

Data Preparation & Cleaning

- Data structure was reviewed to understand key fields, including transaction details, product categories, pricing, and timestamps.
- Initial checks confirmed no obvious issues such as missing data, outliers, or miscategorized entries.
- A new column for **revenue** was added, calculated as price * quantity.
- This new column was appended to the right of the raw dataset rather than modifying the original structure to maintain data integrity.
- Formatting improvements included applying appropriate data types, color-coding relevant fields for clarity, and adding filters for easier review.

Data Exploration

To gain a deeper understanding of the dataset, key exploratory steps were performed:

- **Transaction Volume Analysis:**
 - Identified **total transactions** (149,456) and examined distribution across the six-month period.
 - Assessed transaction trends by **day of the week** and **hour of the day** to pinpoint peak activity periods.
- **Product and Pricing Insights:**
 - Reviewed unit prices, ranging from \$0.80 (flavor syrup) to \$45.00 (premium coffee beans).
 - Analyzed **product categories**, including bakery items, branded merchandise, coffee, coffee beans, drinking chocolate, flavors, loose tea, packaged chocolate, and tea.

Data Analysis & Visualization

Key metrics were summarized using pivot tables and visualized through charts:

- **Revenue Trends:**
 - Revenue by **month** was analyzed and visualized using a line chart.
- **Transaction Patterns:**
 - Number of transactions by **day of the week** and **hour of the day** were summarized using column charts.
- **Product Performance:**
 - Transactions by **product category** were sorted in descending order to highlight top-selling categories.
 - A detailed breakdown of **transactions and revenue by product type** was filtered to the **Top 15 products**, providing insights into the highest-performing items.