FORMERLY KNOWN AS UETR

SESSION- 2024-2026

COLLEGE OF BUSINESS STUDIES

POWER BI PROJECT REPORT

ON

AMAZON SALES DATA

Submitted To:
Dr. Sourabh Poswal

Submitted By:
RADHIKA DHAMIJA
MBA-(A) CU24140038
HARSH KASHYAP 1
MBA-(A) CU24140037
ARYAN CHAUHAN
MBA-(A) CU240140090

Acknowledgment

I would like to express my heartfelt gratitude to everyone who contributed to the successful completion of this report on Amazon sales data analysis. Firstly, I extend my sincere appreciation to my faculty guide, **Dr. Sourabh Poswal** , for their invaluable guidance, support, and expertise throughout this project. Their insights and constructive feedback significantly enriched the analysis and enhanced my understanding of the subject matter.

I also wish to thank my peers and colleagues for their encouragement and collaborative spirit, which made this project more enjoyable and productive.

Furthermore, I acknowledge the importance of the data sourced from the online Excel sheet, which provided a robust foundation for this analysis. Lastly, I appreciate the resources and tools, which facilitated the successful execution of this project.

This report reflects the collective efforts of all those involved, and I am grateful for the support and collaboration that made it possible.

Index

Project Report: Power BI Dashboard Analysis

---

1. Introduction

1.1 Overview of Power BI

Power BI is a robust business intelligence platform by Microsoft that allows users to create interactive visualizations and reports. With its seamless integration with data sources like Excel, SQL databases, and cloud services, Power BI simplifies data analysis and reporting. By using its extensive features, users can extract insights, track key performance indicators (KPIs), and present data dynamically.

1.2 Purpose of the Dashboard

The primary purpose of this dashboard is to analyze product sales, pricing, discounts, and customer ratings. Using Power BI, the dashboard helps uncover patterns, trends, and relationships within the dataset. Specifically, the dashboard is built to answer the following questions:

1. How are discount percentages distributed across different price ranges?
2. What is the distribution of customer ratings for products?
3. How do ratings correlate with discount percentages?
4. What is the relationship between actual prices, discounted prices, and ratings?

These insights help businesses optimize their pricing, marketing, and sales strategies to enhance profitability and customer satisfaction.

1.3 Importance of Data Visualization

Visual data representation is crucial for simplifying complex datasets. By leveraging visuals, the dashboard makes data interpretation intuitive and actionable. For example:

- Identifying top-performing products based on ratings.
- Evaluating the impact of discount strategies on customer satisfaction.
- Understanding how pricing influences buying decisions.

This dashboard demonstrates the value of transforming raw data into insights that guide business decisions.

2. Introduction to the Data

The dataset analyzed in this project represents a comprehensive overview of customer purchasing behavior, product pricing strategies, discount patterns, and customer satisfaction ratings of amazon customers. The primary focus of the data is to provide actionable insights into the performance of products across various pricing tiers and ratings, helping businesses make informed decisions regarding pricing, promotions, and customer satisfaction.

Nature of the Data

The data is structured to highlight key business metrics and trends. It includes multiple variables that provide a multidimensional view of the sales process and customer preferences. Key elements of the dataset include:

- **Customer Count:** Information about the number of unique customers interacting with the product catalog, indicating the overall reach of the business.
- **Product Pricing:** Details of the actual product prices and their corresponding discounted prices, offering insights into pricing strategies and their impact on sales.
- **Discount Percentages:** Analyzed to understand the correlation between pricing tiers and the level of discounting applied.

- **Ratings:** Customer ratings for products, showcasing overall satisfaction levels and helping identify areas for improvement.
- **Lead Time**

Purpose of the Analysis

The dataset serves as the foundation for creating a visual and interactive dashboard in Power BI. The project aims to explore:

1. **Pricing and Discount Strategies:** Understanding how discounts influence product sales across various price ranges and customer satisfaction ratings.
2. **Customer Sentiment:** Evaluating the distribution of customer ratings to identify high-performing products and those requiring improvements.
3. **Trends and Patterns:** Identifying trends in pricing, discounts, and ratings to guide business strategies.

