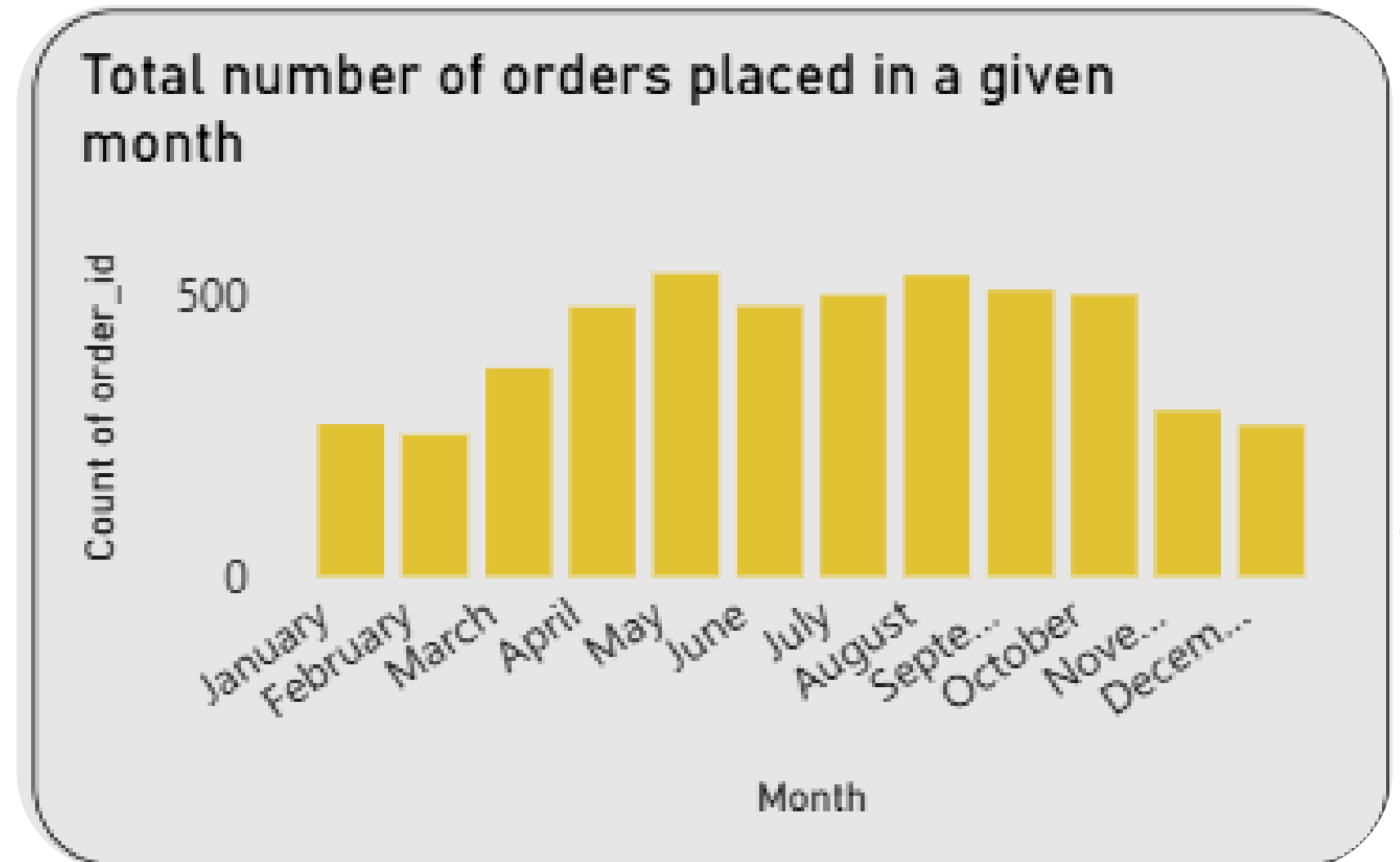
4. Distribution of feedback categories

- Type: Pie Chart
- Purpose: To visualize customer concerns.
- Categories:
  - Delivery, Customer Service, Product Quality, App Experience
- Insight:
  - All feedback types are nearly evenly distributed (~25% each).
  - Indicates multiple areas need improvement, not just one.

Distribution of feedback categories

1.21K (24.26%)
1.27K (25.42%)
1.25K (25%)
1.27K (25.32%)

feedback_cat...
- Delivery
- Customer S...
- Product Qu...
- App Experi...

5. Delivery Status Filter

- Purpose:
  - To filter and analyze orders based on delivery performance categories.
- Visual:
  - slicer(delivery_status)
- Insights:
  - Enables breakdown of orders by: On Time, Slightly Delayed, or Significantly Delayed
  - Helps identify how delivery timing affects customer ratings and feedback
  - Useful for evaluating logistics efficiency and optimizing delivery operations

delivery_status ⌄

☐ On Time

☐ Significantly Delayed

☐ Slightly Delayed
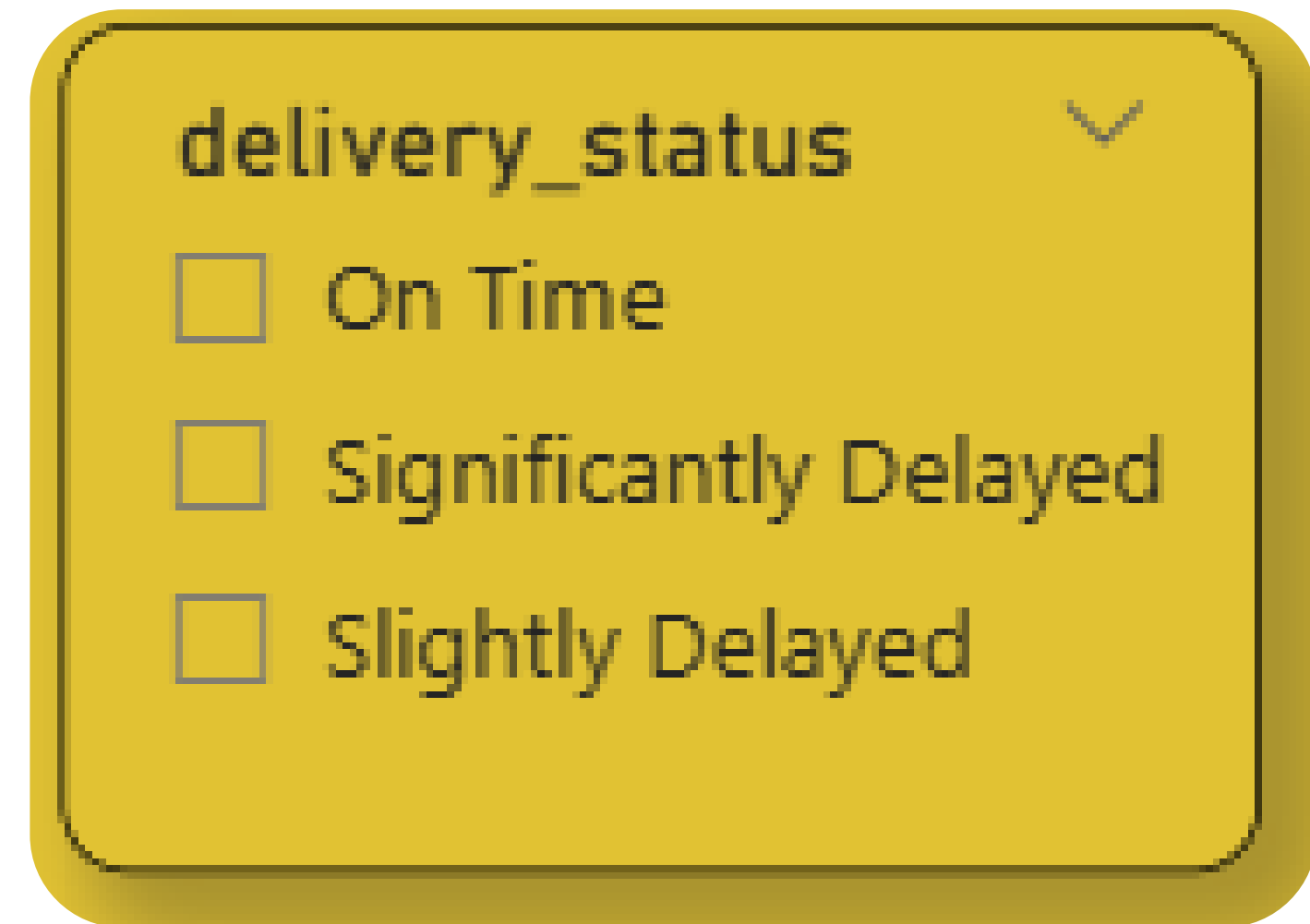
6. Stock received over time for different products

- Purpose:
- To track the quantity of stock received across different products and identify supply trends.
- Visual:
- Clustered column chart (Sum of stock_received) + Line chart (Count of product_name)
- X-axis: product_name | Y-axis: Dual (stock & product count)
- Insights:
  - Products like Biscuits, Curd, and Cola show high stock volume, indicating frequent restocking
  - Lesser restocks for products like Instant Noodles and Frozen Pizza, possibly due to lower demand or storage limits
  - The line trend shows that most categories received stock for 5–10 product variants
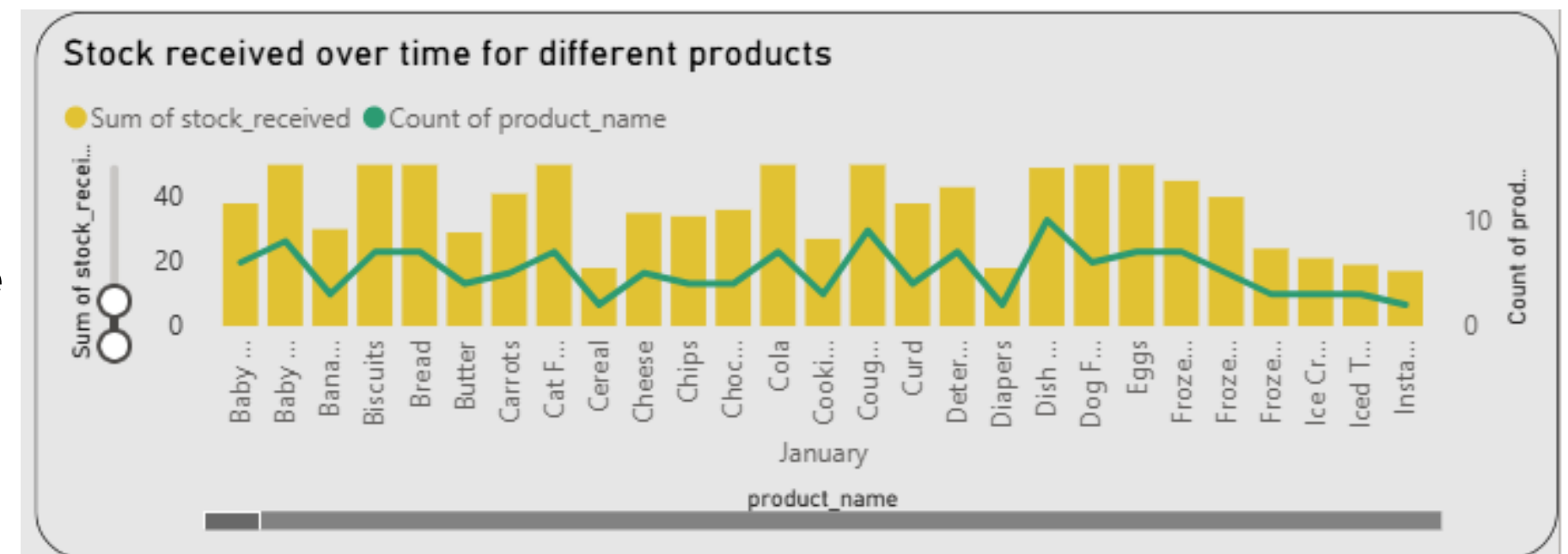
7. Percentage of damaged stock per product

- Purpose:
- To identify which products have the highest percentage of stock damage, helping optimize inventory handling and packaging.
- Visual:
- Horizontal bar chart
- X-axis: Percentage of damaged_stock
- Y-axis: product_name
- Insights:
  - Pet Treats and Pulses have the highest damage percentages, indicating possible handling or packaging issues
  - Dish Soap and Lotion also show significant damage rates — may require improved shipping methods
  - Useful for warehouse teams to prioritize quality control on frequently damaged items