Key Metrics and Variables

The following metrics were derived and analyzed within the dataset:

- **Actual Price vs. Discounted Price:** Comparison to evaluate the effectiveness of discounts and their impact on perceived value.
- **Discount Percentage:** A breakdown of discounts applied across different product categories and pricing tiers.
- **Customer Ratings Distribution:** Highlights the level of customer satisfaction with the product catalog.
- **Product Count per Rating:** Segments the number of products within each rating range (e.g., 3.5, 4.0, 4.5), providing a detailed look at customer sentiment.

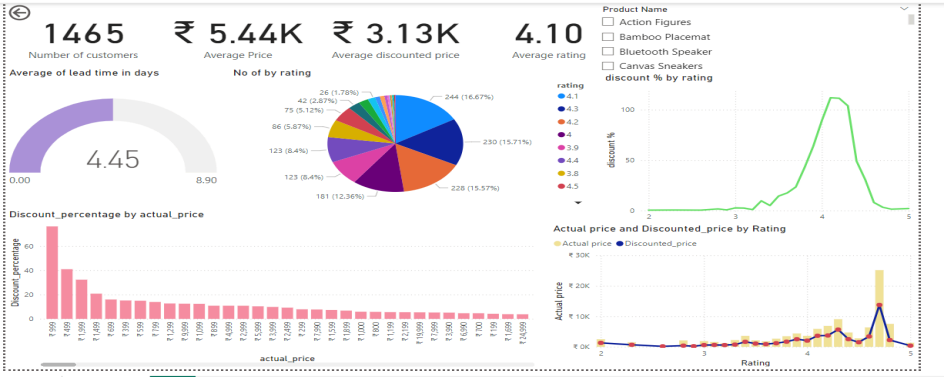
Relevance to the Business Context

The insights generated from the dataset enable businesses to:

- Optimize pricing strategies for different customer segments.
- Develop targeted promotional campaigns based on product performance.
- Improve inventory management by identifying slow-moving and high-demand products.
- Enhance customer satisfaction through data-driven improvements to product offerings.

This project leverages the data to create a user-friendly dashboard, presenting insights in an accessible and actionable format. By utilizing Power BI’s advanced visualization capabilities, the analysis ensures clarity and precision in decision-making, empowering businesses to achieve better operational and strategic outcomes.

3. Overview of the Dashboard



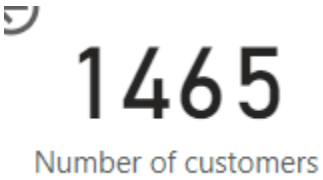
The Power BI dashboard is structured with multiple visuals that collectively provide a 360-degree view of product performance. It comprises:

1. **Key Performance Indicators (KPIs):**
   - Total number of customers: **1,465**
   - Average product price: **₹5.44K**
   - Average discounted price: **₹3.13K**
   - Average rating: **4.10**
2. **Visuals for deeper insights:**
   - Bar chart for discount percentage by actual price.
   - Pie chart showing the distribution of ratings.
   - Line graph analyzing discount percentage by rating.
   - Combined chart comparing actual price and discounted price by ratings.

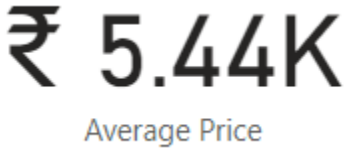
These components are interactive, enabling stakeholders to explore the data dynamically.

Cards Used:

1. **Number of Customers**:
   - **Card Type**: Single Value Card
   - **Value**: **1465**
   - **Description**: Displays the total number of customers, which helps track the number of people who interacted with the products.

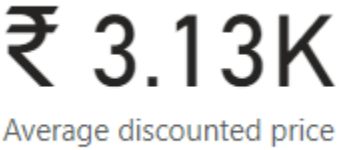


2. **Average Price**:
   - **Card Type**: Single Value Card
   - **Value**: **₹5.44K**
   - **Description**: Shows the average price of all products, giving an overall indication of the product pricing in the dataset.



3. **Average Discounted Price**:

- **Card Type**: Single Value Card
- **Value**: **₹3.13K**
- **Description**: Displays the average price of the products after applying discounts, indicating the pricing trend post-discount.



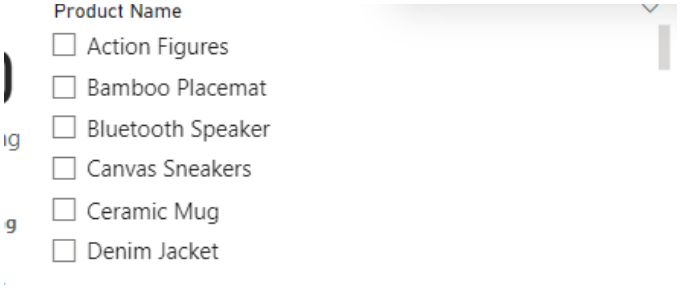
-

4. **Average Rating**:

- **Card Type**: Single Value Card
- **Value**: **4.10**
- **Description**: Shows the average customer rating for the products, helping to understand the general customer satisfaction.

Slicers Used:

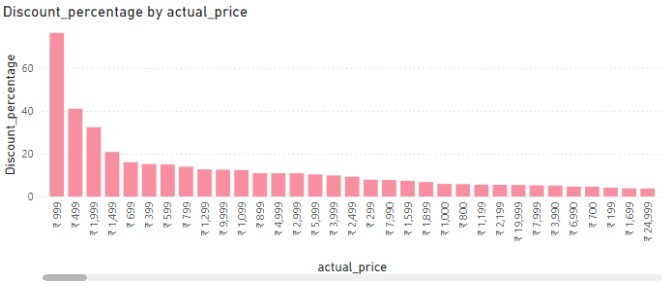
1. **Product Name**:
   - **Type**: List or Dropdown Slicer
   - **Functionality**: Allows users to filter the entire dashboard based on the selected product(s). This slicer provides flexibility to focus on specific products to view related data for that product only.



1. Bar Chart: Discount Percentage by Actual Price

Why this Graph Was Used?

This graph was chosen because a bar chart is effective for comparing numerical values across categories. In this case, it visually compares the **discount percentage** across various **price ranges**, helping stakeholders understand how discounts are applied to different product categories. This is critical for evaluating pricing and promotional strategies.



What Does the Graph Represent?

- **X-Axis (Horizontal):** Displays the actual product prices segmented into ranges (e.g., ₹999, ₹1,499, ₹1,999, etc.).
- **Y-Axis (Vertical):** Represents the discount percentage applied to products within each price range.
- **Bar Heights:** Indicate the magnitude of the discount percentage. Taller bars represent higher discounts.

Insights Gained from the Visualization

1. **Highest Discounts for Low-Priced Products:**

   ○ The graph reveals that products priced at **₹999 or below** receive the largest discounts, often exceeding **50%**.
   ○ This strategy is likely employed to attract budget-conscious customers or to clear inventory of low-priced items.

2. **Discounts Decrease for Higher Price Ranges:**

   ○ Products priced at **₹5,999 and above** show significantly smaller discounts, often below **10%**.
   ○ This suggests that higher-priced products are targeted at customers who are less price-sensitive or view these items as premium offerings.

3. **Price-Based Segmentation:**

○ The consistent pattern of higher discounts for lower-priced items and smaller discounts for higher-priced items reflects a deliberate segmentation strategy.

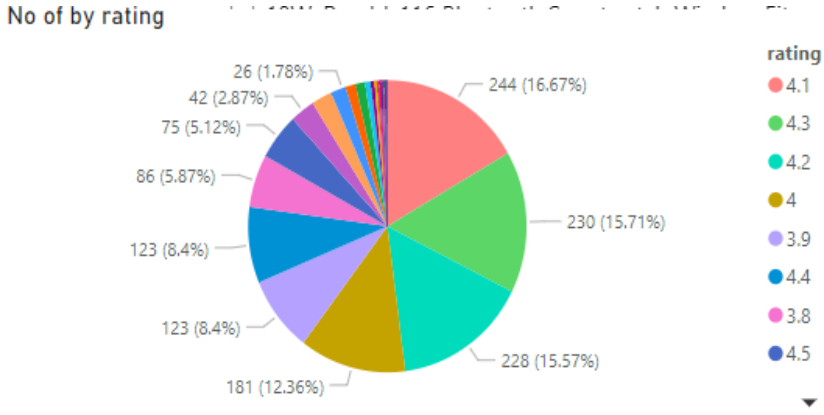
How Does This Support Business Decisions?