Percentage of damaged stock per product
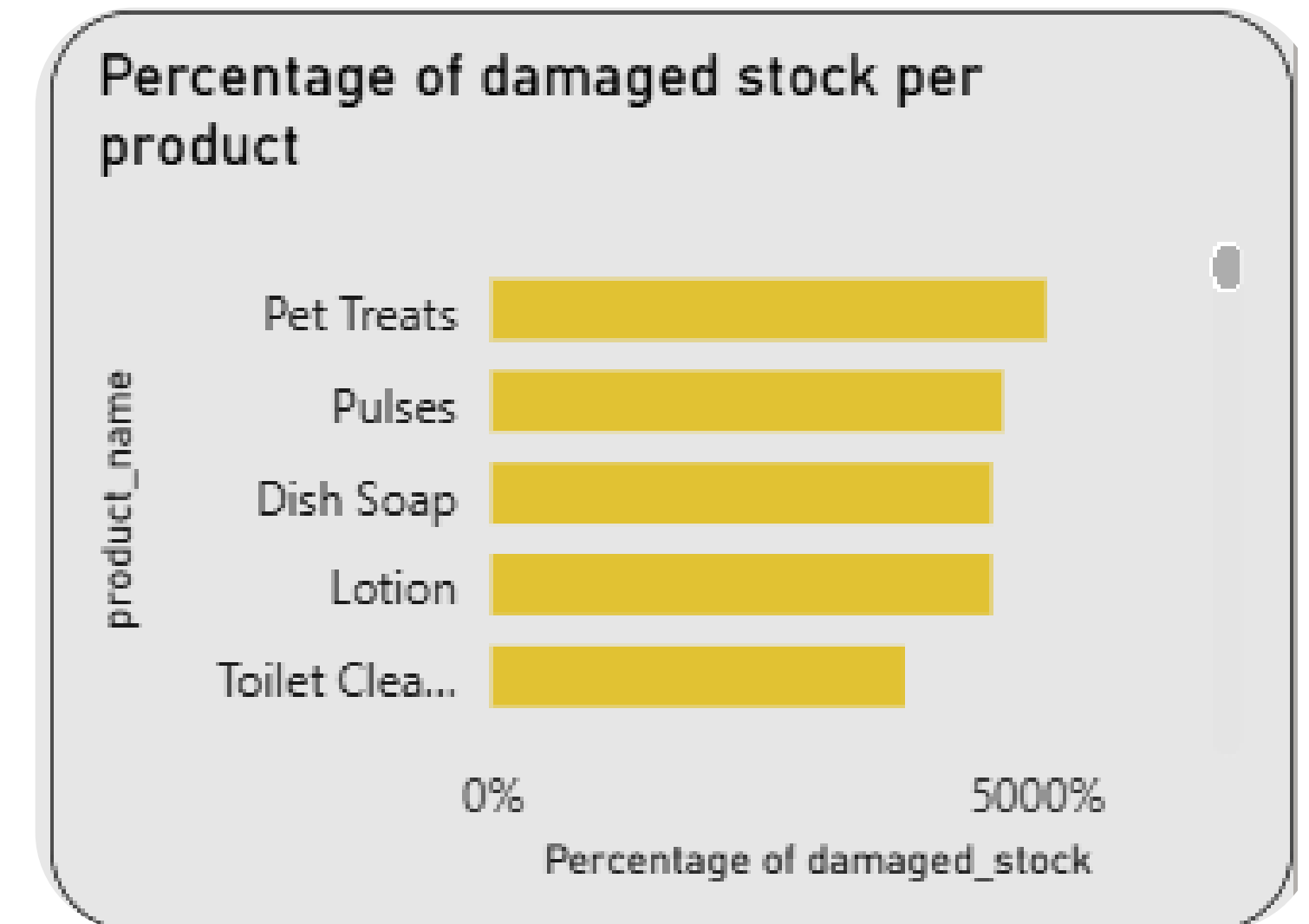
8. Table with all campaigns (campaign_id, campaign_name, spend, revenue_generated)

- Purpose:
  - To evaluate how effective each marketing campaign was in terms of return on investment (ROI) by comparing spend vs revenue generated.
- Visual:
  - Data table showing campaign_name, spend, and revenue_generated
- Insights:
  - Membership Drive had the highest ROI, generating ₹9,253.77 from only ₹1,189.21 in spend
  - New User Discounts had high spend but relatively lower returns, suggesting scope for optimization or targeting changes

| campaign_id	campaign_name	spend	revenue_generated
243	Referral Program	3,641.18	5,970.43
342	Flash Sale	2,900.46	7,711.31
417	Email Campaign	2,704.17	5,202.30
472	Festival Offer	1,542.65	4,316.40
648	New User Discount	4,892.22	2,261.53
788	New User Discount	2,896.44	4,004.21
1005	Membership Drive	1,189.21	9,253.77

9. KPI visual to display the average order value

- Purpose:
  - To measure the average amount spent per order and compare it against a predefined business goal.
- Visual:
  - KPI card showing Avg_order_value with goal tracking and percentage over target
- Insights:
  - The average order value is ₹6.54K, which significantly exceeds the set goal of ₹1.5K
  - Performance is +336.12% above target, indicating customers tend to place high-value orders
  - This insight supports marketing strategies that promote premium bundles or larger carts

10. Total revenue generated from all campaigns

- Purpose:
  - To display the cumulative revenue generated from all marketing campaigns.
- Visual:
  - Card showing total revenue in millions (32.19M)
- Insights:
  - The total revenue from all campaigns is ₹32.19 million, reflecting strong overall marketing performance
  - Useful as a summary metric to evaluate campaign contributions collectively
  - Can be further analyzed using individual campaign ROIs and time-based trends

Total revenue generated from all campaigns

32.19M
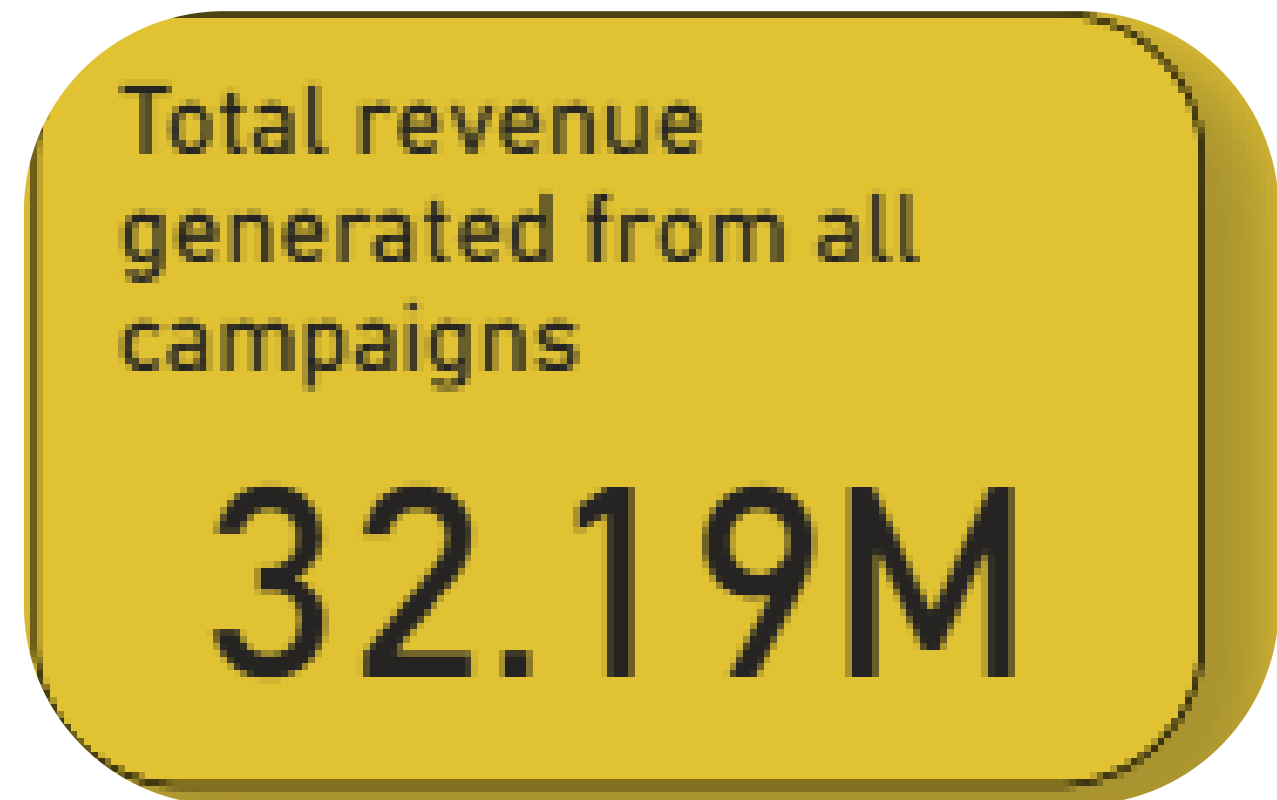
11. Stacked bar chart to compare order quantity per product

- Purpose:
  - To compare how frequently each product was ordered, based on quantity and product count.
- Visual:
  - Stacked horizontal bar chart
  - Y-axis: quantity | X-axis: Count of product_id
- Insights:
  - Most products were ordered in quantities of 2 or 3, indicating medium-sized purchase patterns
  - Products with quantity = 3 had the highest number of orders, suggesting higher customer preference or bundling
  - Helps identify popular order sizes for inventory and packaging optimization
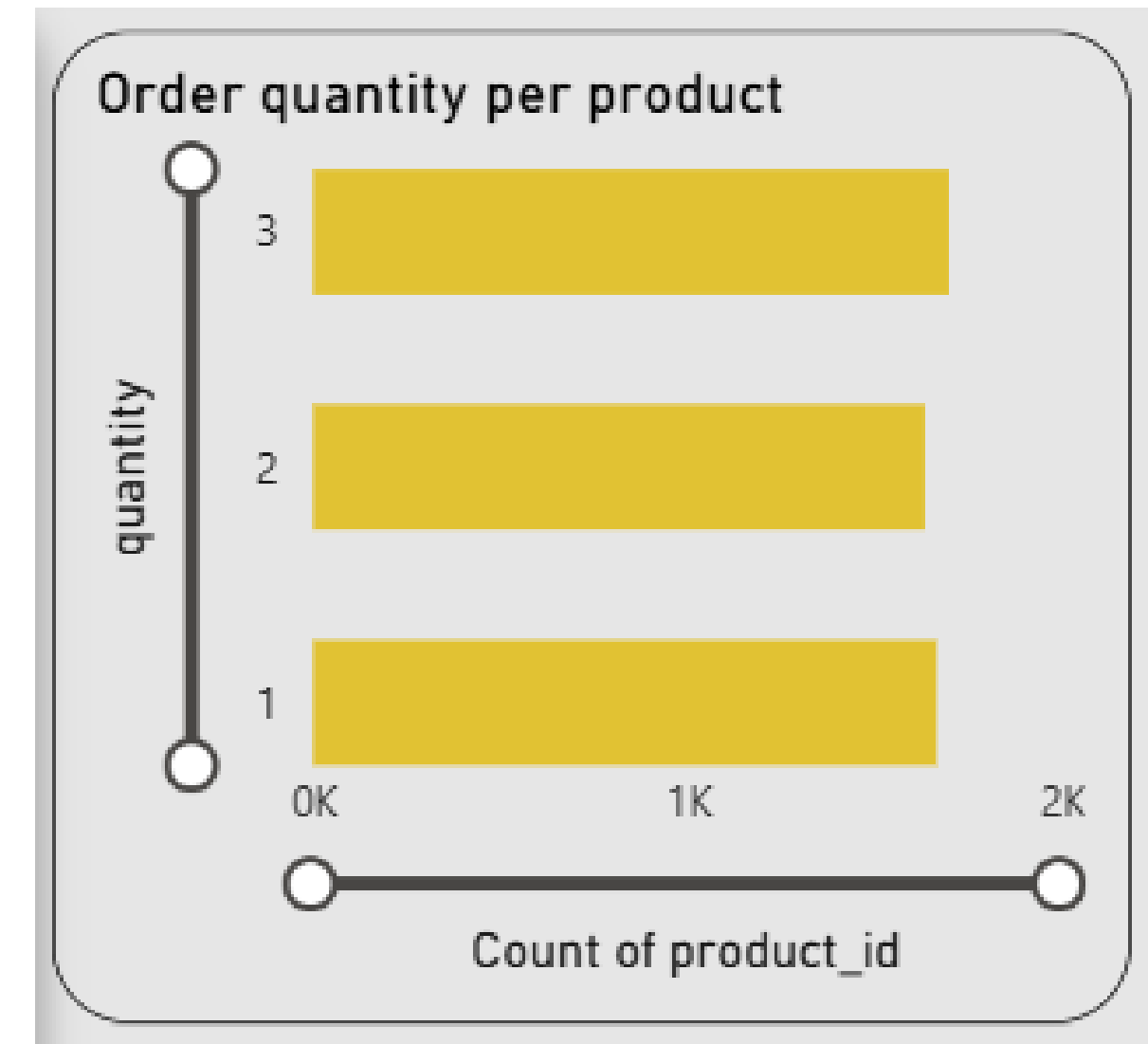
12. Total sales revenue per product