- **Optimizing Promotions:** This insight helps businesses allocate their promotional budgets more effectively. For instance, discounts could be adjusted based on customer demand and product performance within each price range.
- **Inventory Management:** Larger discounts on low-cost items might indicate surplus inventory. Businesses can use this information to manage stock levels and forecast demand.

2. Pie Chart: Ratings Distribution

Why this Graph Was Used?

A pie chart is an excellent choice for illustrating **proportional data**. It allows users to see the distribution of customer ratings across various categories at a glance. This visual helps assess the overall customer satisfaction levels for the product portfolio.



What Does the Graph Represent?

- **Segments:** Each slice of the pie represents a specific rating (e.g., 4.1, 4.0, 3.9, etc.).
- **Size of Each Segment:** Indicates the proportion of products that fall within that rating category.
- **Percentage Labels:** Show the exact share of each rating, making it easier to interpret.

Insights Gained from the Visualization

1. **High Ratings Dominate:**

   ○ The largest slices of the pie correspond to ratings **4.1 (16.67%)** and **4.0 (15.71%)**, followed closely by **4.2 (12.36%)** and **4.3 (12.57%)**.
   ○ This indicates that the majority of products are rated highly, reflecting a strong overall customer satisfaction level.

2. **Low Ratings Are Rare:**

○ Ratings below **3.8** constitute a small fraction of the dataset. This suggests that poorly performing products are either minimal in number or phased out quickly.

3. **Opportunities for Improvement:**

○ While most products are rated highly, the presence of lower-rated items (e.g., ratings below 4.0) provides an opportunity to identify and address issues with specific products.

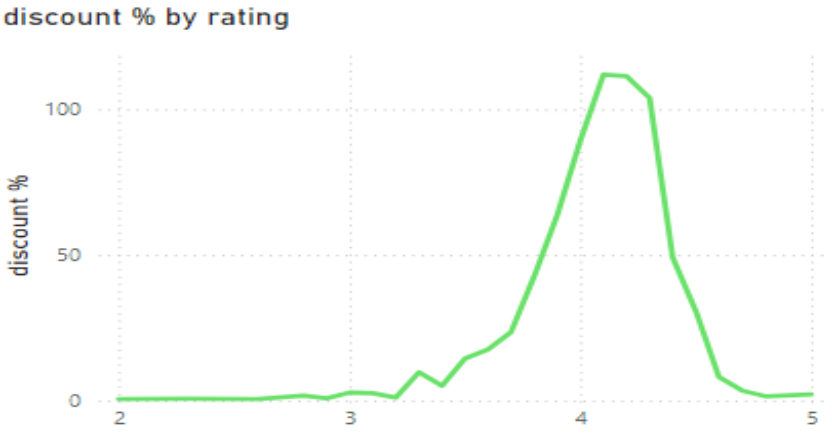
How Does This Support Business Decisions?

- **Focus on Quality Control:** The presence of a few low-rated products highlights the need for targeted quality improvement efforts.
- **Benchmarking Success:** The predominance of high ratings confirms that existing strategies are effective in maintaining customer satisfaction.
- **Promotional Targeting:** High-rated products could be used in marketing campaigns to enhance brand perception.

3. Line Graph: Discount Percentage by Rating

Why this Graph Was Used?

A line graph was chosen to show the **relationship** between product ratings and the discount percentage. It is ideal for identifying trends and patterns across a continuous variable, such as ratings.



What Does the Graph Represent?

- **X-Axis (Horizontal):** Displays the product ratings, ranging from **3.0 to 5.0**.
- **Y-Axis (Vertical):** Represents the discount percentage applied to products.
- **Line Curve:** Tracks how discount percentages change as ratings increase.

Insights Gained from the Visualization

1. **Peak Discounts for Moderate Ratings (3.5–4.0):**

   ○ Products rated between **3.5 and 4.0** receive the highest discounts, often exceeding **50%**.
   ○ This pattern suggests an effort to boost sales or improve the appeal of average-rated products.

2. **Low Discounts for High Ratings (4.5 and Above):**

○ Discounts drop significantly for products rated **4.5 or higher**, indicating that high-rated products are likely in demand and require minimal promotions.

3. **Strategic Discounting:**

○ The focus on discounting moderately-rated products may reflect a strategy to clear inventory or incentivize purchases of less popular items.

How Does This Support Business Decisions?

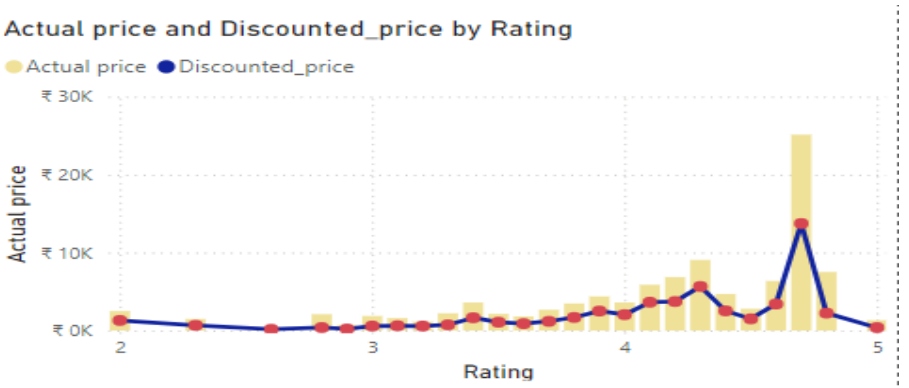
- **Targeted Promotions:** Businesses can focus discounts on products rated **3.5–4.0** to drive sales, while relying on the inherent appeal of higher-rated products to sustain demand.

- **Product Strategy:** The correlation between ratings and discounts helps in identifying which products may need further improvements to reduce reliance on heavy discounting.

4. Combined Chart: Actual Price and Discounted Price by Rating

Why this Graph Was Used?

A combined chart was used to simultaneously compare two variables: **actual price** and **discounted price**, with respect to product ratings. This dual-axis visualization is ideal for identifying relationships between multiple data points.



What Does the Graph Represent?

- **Bar Chart (Blue Bars):** Represents the **actual price** of products across different ratings.
- **Line Chart (Red Line):** Tracks the **discounted price** for products across the same ratings.
- **X-Axis (Horizontal):** Represents product ratings.
- **Y-Axis (Left – for Bars):** Shows actual prices.
- **Y-Axis (Right – for Line):** Shows discounted prices.

Insights Gained from the Visualization

1. **Large Discounts for Moderate Ratings:**

   ○ Products rated between **3.5 and 4.0** show a significant drop between actual prices and discounted prices, reinforcing the observation that these items are heavily discounted.

2. **Minimal Discounts for High Ratings:**

   ○ For products rated **4.5 and above**, the gap between actual and discounted prices narrows, indicating smaller discounts.

3. **Correlation Between Ratings and Price:**

○ Highly-rated products tend to have higher actual and discounted prices, suggesting that customers are willing to pay a premium for quality.

How Does This Support Business Decisions?

- **Pricing Strategy:** Businesses can justify premium pricing for highly-rated products while using discounts strategically for average-rated items.
- **Product Value Proposition:** Understanding the pricing dynamics across ratings helps align marketing strategies with customer expectations.

Below are the interpretations and insights for the three graphs you provided. I've used the required format:

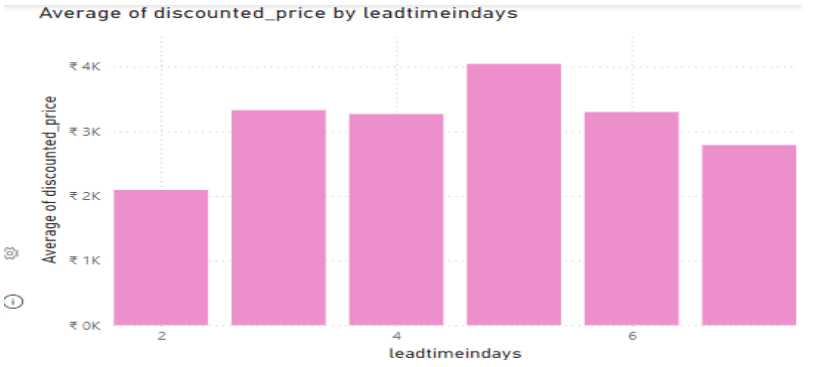
5. Bar Chart: Average Discounted Price by Lead Time in Days

Why this Graph Was Used?

A bar chart was chosen to visualize the average discounted price in relation to the lead time. This format effectively compares averages across discrete time intervals.