- Purpose:
  - To identify which products generated the highest sales revenue and contributed most to overall income.
- Visual:
  - Horizontal bar chart
  - X-axis: Total_Revenue_Per_Product
  - Y-axis: product_name
- Insights:
  - Vitamins and Pet Treats were the top revenue-generating products, crossing ₹200K
  - Other high performers include Cough Syrup and Toilet Cleaner, indicating strong customer demand
  - Insights help guide product prioritization, promotions, and stock planning
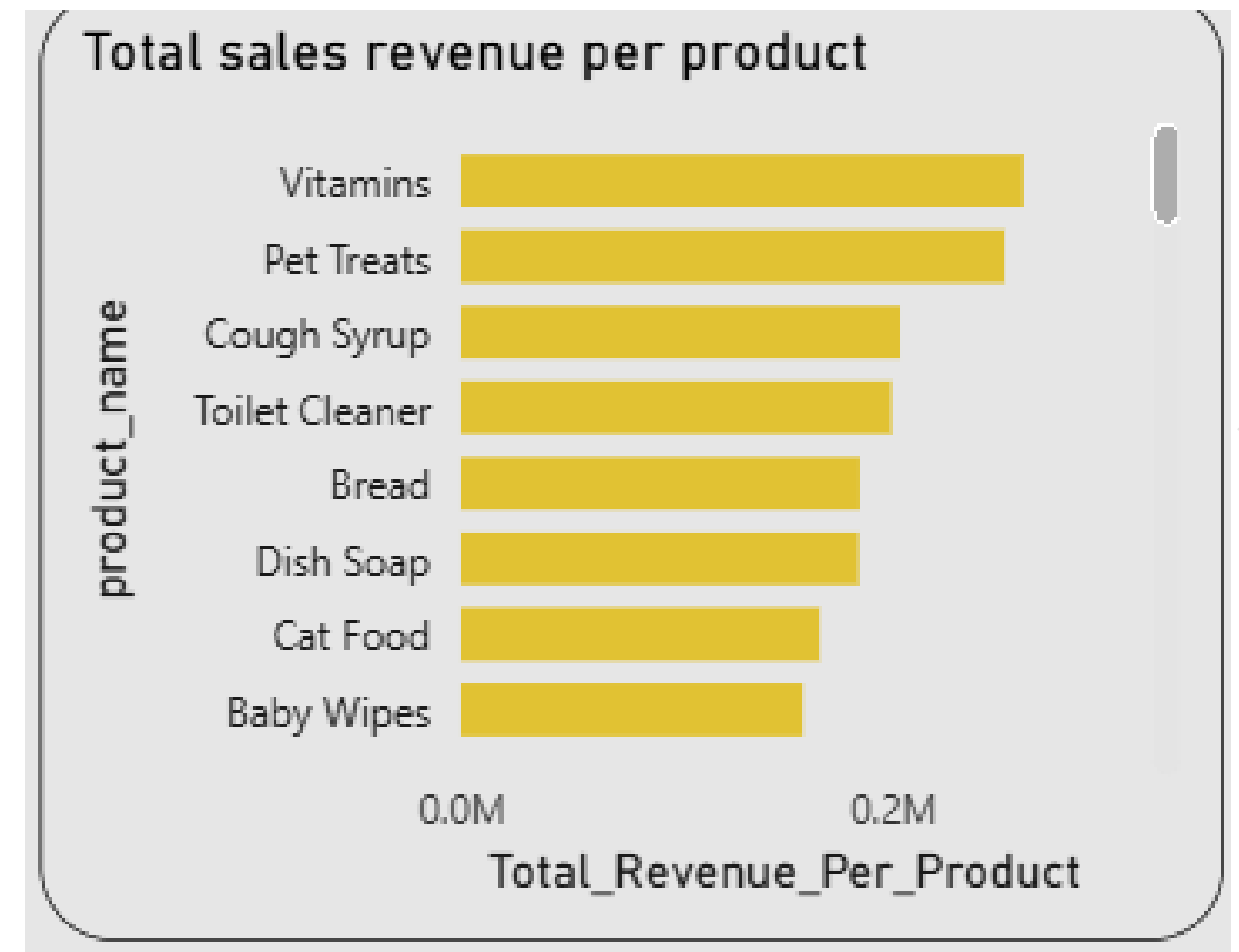
13. Measure to calculate the total delivery time

- Purpose:
  - To calculate and analyze the total delivery time taken from the promised time to the actual delivery, helping assess operational efficiency and identify delivery delays.
- Visual:
- Card Visual
  - Shows the aggregated total delivery time across all orders.
  - Format: dd-mm-yyyy hh:mm:ss (interpreted as a duration rather than an actual date).
- Measure Used:
  - Total_delivery_time = SUMX(blinkit_delivery_performance,blinkit_delivery_performance[actual_time]-blinkit_delivery_performance[promised_time])
- Insights:
  - The aggregated total delivery time is 14 days, 10 hours, 15 minutes.
  - Since this duration is represented as 14-01-1900 10:15:00, it's likely based on Excel datetime logic (where 01-01-1900 is 0), meaning 14 days and 10:15:00 duration.
  - Helps in evaluating total delivery workload or cumulative delay.
- Business Use:
  - Identify areas with frequent delivery delays.
  - Plan for resource allocation.
 - Enhance customer satisfaction by reducing delivery lag.

Total_delivery_time

14-01-1900
10:15:00

14. Visualization of customer segments

- Purpose:
  - To understand the distribution of customers across different segments — Regular, Premium, New, and Inactive — for better marketing and engagement strategies.
- Visual:
  - Pie Chart
  - Legend: customer_segment
  - Values: Count of customers in each segment
  - Data labels: Number of customers and their percentage share
- Insights:
  - The customer base is evenly distributed across all segments, with no group exceeding 26%:
  - Regular: 639 customers (25.56%)
  - Premium: 633 customers (25.32%)
  - New: 628 customers (25.12%)
  - Inactive: 600 customers (24%)
  - The New and Premium segments together account for over 50%, indicating strong growth and high-value customers.

Visualization of customer segments
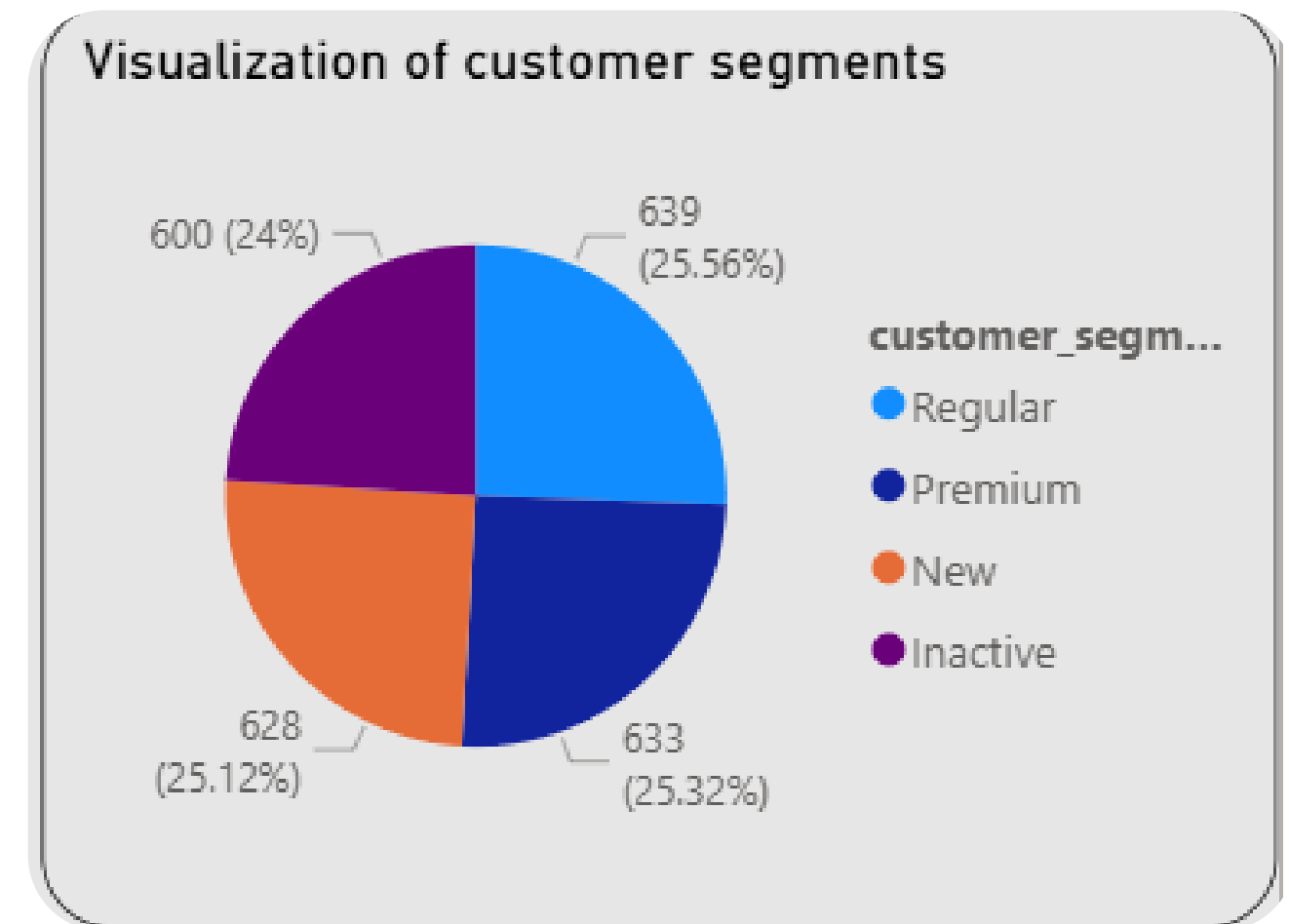
15. Frequency of orders per pincode

- Purpose:
  - To analyze the distribution of customer orders by geographic location (Pincode-wise) and identify areas with high or low order volumes for targeted operations and marketing.
- Visual:
  - Table Visual
  - Columns:
  - Pincode_New – Region-wise customer location
  - Sum of total_orders – Total number of orders placed from each pincode
  - Grand Total: 26,229 orders across all listed regions
- Insights:
  - Orders are spread across multiple pincodes, with most regions contributing modestly (3–20 orders each).
  - Pincode 009477 shows the highest number of orders among the listed, indicating potential customer density or high engagement.
  - Several regions like 000701 and 003231 show minimal activity (only 3 orders), suggesting low market penetration or delivery constraints in those areas.
  - This data can serve as a base for heatmap visualizations to clearly highlight high-demand zones.
- Business Use:
  - Identify high-performing zones for delivery optimization and priority servicing.
  - Detect low-performing areas that may need promotional strategies, logistics improvements, or customer feedback surveys.
  - Assist in location-based decision-making for warehouse placement, last-mile delivery strategy, or hyperlocal marketing.

| Pincode_New	Sum of total_orders
000074, India	8
000516, India	6
000605, India	19
000701, India	3
001690, India	6
002279, India	12
003231, India	3
003256, India	5
005102, India	14
005549, India	10
005657, India	11
006107, India	17
007488, India	14
007809, India	12
008125, India	17
008569, India	7
009477, India	20
Total	**26229**