What Does the Graph Represent?

- **Bars (Pink):** Display the average discounted price for products across different lead time intervals.
- **X-Axis (Horizontal):** Represents lead time in days.
- **Y-Axis (Vertical):** Represents the average discounted price in currency.



Insights Gained from the Visualization

- **Highest Discounts Around Day 4:** Products with a lead time of 4 days tend to have the highest average discounted price.

- **Lower Discounts for Shorter Lead Times:** Products with a lead time of 2 days have the lowest average discounted price, possibly due to expedited delivery charges.

- **Moderate Discounts for Longer Lead Times:** Discounts gradually decrease after 4 days.

How Does This Support Business Decisions?

- **Pricing Strategy:** Encourage customers to select a lead time of 4 days by offering greater discounts.
- **Logistics Optimization:** Plan inventory and delivery schedules around the most discounted lead times to balance costs and customer satisfaction.

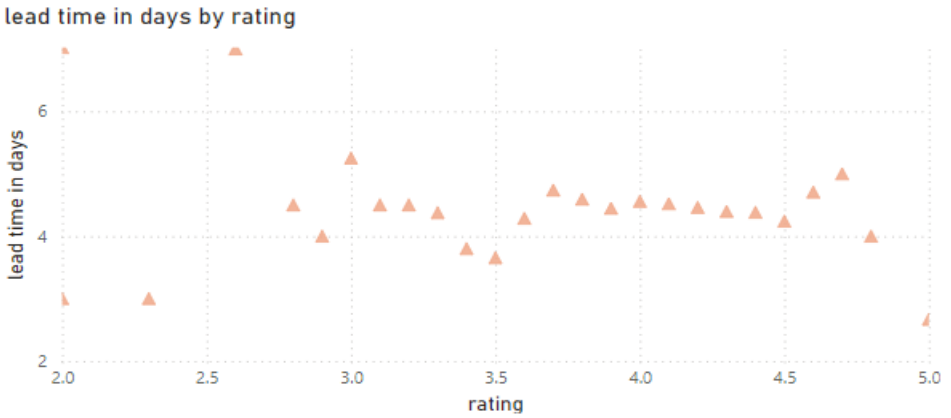
6. Scatter Plot: Lead Time in Days by Rating

Why this Graph Was Used?

A scatter plot was chosen to identify trends or relationships between lead time in days and product ratings. This visualization is ideal for uncovering patterns in scattered data.

What Does the Graph Represent?

- **Points (Triangles):** Each point represents a product with a specific rating and its associated lead time.
- **X-Axis (Horizontal):** Displays product ratings on a scale of 2 to 5.
- **Y-Axis (Vertical):** Represents lead time in days.



Insights Gained from the Visualization

- **Clustered Around 3.5 to 4.5 Ratings:** Most products fall within this range of ratings, with a moderate lead time (3-5 days).

- **Short Lead Times for Lower Ratings:** Products with ratings below 3.0 often have shorter lead times, likely due to reduced demand.

- **Stable Lead Times for Higher Ratings:** Products rated above 4.5 maintain consistent lead times, reflecting streamlined logistics for high-demand items.

How Does This Support Business Decisions?

- **Customer Experience:** Focus on improving the delivery experience for lower-rated products.
- **Inventory Management:** Prioritize efficient lead times for highly-rated products to maintain customer satisfaction.

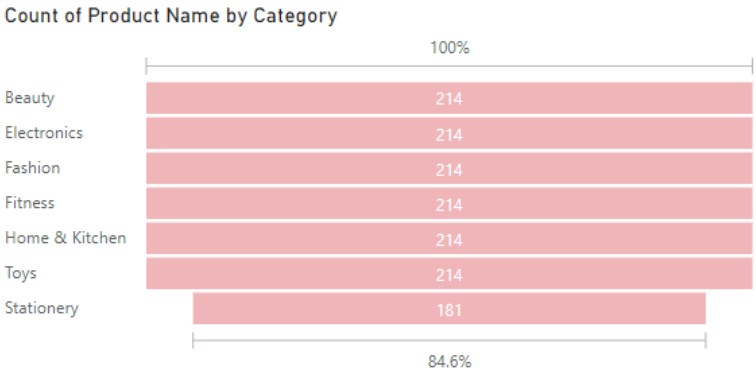
7. Bar Chart: Count of Product Name by Category

Why this Graph Was Used?