16. Measure to calculate the Return on Ad Spend (ROAS) using (revenue_generated / spend)

- Purpose:
  - To measure the effectiveness of advertising spend by calculating how much revenue is generated for every ₹1 spent on ads.
- Visual:
  - Card Visual
  - Metric Displayed: ROAS (Return on Ad Spend)
  - Value: 1.97
- Insights:
  - A ROAS of 1.97 means that the business earns ₹1.97 in revenue for every ₹1 spent on advertising.
  - This indicates a moderately efficient marketing strategy. While it's above break-even (ROAS > 1), there is still scope for optimization to further improve profitability.
  - A good benchmark for ROAS depends on the industry — eCommerce often targets 4:1 or higher.
- Business Use:
  - Helps assess ad campaign performance and make data-driven decisions on budget allocation.
  - Aids in comparing ROAS across different channels (e.g., Google Ads, Instagram, Facebook).
  - Can guide future marketing strategies, including scaling successful campaigns or discontinuing underperforming ones.
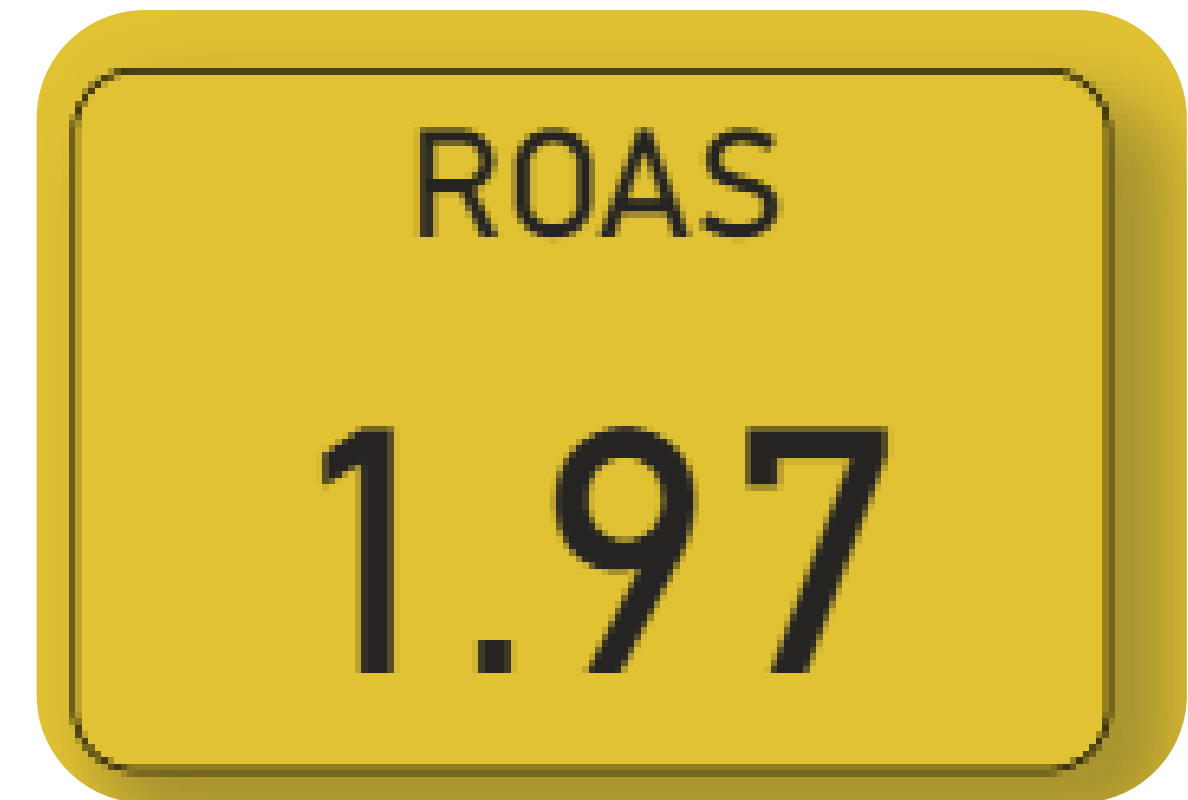
17. Scatter plot ofAverage of delivery_time_minutes by distance_km

- Purpose:
  - To explore the relationship between delivery distance (distance_km) and the average delivery time (delivery_time_minutes) to assess delivery efficiency across varying distances.
- Visual:
  - Scatter Plot
  - X-axis: distance_km
  - Y-axis: Average of delivery_time_minutes
  - Data points: Each represents an individual order or group of orders
- Insights:
  - Most deliveries are made within a 0–5 km range, with average delivery times typically between 5 and 15 minutes.
  - The delivery time shows mild variation but doesn't increase significantly with distance — suggesting that delivery operations are efficient even as distance increases.
  - A few outlier points above 20 minutes may indicate delays due to traffic, delivery partner unavailability, or operational issues.
- Business Use:
  - Useful to validate and improve delivery logistics by analyzing where time increases disproportionately with distance.
  - Helps identify whether certain areas (or ranges) consistently result in delays, guiding route optimization or resource redistribution.
  - Can be enhanced with filters by delivery partner, order volume, or time of day to pinpoint bottlenecks more accurately.

Average of delivery_time_minutes by distance_km
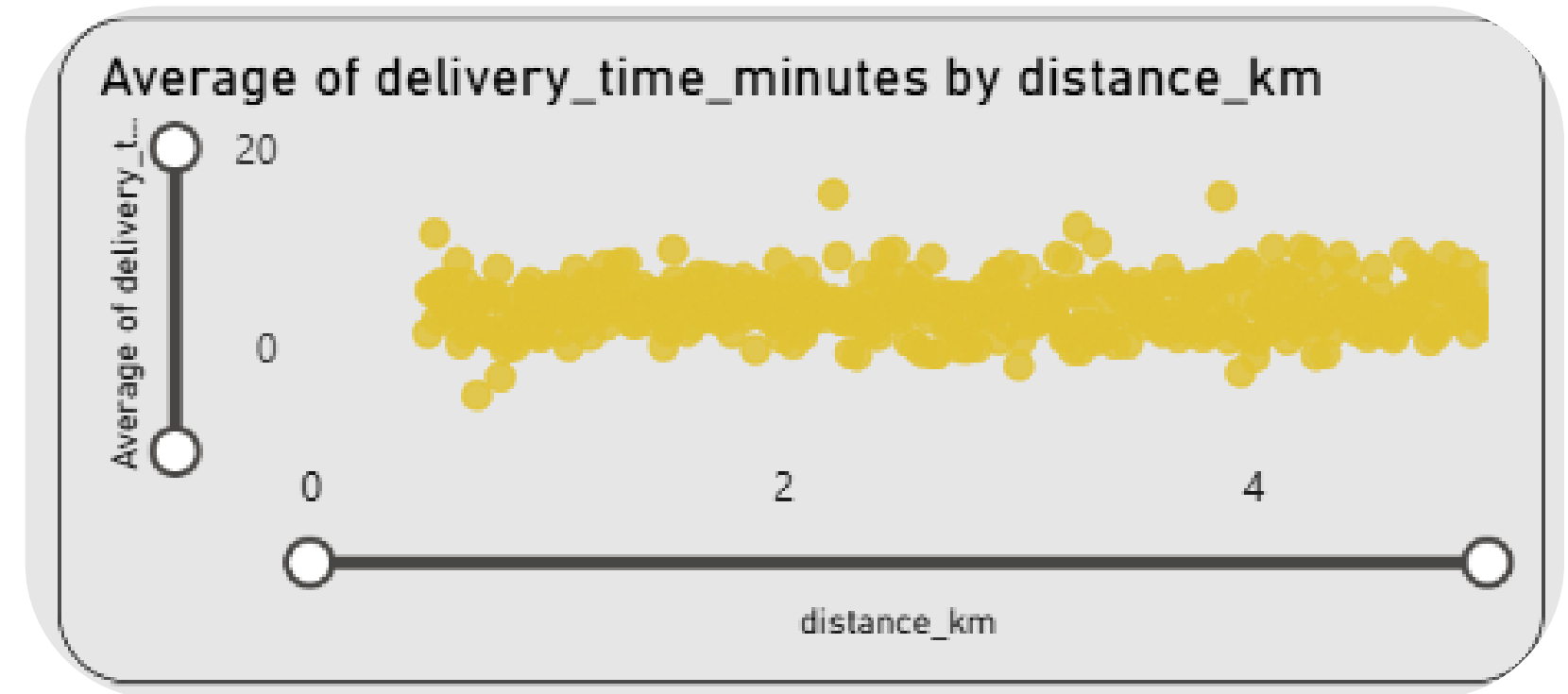
18. Measure to calculate customer retention rate using total_orders

- Purpose:
  - To evaluate how effectively the business is retaining its customers over time, which is a key indicator of customer satisfaction, loyalty, and long-term revenue potential.
- Visual:
  - Card Visual
  - Metric Displayed: Customer_Retention_Rate
  - Value: 94.20%
- Insights:
  - A high retention rate of 94.20% indicates that the majority of customers are coming back for repeat purchases.
  - This reflects strong brand trust, satisfaction, and possibly effective loyalty programs or consistent service quality.
  - A small churn (5.8%) still exists and can be analyzed for specific segments or reasons (e.g. delivery delays, pricing, product issues).
- Business Use:
  - Highlights customer loyalty and effectiveness of retention strategies.
  - Guides future investments in customer engagement, support, and personalized marketing.
  - Aids in identifying churn risk and planning targeted win-back campaigns.

Customer_Retention_Rate

94.20%
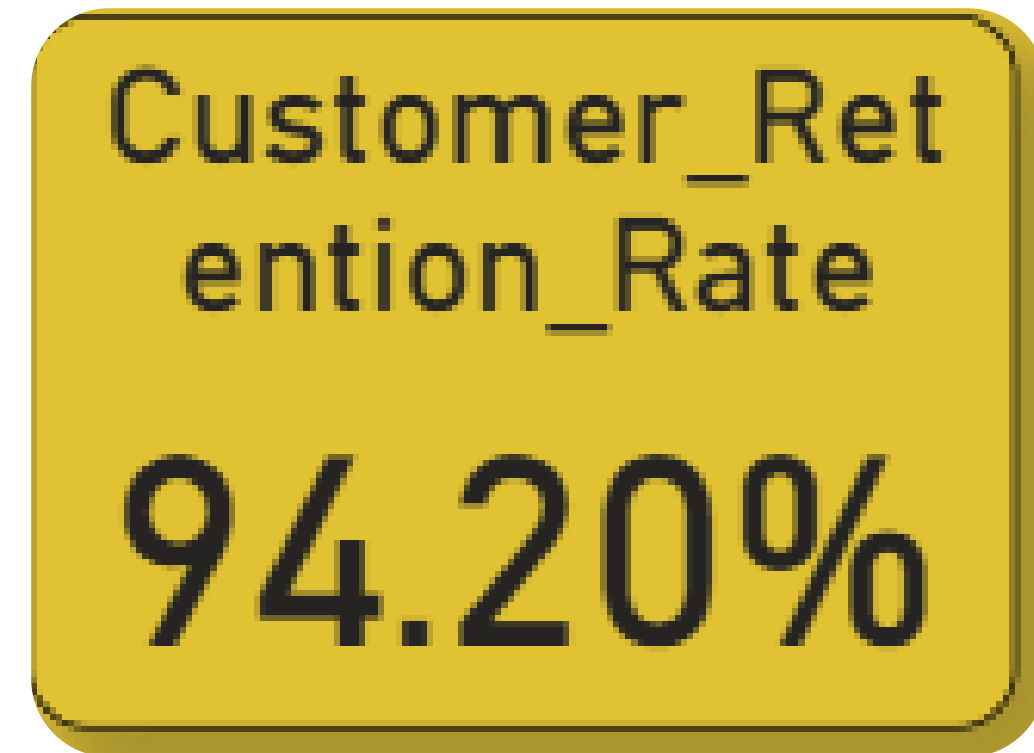
19. Forecast for future stock levels based on historical stock received data

- Purpose:
  - To predict future stock levels using historical data on stock received, enabling strategic planning for procurement, storage, and demand management over the next 30 months.
- Visual:
  - Forecast Line Chart with Area Fill
  - X-axis: Year (includes past years and forecast horizon up to 2026)
  - Y-axis: Sum of stock_received
  - Actuals: Solid line with filled area (e.g., 2023–2024)
  - Forecast: Dashed trendline (e.g., 2025–2026) using Power BI's built-in time series forecasting model
  - Forecast Logic:
  - Based on historical monthly/annual trends of stock_received
  - Applies seasonal smoothing and exponential forecasting (internally by Power BI)
  - Accounts for recent changes in trend to project expected values
- Insights:
  - The forecast shows a gradual decline in stock received:
  - From 14.6K units in 2023 to 14.2K in 2024
  - With a projected decrease through 2026, possibly falling below 13.5K units
  - This indicates a downward trajectory in inventory replenishment if no interventions are made.
  - Highlights the need to evaluate supplier consistency, seasonal demand, or shifts in business priorities.
- Business Use:
  - Allows the inventory team to pre-emptively manage stock shortages or surpluses.
  - Supports automated reorder scheduling, buffer stock planning, and warehouse capacity optimization.
  - Helps finance and operations align budgets and logistics based on forecasted inventory flow.

Forecast of Stock Received (Next 30 Months)
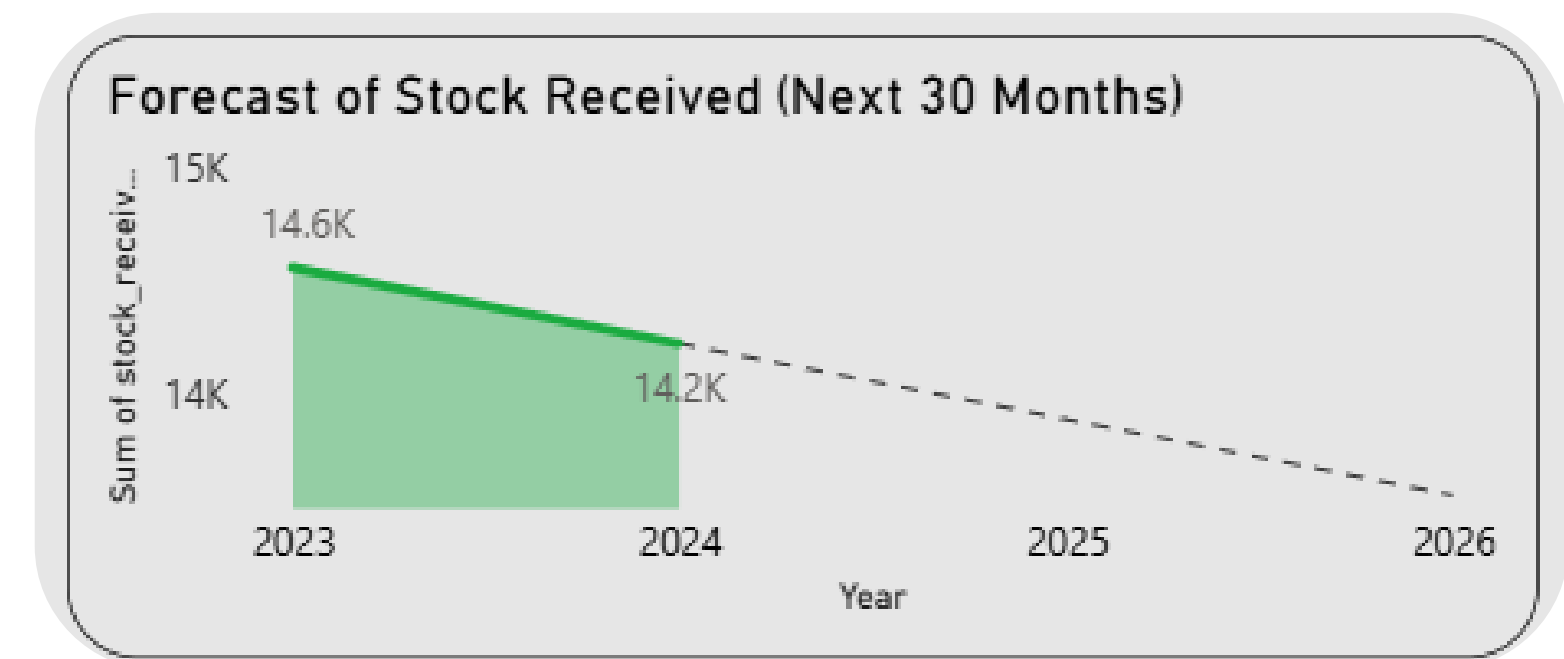
20. Top 5 best-selling products based on quantity ordered

- Purpose:
    - To identify the most popular products by quantity sold, helping guide inventory planning, promotions, and sales strategies.
- Visual:
    - Horizontal Bar Chart
    - Y-axis: product_name
    - X-axis: TotalQuantity
    - Top 5 products based on total quantity ordered
    - Bars are sorted in descending order
- Insights:
    - Pet Treats is the highest-selling product, with nearly 500 units sold.
    - Other strong performers include:
    - Toilet Cleaner
    - Dish Soap
    - Vitamins
    - Cough Syrup
    - All five products have comparable sales volumes, indicating consistent customer demand across essential and wellness categories.
    - This trend highlights pet care, hygiene, and health products as key revenue contributors.

Top 5 best-selling products based on quantity ordered
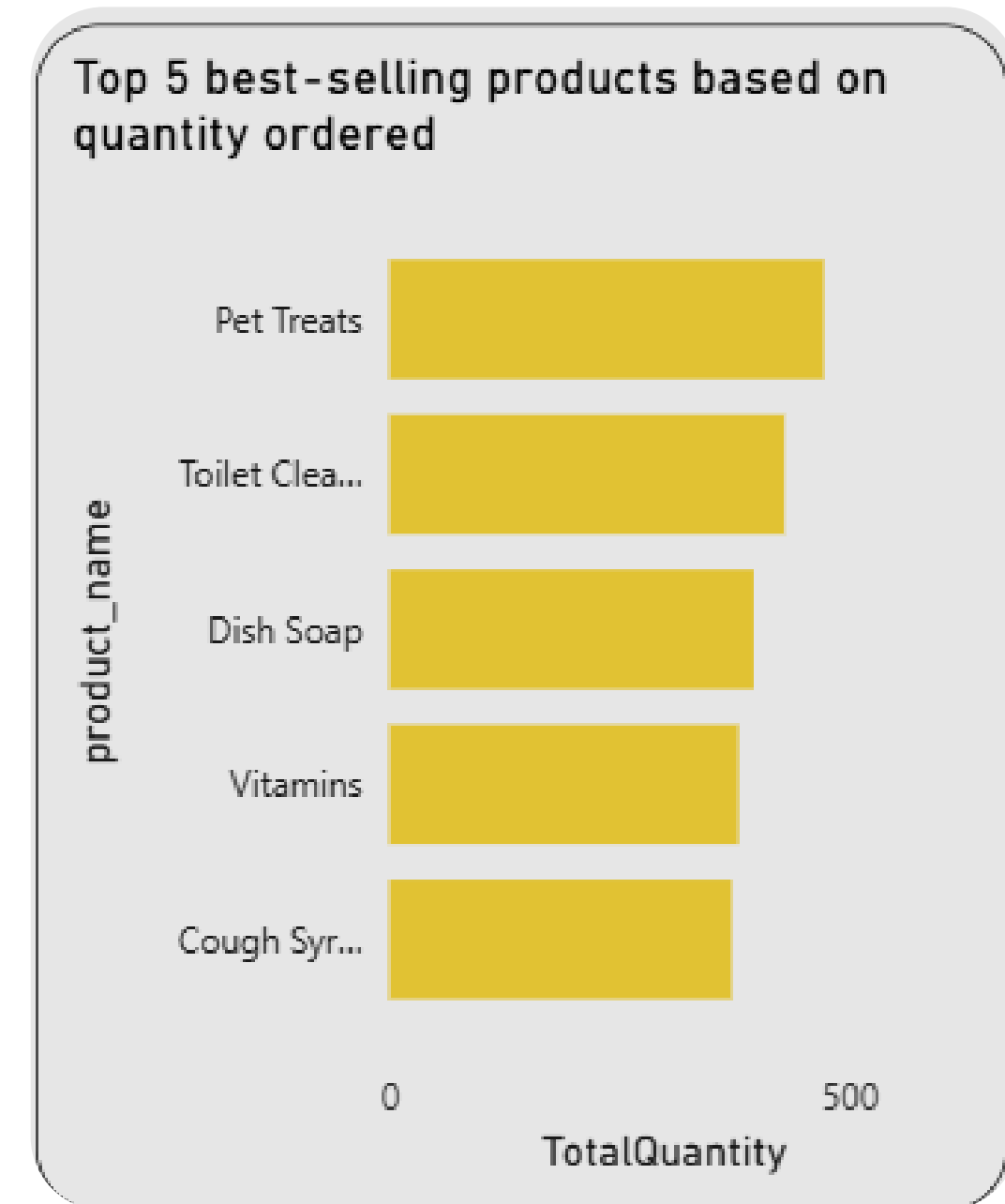
21. Time-series analysis of daily order counts using order_date

- Purpose:
  - To analyze order trends over time and observe how daily order volume has changed between years, aiding in forecasting and performance evaluation.
- Visual:
  - Line Chart (Time Series)
  - X-axis: Year (extracted from order_date)
  - Y-axis: Order_count (aggregated)
  - Data plotted yearly to show change in total daily order counts between 2023 and 2024
  - Yellow line indicates trend
- Insights:
  - Order volume increased from around 2,460 in 2023 to nearly 2,570 in 2024.
  - The upward trend suggests improved customer acquisition, order frequency, or market expansion.
  - Although the increase is moderate, it indicates healthy growth and operational stability over the past year.
- Business Use:
  - Helps in forecasting future demand and preparing supply chain accordingly.
  - Useful for identifying seasonal patterns, promotional impact, or growth periods.
  - Can guide resource planning, such as delivery staff allocation and inventory readiness.

Time-series analysis of daily order counts using order_date
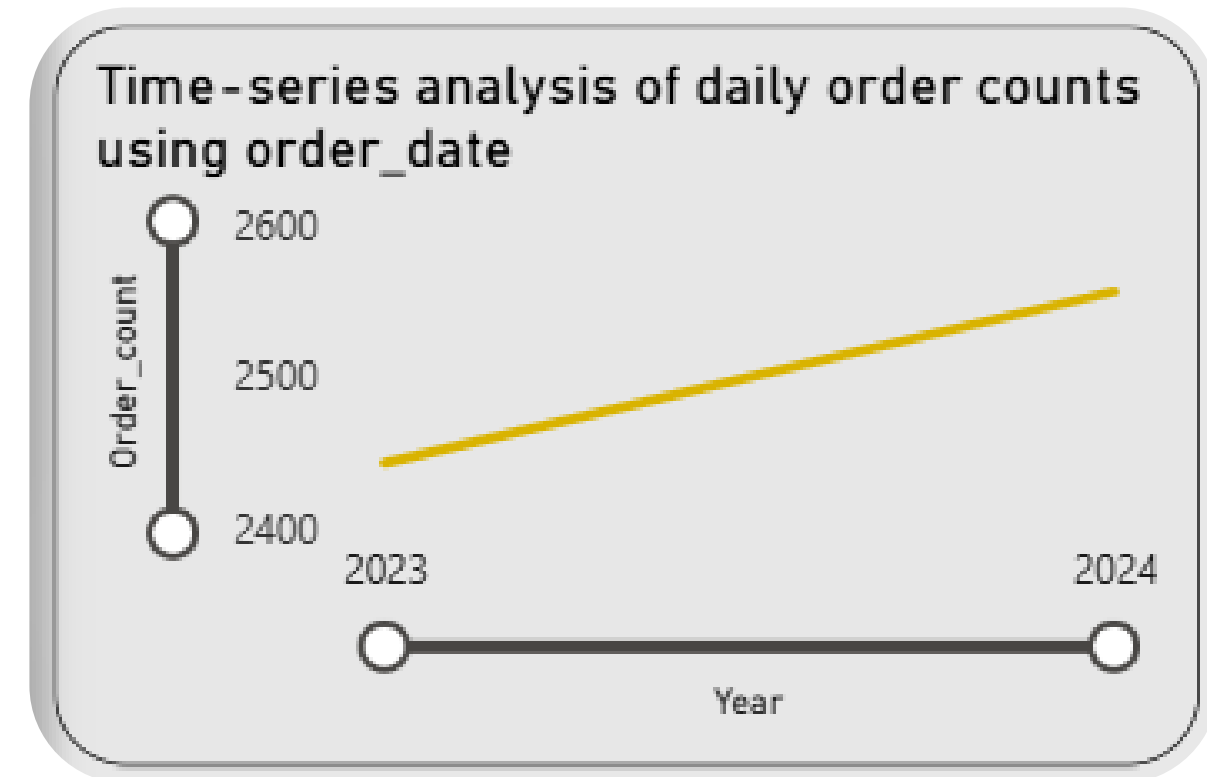
22. Most frequently ordered product using DAX

- Purpose:
  - To identify which products are ordered most frequently, helping understand customer preferences and purchasing patterns.
- Visual:
  - Chart Type: Horizontal bar chart
  - X-axis: Order_Count (number of times ordered)
  - Y-axis: product_name (ordered by frequency)
- Insights:
  - Pet Treats were the most frequently ordered product, indicating strong recurring demand from pet owners
  - Other high-frequency products include Vitamins and Cough Syrup, suggesting these are regular household needs
  - The data reveals customer buying habits that can inform inventory management and subscription program opportunities

Most frequently ordered product

Pet Treats

23. Matrix visualization to show order quantity per product category

- Purpose:
  - To analyze order quantities across different product categories, identifying high-volume segments and inventory requirements.
- Visual:
  - Chart Type: Matrix visualization
  - Rows: Main product categories (Baby Care, Cold Drinks & Juices, Dairy & Breakfast, Fruits & Vegetables)
  - Columns: Sub-categories/products (Baby Food, Baby Wipes, etc.)
  - Values: Order quantity (count or sum)
- Insights:
  - Baby Care category shows strong demand, with 176 units of Baby Wipes ordered (highest volume)
  - Baby Food follows with 115 units ordered
  - Other categories (Cold Drinks, Dairy, Fruits/Vegetables) show no recorded orders in this dataset, which may indicate:
  - Data collection issues
  - New product categories without sales yet
  - Seasonal demand patterns