A bar chart was used to showcase the distribution of product names across different categories. This format highlights counts effectively.

What Does the Graph Represent?

- **Bars (Pink):** Represent the number of product names in each category.
- **X-Axis (Horizontal):** Categories of products.
- **Y-Axis (Vertical):** Count of product names.



Insights Gained from the Visualization

- **Balanced Distribution Across Categories:** Most categories, such as Beauty, Electronics, Fashion, Fitness, Home & Kitchen, and Toys, have an equal number of product names (214 each).

- **Slightly Lower Count in Stationery:** Stationery has fewer product names (181), indicating a smaller representation in this category.

How Does This Support Business Decisions?

- **Category Expansion:** Increase product diversity in the Stationery category to balance the portfolio.
- **Resource Allocation:** Focus marketing efforts equally across categories with similar representation.

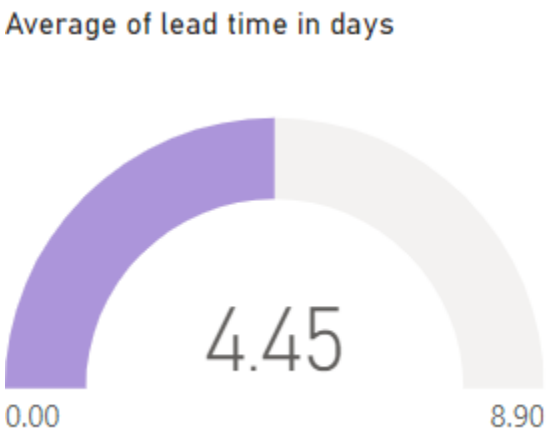
8. Gauge Chart: Average Lead Time in Days

Why this Graph Was Used?

A gauge chart is an effective tool to highlight a single, key metric—in this case, the average lead time in days.

What Does the Graph Represent?

- **Gauge (Semi-circle):** Displays the range of lead times from 0.00 to 8.90 days.
- **Value (4.45):** Indicates the average lead time across all entries.



Insights Gained from the Visualization

- **Moderate Average Lead Time:** The lead time of 4.45 days suggests a balanced delivery performance.

- **Benchmarking:** This value can be used to set targets or compare against competitors.

How Does This Support Business Decisions?

- **Optimization:** Investigate factors contributing to lead times above the average to streamline delivery.
- **Customer Experience:** Communicate this average lead time to set accurate customer expectations.

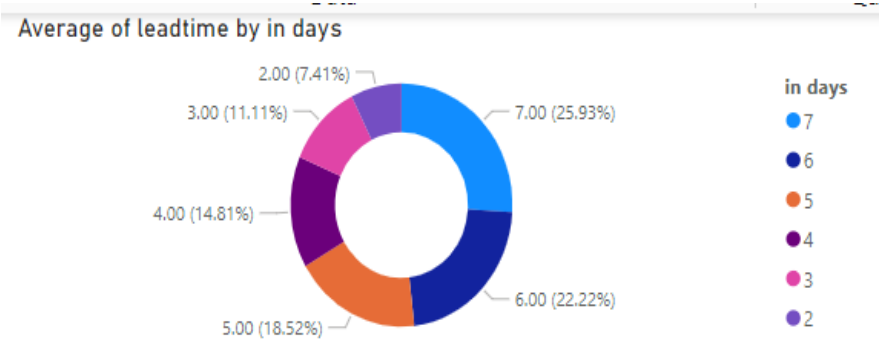
9. Donut Chart: Distribution of Average Lead Time by Day

Why this Graph Was Used?

A donut chart provides a clear breakdown of proportions, making it easy to analyze the percentage of entries falling into different lead time categories.

What Does the Graph Represent?

- **Slices (Colors):** Each slice represents a specific lead time (in days), ranging from 2 to 7 days.
- **Percentages:**
  - **7 Days (25.93%):** Largest segment, indicating a significant proportion of items taking the longest lead time.
  - **6 Days (22.22%):** Second-largest, showing a slightly shorter but still lengthy lead time.
  - **5 Days (18.52%) and 4 Days (14.81%):** Represent moderate lead times.
  - **3 Days (11.11%) and 2 Days (7.41%):** Indicate quicker lead times, but they form the smallest segments.