Matrix visualization to show order quantity per product category

category	Baby Food	Baby Wipes	B
Baby Care	115	176	
Cold Drinks & Juices			
Dairy & Breakfast			
Fruits & Vegetables			
Total	**115**	**176**	

# 24. Customer lifetime value using avg_order_value and total_orders

- Purpose:
  - To measure and analyze the total revenue generated by customers over their entire relationship with the business, helping identify high-value customer segments and guide retention strategies.
- Visual:
  - Card Visualization (for total CLV)
  - Supporting Visuals (Optional):
  - Line Chart: CLV trend over time
  - Bar Chart: CLV by customer segment (e.g., new vs. returning, geographic region)
  - Table: Top 10 high-CLV customers
- Insights:
  - Total CLV: ₹29.03M – Indicates strong long-term revenue generation from customers.
  - High CLV suggests:
  - Effective customer retention strategies
  - Loyal customer base with repeat purchases
  - Potential for upselling/cross-selling opportunities
  - Areas to investigate:
  - Which customer segments contribute most to CLV?
  - How does CLV vary by acquisition channel?
  - Are there declining trends in newer cohorts?

customer lifetime value

29.03M

25. Dashboard to track order trends across different categories

- Purpose:
  - To analyze both volume and value of orders across different product categories over time, providing a dual insight into sales activity and customer spending behavior.
- Visual:
  - Line Chart
  - X-axis: Hierarchy of Month → category
  - Primary Y-axis: TotalQuantity
  - Secondary Y-axis: Average Order Value
  - Tracks both how much is sold and how much is earned per order, category-wise per month.
- Insights:
  - Reveals which categories are most popular (high total quantity) and which are most profitable per transaction (high average order value).
  - For example, a category with moderate quantity but high average order value (e.g., Personal Care) may be a premium product line.
  - Helps compare sales trends over time and uncover seasonal shifts or price sensitivity in different categories.

Dashboard to track order trends across different categories

● TotalQuantity ● Average Order Value
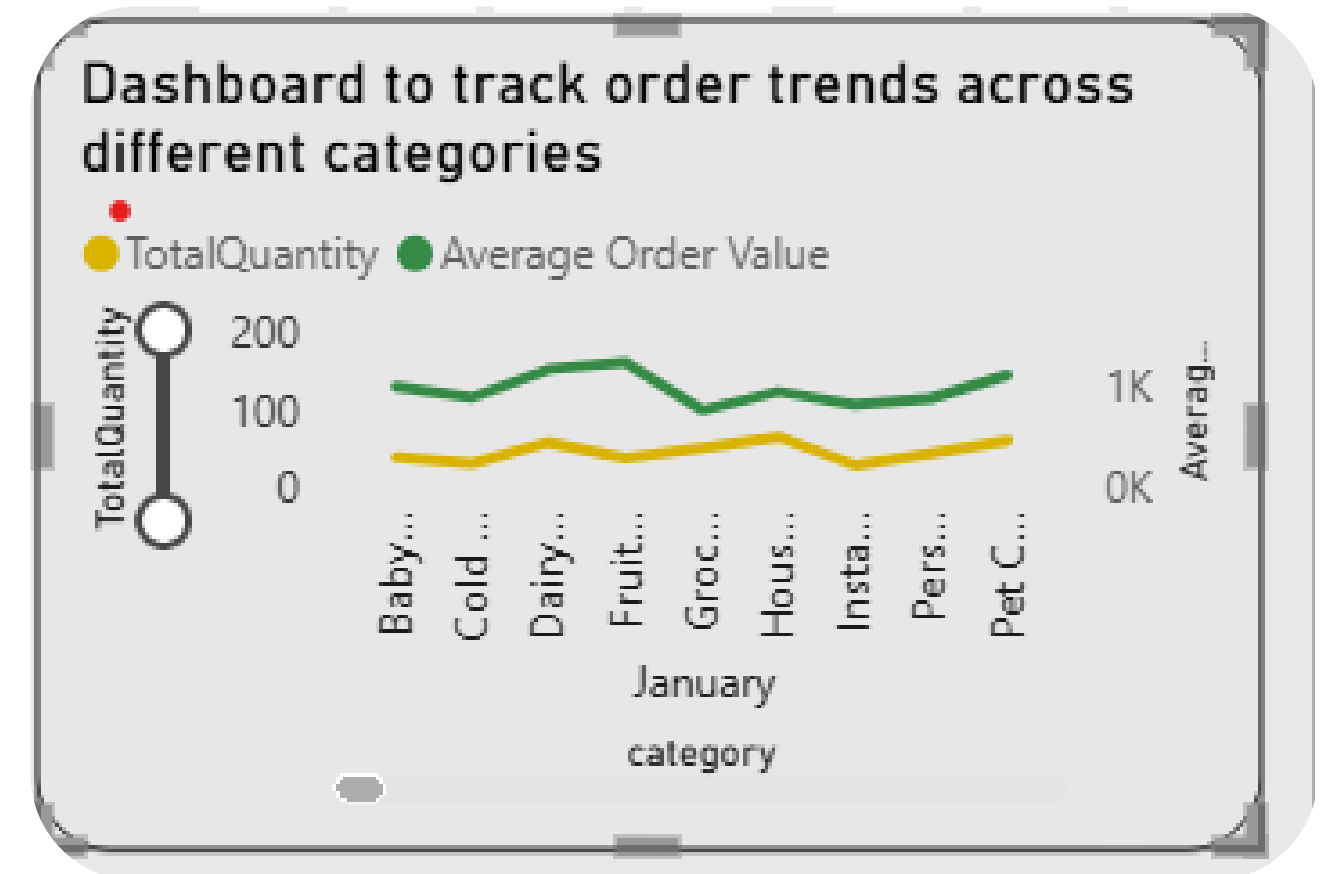
26. Funnel chart to track the campaign conversion process

- Purpose:
  - To analyze the conversion efficiency of a marketing campaign by visualizing the drop-off from impressions to clicks and finally to conversions.
- Visual:
  - Funnel Chart
  - Stages Tracked:
    - Impressions (Top – 100%)
    - Clicks (Middle – small drop)
    - Conversions (Bottom – sharp drop to 1%)
  - Each stage shows how many users continued to the next step
- Insights:
  - The funnel starts with a large number of impressions, indicating good campaign reach.
  - A significant drop from impressions to clicks suggests low engagement or poor ad targeting/message.
  - An even sharper decline from clicks to conversions, down to just 1%, indicates:
  - Landing page issues, poor user experience
  - Lack of compelling call-to-action or misalignment between ad and offer
  - Possible pricing or trust-related concerns from users

Funnel chart to track the campaign conversion process

100%

impressions

Clicks

conversions

1%
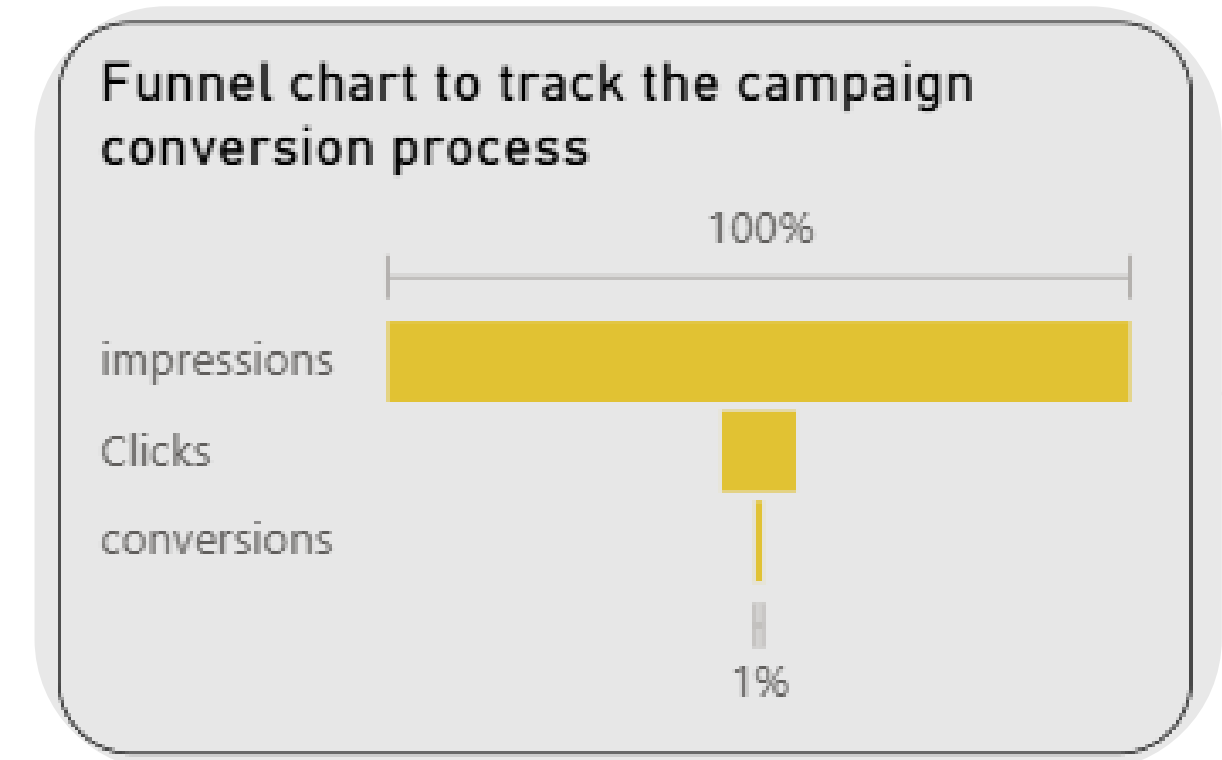
27. Dynamic slicer for filtering orders by delivery status (on-time vs delayed)

- Purpose:
  - To allow users to interactively filter reports and dashboards based on the delivery performance of orders, helping analyze trends and performance across On-Time, Slightly Delayed, and Significantly Delayed deliveries.
- Visual:
  - Slicer (Dropdown / Checkbox Style)
  - Field: delivery_status
  - Values Available:
    - ✅ On Time
    - ✅ Slightly Delayed
    - ✅ Significantly Delayed
  - Users can select one or more statuses to dynamically update visualizations across the report (e.g., bar charts, maps, tables)
- Insights :
  - Analyze delivery efficiency and how it affects customer satisfaction, retention, or reviews.
  - Easily compare key metrics (e.g., delivery time, order value, customer segment) across different delay categories.
  - Helps identify patterns or causes of delays when used in combination with visuals filtered by this slicer.

delivery_status ∨

☐ On Time

☐ Significantly Delayed

☐ Slightly Delayed
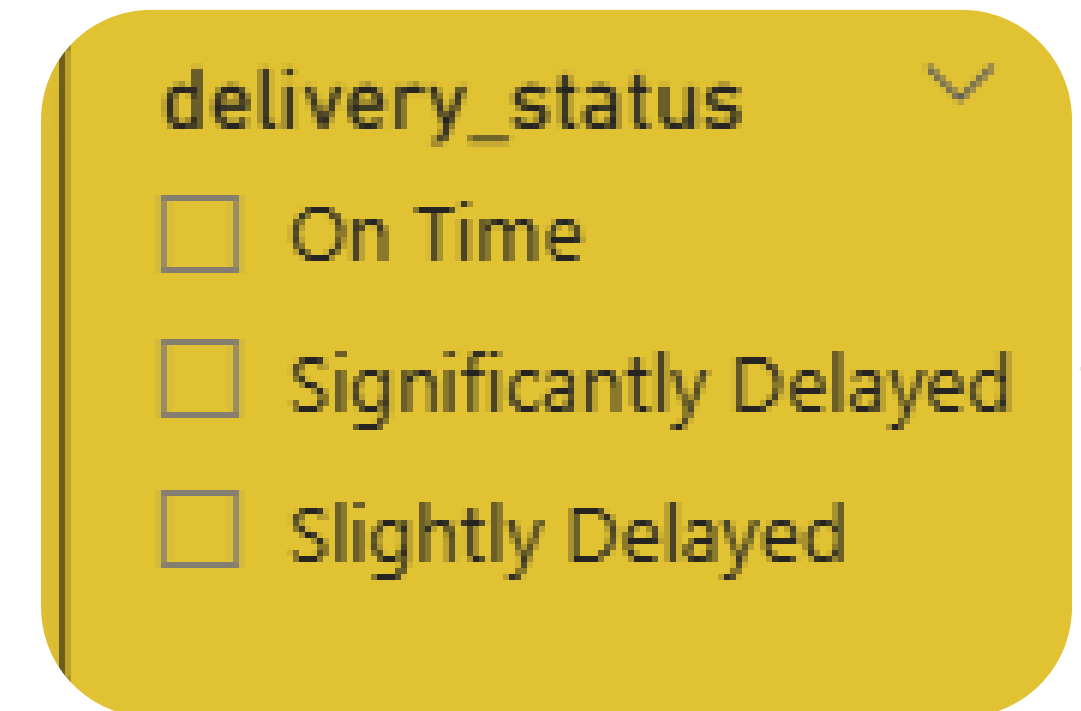
28. Map visualization showing order density per area

- Purpose:
  - To visually represent the geographic distribution of orders across different regions in India, helping identify high-demand zones and optimize delivery operations.
- Visual:
  - Map Visualization (Scatter/Location Map)
  - Dots represent order density by area (likely based on Pincode, City, or Coordinates)
  - Larger or denser clusters indicate higher order volumes
  - Coverage primarily focused on India, with visible data points across metro and regional areas
- Insights:
  - High concentration of orders is seen in urban centers like Delhi, Mumbai, Bengaluru, and Hyderabad.
  - Some sparser regions still show activity, indicating rural or semi-urban demand as well.
  - The distribution helps reveal operational load across areas and informs decisions on expansion, delivery partners, or local warehousing.

Map visualization showing order density per area
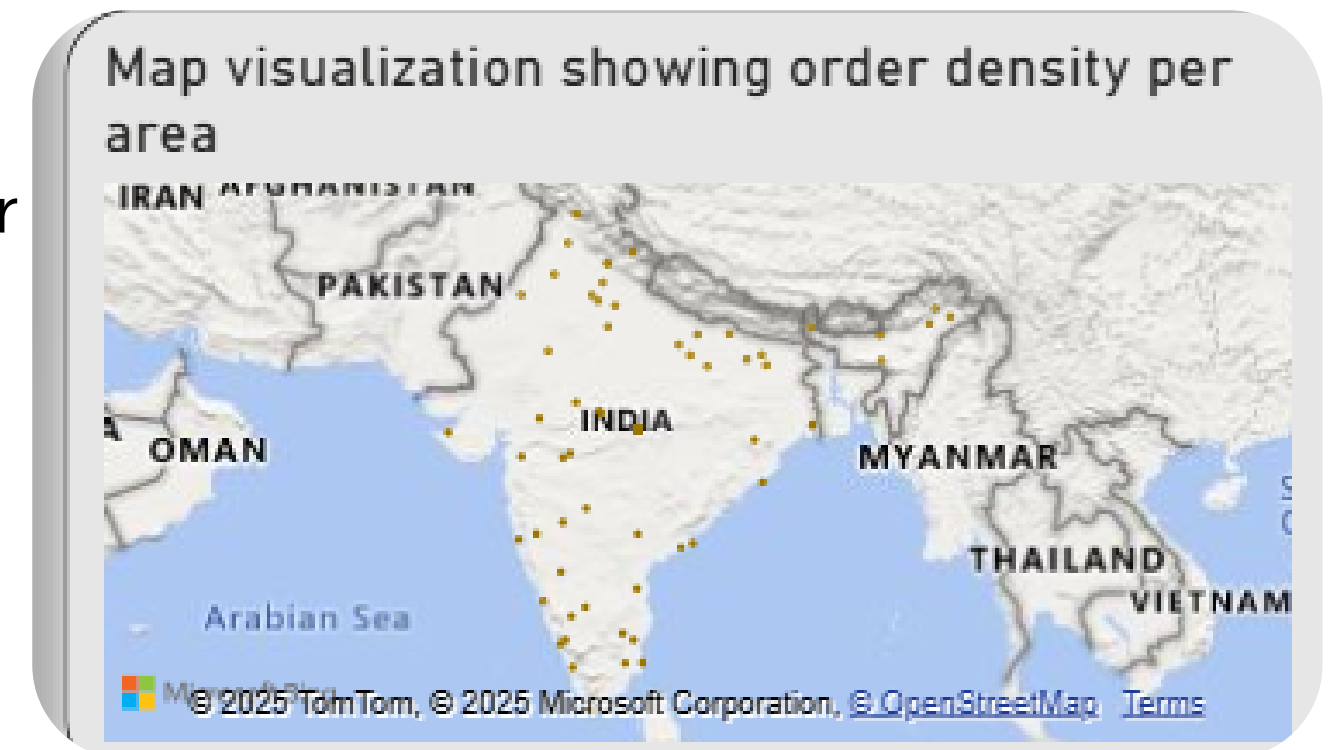
29. Drill-down functionality for order trends by customer segment

- Purpose:
  - To analyze and compare order trends across customer segments over time, using drill-down functionality to uncover deeper insights at yearly, quarterly, or monthly levels.
- Visual:
  - Clustered Column Chart with Drill-Down Enabled
  - X-axis: Year (drillable to Quarter, Month, etc.)
  - Y-axis: Count of orders
  - Legend: customer_segment
    - 🟡 Inactive
    - 🟢 New
    - ⚫ Premium
    - ⚪ Regular
  - Each bar is split by segment, showing how each type of customer contributes to total order volume over time.
- Insights:
  - Across both 2023 and 2024, Regular and Premium customers show higher order activity compared to New or Inactive users.
  - The presence of consistent orders from Inactive customers may indicate occasional returning users.
  - Growth from New customers in 2024 can be tracked using drill-down to months, helping evaluate marketing success and customer acquisition campaigns.
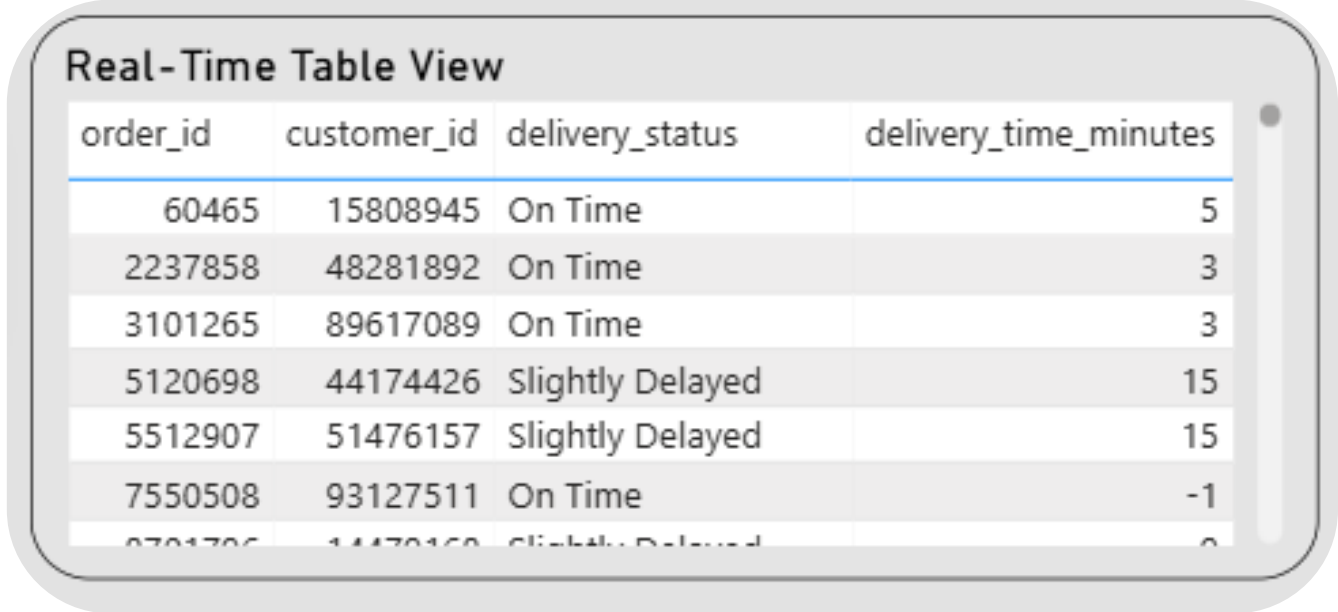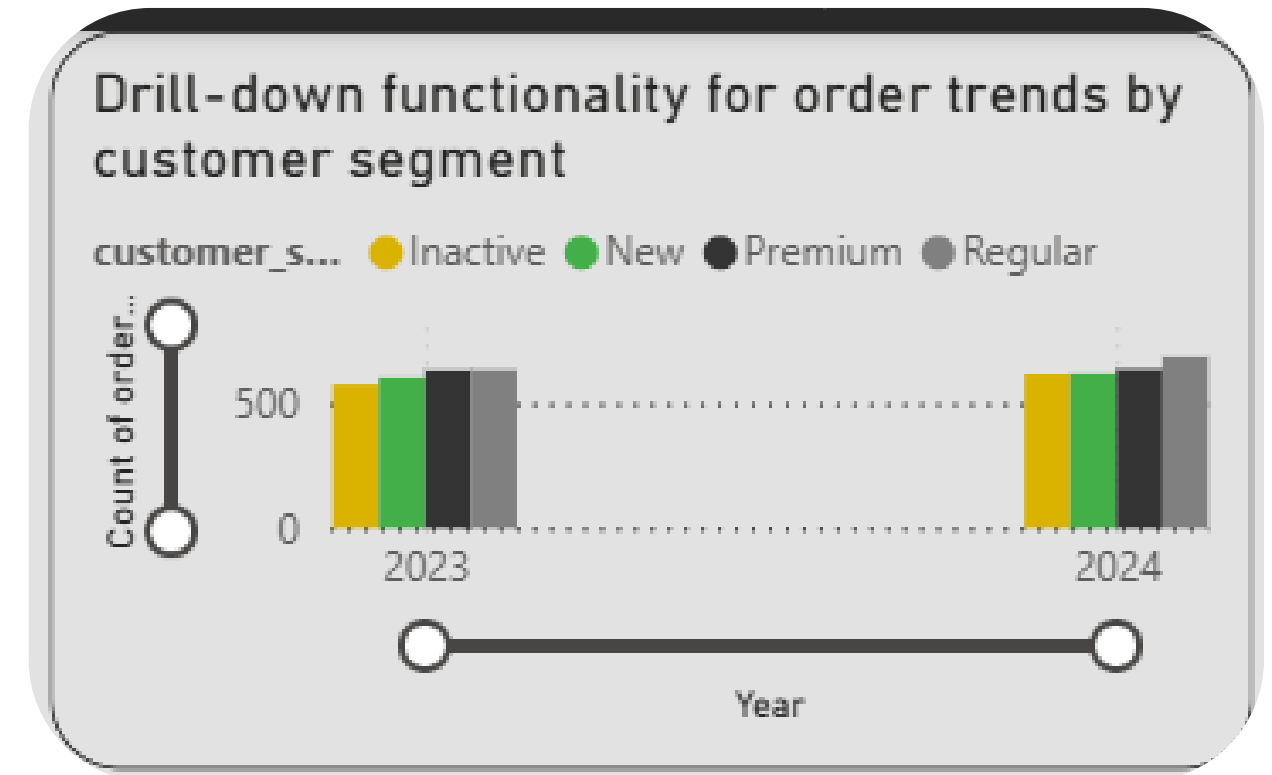
30. Top N campaigns with the highest ROAS

- Purpose:
  - To identify and rank marketing campaigns based on ROAS (Return on Ad Spend), helping evaluate the effectiveness and profitability of different promotional efforts.
- Visual:
  - Horizontal Bar Chart
  - X-axis: TopCampaignsROAS
  - Y-axis: campaign_name
  - Bars represent ROAS values for each campaign
  - Focuses on the Top N performing campaigns based on ROAS