Insights Gained from the Visualization

- **Dominance of Longer Lead Times:** Nearly half (48.15%) of the entries fall into the 6-7 day range, which might indicate a challenge with quicker deliveries.
- **Low Representation for Shorter Lead Times:** Only 18.52% fall into the 2-3 day range, showing an opportunity to improve expedited services.

How Does This Support Business Decisions?

- **Focus on Quick Deliveries:** Enhance systems to increase the proportion of deliveries within 2-3 days.
- **Performance Analysis:** Identify and mitigate bottlenecks causing longer lead times (6-7 days).

Key Takeaways Across All Visuals

From these four detailed visualizations, the following overarching insights emerge:

- **Customer Ratings:** High-rated products dominate, reflecting strong customer satisfaction.

- **Pricing and Discounts:** Discounts are strategically applied to low-priced and moderate-rated products, indicating a data-driven promotional strategy.

- **Product Optimization:** Lower-rated products provide opportunities for quality improvements.

By leveraging these insights, businesses can refine their strategies for pricing, promotions, inventory management, and customer satisfaction.

4. Features of Power BI Used

1. **Dynamic Visualizations:**
   - Bar charts, pie charts, and line graphs provide a mix of categorical, proportional, and trend-based insights.

2. **Filters and Slicers:**
   - Interactive slicers allow users to segment data by rating, price range, and other criteria.

3. **Custom Measures:**
   - Calculations like average price, discounted price, and discount percentage were created for dynamic analysis.

4. **Interactivity:**
   - Clicking on chart segments updates related visuals, enabling deeper exploration.

5. Recommendations

The dashboard provides numerous insights that can guide business strategy. Below are detailed recommendations based on the Power BI features and the data presented:

1. Enhance Pricing Strategies

- **Observation:** The average price is ₹5,440, but significant discounts bring the average discounted price down to ₹3,130.
- **Recommendation:**
  - Introduce dynamic pricing to adjust prices based on customer demand and competition.
  - Optimize pricing tiers for low-cost products to maximize sales volume while maintaining profitability.

2. Focus on Discount Optimization

- **Observation:** Discounts vary significantly based on price and rating. Products rated 3.5–4.0 receive the highest discounts.
- **Recommendation:**
  - Conduct periodic reviews of discount strategies to ensure they align with sales goals.

○ Gradually reduce discounts on 3.5–4.0 rated products as their popularity increases, improving profit margins.

3. Address Product Quality Issues

- **Observation:** A small percentage of products have ratings below 3.8, indicating quality issues.
- **Recommendation:**
  - ○ Collect detailed customer feedback to identify recurring issues with low-rated products.
  - ○ Prioritize improvements in product design, performance, or packaging to boost ratings.

4. Leverage High-Rated Products

- **Observation:** Products rated 4.5 and above are highly valued but receive minimal discounts.
- **Recommendation:**
  - ○ Use these products in marketing campaigns as flagship items to build trust and attract new customers.
  - ○ Offer loyalty rewards (e.g., discounts on future purchases) to customers who purchase these products.

5. Refine Inventory Management

- **Observation:** Low-priced items receive the highest discounts, suggesting they are overstocked.
- **Recommendation:**
  - Use historical sales data to predict demand and adjust inventory levels.
  - Plan clearance sales for slow-moving inventory to free up storage space.

6. Enhance Customer Segmentation

- **Observation:** Customer behavior varies significantly across price ranges and ratings.
- **Recommendation:**
  - Use customer segmentation to target different groups with tailored offers.
  - Focus on upselling and cross-selling by bundling high-rated premium products with moderately rated items.

7. Improve Dashboard Interactivity

- **Observation:** Filters and slicers are used effectively but can be further enhanced.

- **Recommendation:**
  - Add region-specific slicers to analyze geographic trends in pricing and ratings.
  - Incorporate time-series filters to track changes in sales and ratings over time.

8. Invest in Data-Driven Decision Making

- **Observation:** Power BI features like tooltips and DAX enable granular insights.
- **Recommendation:**
  - Train key business teams in Power BI to maximize the use of its advanced features.
  - Regularly update the dataset and dashboard to reflect real-time performance.