Top N campaigns with the highest ROAS

- Insights:
  - All top campaigns shown — including Flash Sale, Festival Offer, and App Push Notification — have ROAS values nearing 1, indicating almost 1x return per currency spent.
  - Flash Sale and Festival Offer stand out, suggesting time-sensitive or seasonal campaigns perform well.
  - Email Campaigns and Push Notifications also deliver strong performance, making them cost-efficient channels for driving revenue.
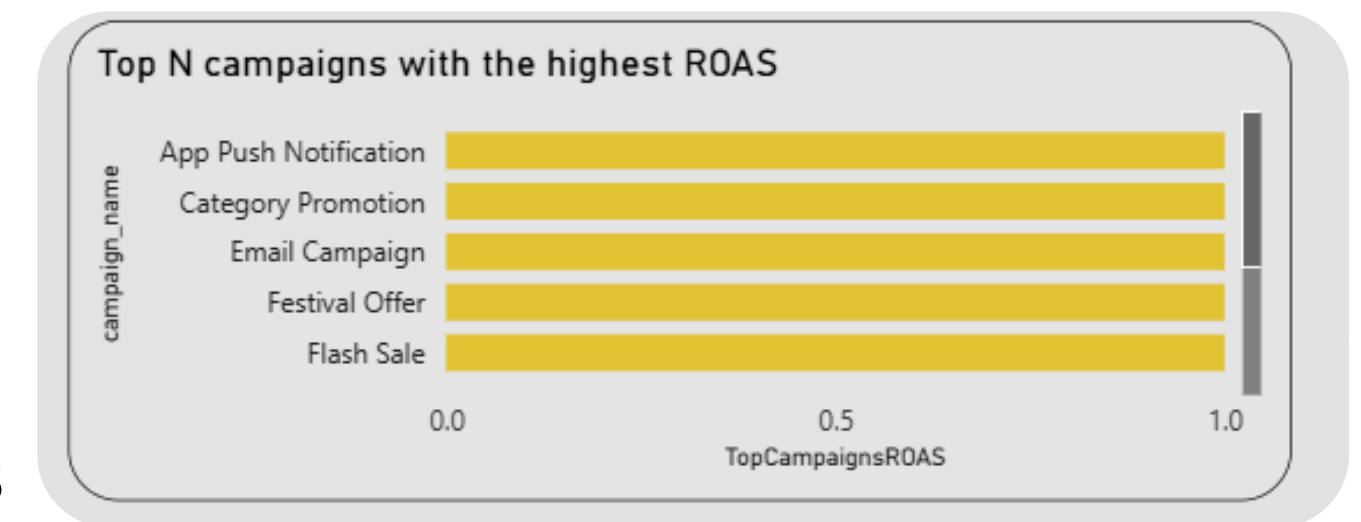
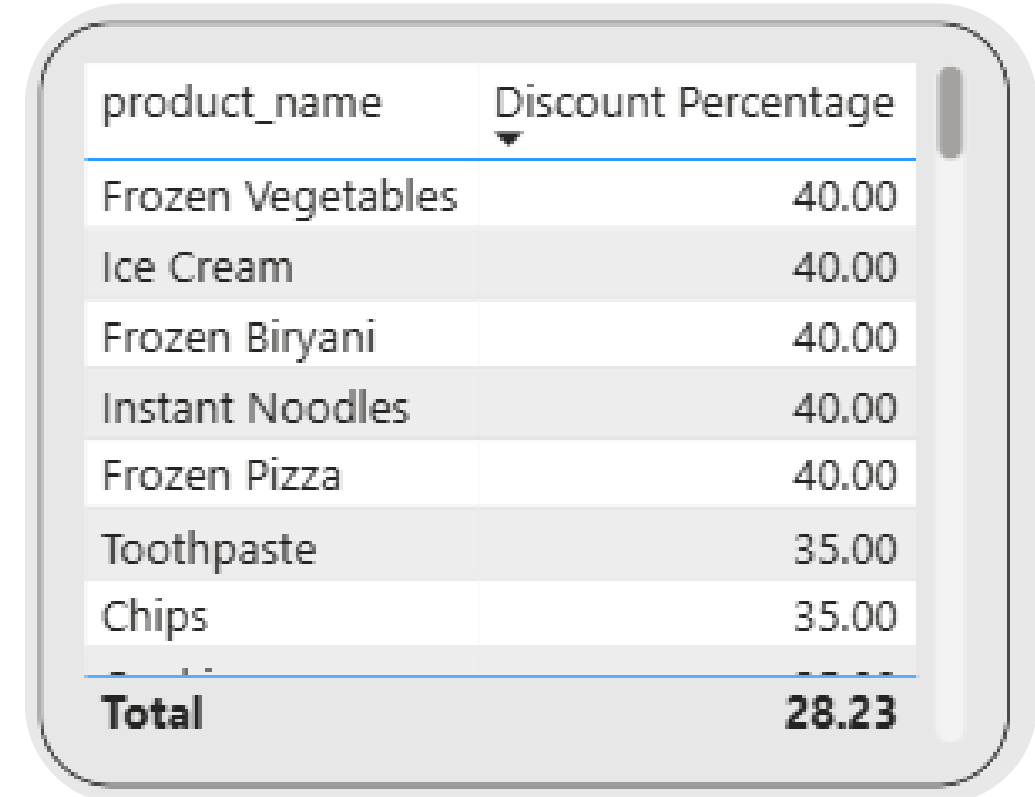# 31. Real-time Table to monitor ongoing deliveries and status updates

- Purpose:
  - To provide a live snapshot of individual order records, enabling quick monitoring and analysis of delivery performance, customer activity, and time efficiency at the order level.
- Visual:
  - Table Visual
  - Columns displayed:
    - order_id
    - customer_id
    - delivery_status
    - delivery_time_minutes
  - Data is presented in a scrollable table format for granular visibility.
- Insights:
  - Most orders are marked "On Time", indicating strong delivery efficiency.
  - A few entries show "Slightly Delayed" with longer delivery durations (~15 minutes).
  - Anomalies (e.g., negative delivery times like -1) highlight potential data entry errors or ETL issues that need validation.
  - This table supports drill-through analysis—users can select a row to investigate further in related visuals (like maps, customer profiles, or delivery delay breakdowns).

Real-Time Table View

order_id	customer_id	delivery_status	delivery_time_minutes
60465	15808945	On Time	5
2237858	48281892	On Time	3
3101265	89617089	On Time	3
5120698	44174426	Slightly Delayed	15
5512907	51476157	Slightly Delayed	15
7550508	93127511	On Time	-1

32. Product-wise discount percentages using price and MRP

- Purpose:
  - To analyze the discount strategy across products, helping identify which items receive higher promotional offers and understand their impact on sales and profitability.
- Visual:
  - Table Visual
  - Columns:
    - product_name
    - Discount Percentage
  - A total row at the bottom shows the average discount across all listed products (28.23%).
- Insights:
  - Several frozen and ready-to-eat items (e.g., Frozen Vegetables, Ice Cream, Biryani, Pizza, Noodles) are consistently offered at 40% discounts, indicating a possible high competition or clearance strategy in this category.
  - Toothpaste and Chips show a lower discount of 35%, possibly reflecting lower margin flexibility or strong brand demand.
  - The overall average discount offered across products is 28.23%, which may be used as a benchmark for future campaigns.

| product_name	Discount Percentage
Frozen Vegetables	40.00
Ice Cream	40.00
Frozen Biryani	40.00
Instant Noodles	40.00
Frozen Pizza	40.00
Toothpaste	35.00
Chips	35.00
Total	**28.23**

33. Decomposition tree visualization to analyze product sales

- Purpose:
  - To break down total product orders step-by-step across various dimensions such as category, customer segment, date, and brand.
- Visual:
  - Decomposition Tree
  - Root: Count of order_id
  - Drill-down fields: category, customer_segment, year, month, day, brand
- Insights:
  - Highest number of orders observed in Dairy & Breakfast category (566), followed by Cold Drinks and Fruits
  - Within Inactive customers, most orders occurred in 2023, specifically in March on the 28th
  - Final-level breakdown shows brand-level contributions like Banik-Ratta and Chaudhuri PLC
  - This tool helps perform dynamic drill-downs to identify key contributors to sales at any level (e.g., top day, brand, or segment)
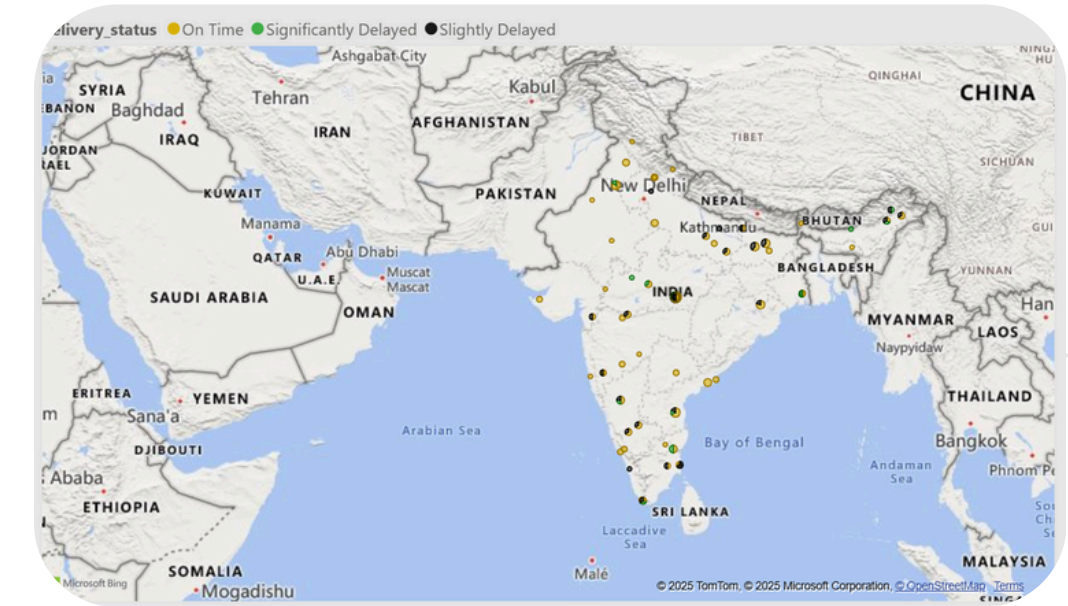
34. Filtering Orders by Date Range

- Purpose:
  - To filter data based on a specific time window, enabling time-based trend and comparison analysis.
- Visual:
  - Date range slicer on order_date
  - Selected range: 16-03-2023 to 04-11-2024
- Insights:
  - Allows you to dynamically adjust the reporting period and observe changes in order volume, revenue, or delivery metrics over time
  - Useful for identifying seasonal trends, campaign effects, and order surges during specific months
  - Enhances report interactivity by providing time-based filtering across all visuals
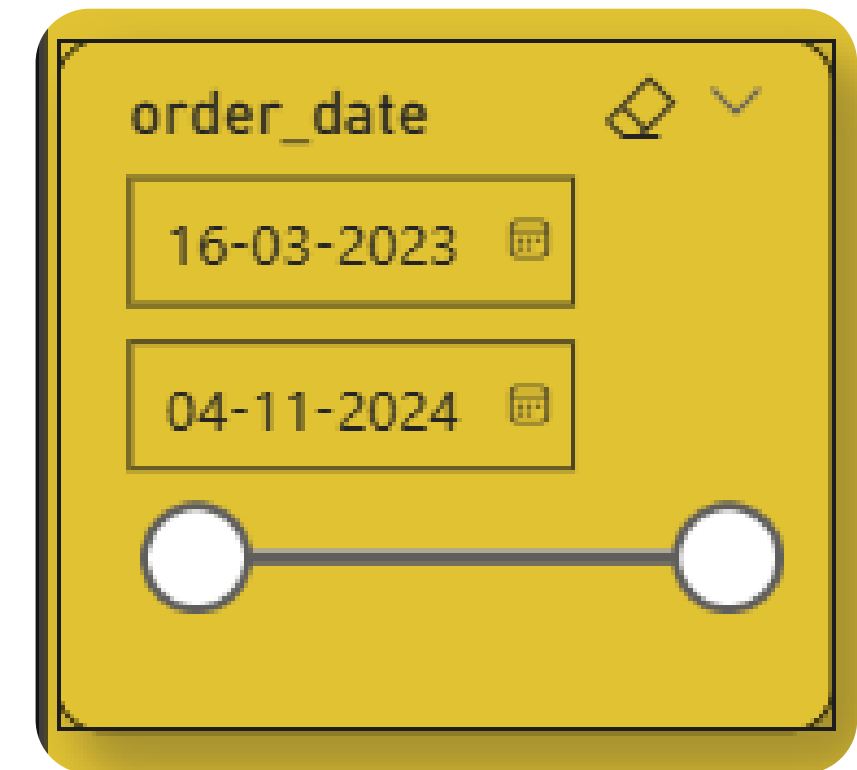
36. Delivery Status by Location

- Purpose:
  - To visualize the geographical distribution of order delivery performance across different regions.
- Visual:
  - Map with color-coded points by delivery_status:
  - 🟡 On Time
  - 🟢 Slightly Delayed
  - ⚫ Significantly Delayed
- Insights:
  - Majority of deliveries across southern and western India are on time
  - Clusters of delays (both slight and significant) observed around eastern India, central regions, and some northeast zones
  - Map helps identify logistics bottlenecks by geography and plan regional improvements in delivery routes or warehouse support

Thank You!