Coffee Shop Sales Data Analysis

Analyzed by

Asadullah Khan

Dataset Name

Coffee Shop Sales

Project Given by

Internncraft

Date of Submission

09 July 2024

Table of Contents

Introduction

Project Overview

The "Coffee Shop Sales" project aims to conduct an analysis of transactional data from a coffee shop to derive insights into product profitability and operational strategies. The dataset consists of detailed transaction records having parameters such as transaction date, time, quantity sold, store location, product details, and more. By analyzing these data points, the project seeks to uncover patterns and trends that influence the financial performance of products and services offered by the coffee shop.

Objectives

The primary objectives of this project include:

Data Cleaning and Integrity: Ensure the dataset is free of inconsistencies, missing values, and duplicates to maintain data accuracy.

Sales Analysis: Identify products and services that contribute positively sales.

Strategic Recommendations: Develop actionable strategies to increase sales.

Scope

The scope of the project encompasses:

- Analysis of transactional data spanning a specified period to understand sales trends and customer preferences.
- Identification of factors impacting product sales, including pricing, store location, and product categories.
- Recommendations aimed at optimizing product offerings and operational efficiencies to enhance overall financial performance.

This introduction sets the stage by outlining the project's goals, objectives, and the scope of analysis based on the provided transaction data. Feel free to adjust or expand upon it as per your specific project requirements.

Data Overview

Description of the Dataset

The dataset for the "Coffee Shop Sales" project contains transactional records having various aspects of customer interactions and product sales at multiple locations. Each row in the dataset represents a specific transaction with the following key attributes:

- **transaction_id:** Unique identifier for each transaction.
- **transaction_time:** Time of day when the transaction occurred (in the format HH:MM:SS).
- **transaction_qty:** Quantity of products sold in the transaction.
- **store_id:** Identifier for the store where the transaction occurred.
- **store_location:** Location of the store where the transaction took place.
- **product_id:** Identifier for the product sold.
- **unit_price:** Price per unit of the product.
- **product_category:** Broad category to which the product belongs (e.g., Flavors, Beverages).
- **product_type:** Type of product within its category (e.g., Regular syrup, Espresso).
- **product_detail:** Specific details about the product (e.g., Chocolate syrup, Decaf Espresso).

Data Sources

The data were sourced from transactional systems within the coffee shop, capturing daily sales activities across multiple stores and product offerings. Each record provides insights into customer preferences, popular products, and sales trends, essential for conducting detailed analysis to optimize profitability and operational efficiency.

Data Cleaning

Review of Data Cleaning Process

Data cleaning is crucial to ensure the accuracy and reliability of analysis. The following steps were undertaken using Python, numpy, pandas, and other libraries:

Handling Missing Values

Missing values were identified and addressed using appropriate methods:

Identifying Missing Values: Utilized pandas' `.isnull()` function to identify columns and rows with missing data.

```
[5]   1 # Checking for null values
      2 df.isnull().sum()

   transaction_id      0
   transaction_date    0
   transaction_time    0
   transaction_qty     0
   store_id            0
   store_location      0
   product_id          0
   unit_price          0
   product_category    0
   product_type        0
   product_detail      0
   dtype: int64
```

Addressing Duplicates

Duplicates were managed to maintain data integrity:

Removing Duplicates: Eliminated duplicate records using pandas' `.drop_duplicates()` method based on transaction identifiers or other unique attributes.

```
[6]   1 # Checking for duplicated values
      2 df.duplicated().sum()

   0
```

Ensuring Data Integrity

To ensure the integrity of the dataset:

Data Type Standardization: Ensured all data types were appropriate for their respective variables (e.g., dates as datetime objects, numeric values as floats or integers).

```
Ensure data integrity and consistency

   1 df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 149116 entries, 0 to 149115
Data columns (total 11 columns):
 #   Column            Non-Null Count   Dtype
---  ------            --------------   -----
 0   transaction_id    149116 non-null  int64
 1   transaction_date  149116 non-null  datetime64[ns]
 2   transaction_time  149116 non-null  object
 3   transaction_qty   149116 non-null  int64
 4   store_id          149116 non-null  int64
 5   store_location    149116 non-null  object
 6   product_id        149116 non-null  int64
 7   unit_price        149116 non-null  float64
 8   product_category  149116 non-null  object
 9   product_type      149116 non-null  object
 10  product_detail    149116 non-null  object
dtypes: datetime64[ns](1), float64(1), int64(4), object(5)
memory usage: 12.5+ MB
```

Converting 'object' type columns to 'category' type columns

Converting to the 'category' type to optimize memory usage and improve performance.

```python
1 df['store_location'] = df['store_location'].astype('category')
2 df['product_category'] = df['product_category'].astype('category')
3 df['product_type'] = df['product_type'].astype('category')
4 df['product_detail'] = df['product_detail'].astype('category')
5
6
7 df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 149116 entries, 0 to 149115
Data columns (total 11 columns):
 #   Column            Non-Null Count   Dtype
---  ------            --------------   -----
 0   transaction_id    149116 non-null  int64
 1   transaction_date  149116 non-null  datetime64[ns]
 2   transaction_time  149116 non-null  object
 3   transaction_qty   149116 non-null  int64
 4   store_id          149116 non-null  int64
 5   store_location    149116 non-null  category
 6   product_id        149116 non-null  int64
 7   unit_price        149116 non-null  float64
 8   product_category  149116 non-null  category
 9   product_type      149116 non-null  category
 10  product_detail    149116 non-null  category
dtypes: category(4), datetime64[ns](1), float64(1), int64(4), object(1)
memory usage: 8.5+ MB
```

Activate Windows
Go to Settings to activate Windows.

Formatting the 'transaction_time' column into proper timedate type

Formatting the 'transaction_time' column

```python
1 # Convert 'transaction_time' to proper time format
2 df['transaction_time'] = pd.to_datetime(df['transaction_time'], format='%H:%M:%S', errors='coerce').dt.time
```

Formatting the 'object' type columns

Formatting the object type columns

```python
1 # Convert all non-string values in object columns to strings
2 for col in df.select_dtypes(include=['object']).columns:
3     df[col] = df[col].astype(str)
4
5 # Strip leading/trailing spaces from object type columns
6 for col in df.select_dtypes(include=['object']).columns:
7     df[col] = df[col].str.strip()
8
```

Sales Analysis

Visualization of Sales Data

Visualizing sales data provides insights into trends and performance metrics:

Time Series Analysis
Plotting sales over time (daily, monthly) to identify seasonal trends or fluctuations.

- **Sales on Daily Basis:**

1. **Fluctuating Sales**: The chart shows variability in daily sales, with peaks and troughs indicating fluctuating customer demand.

2. **Periodic Variations**: Regular patterns in sales suggest opportunities for targeted marketing and promotional activities during high-demand periods.

- **Sales on Monthly Basis**

From the "Sales by Month" chart displaying data from January 2021 to June 2021, several key insights can be drawn:

1. **Consistent Upward Trend:** The chart illustrates a consistent upward trend in sales over the six-month period. Starting at approximately $90,000 in January, sales steadily increase month-over-month, reaching around $180,000 in June. This trend indicates a positive growth trajectory for the coffee shop's sales.

2. **Seasonal Variations**: While the overall trend is upward, there may be seasonal variations influencing sales. For instance, sales tend to rise as the months progress from winter into summer, which could be influenced by factors such as weather, seasonal promotions, or customer preferences.

3. **Forecasting and Planning**: The consistent upward trend in sales provides valuable insights for forecasting and planning purposes. Coffee shop management can use this data to anticipate future sales patterns, adjust inventory levels accordingly, and strategically plan marketing campaigns and promotions to capitalize on peak sales periods.

Store Location Analysis

Visualizations based on location to compare sales across different store locations.
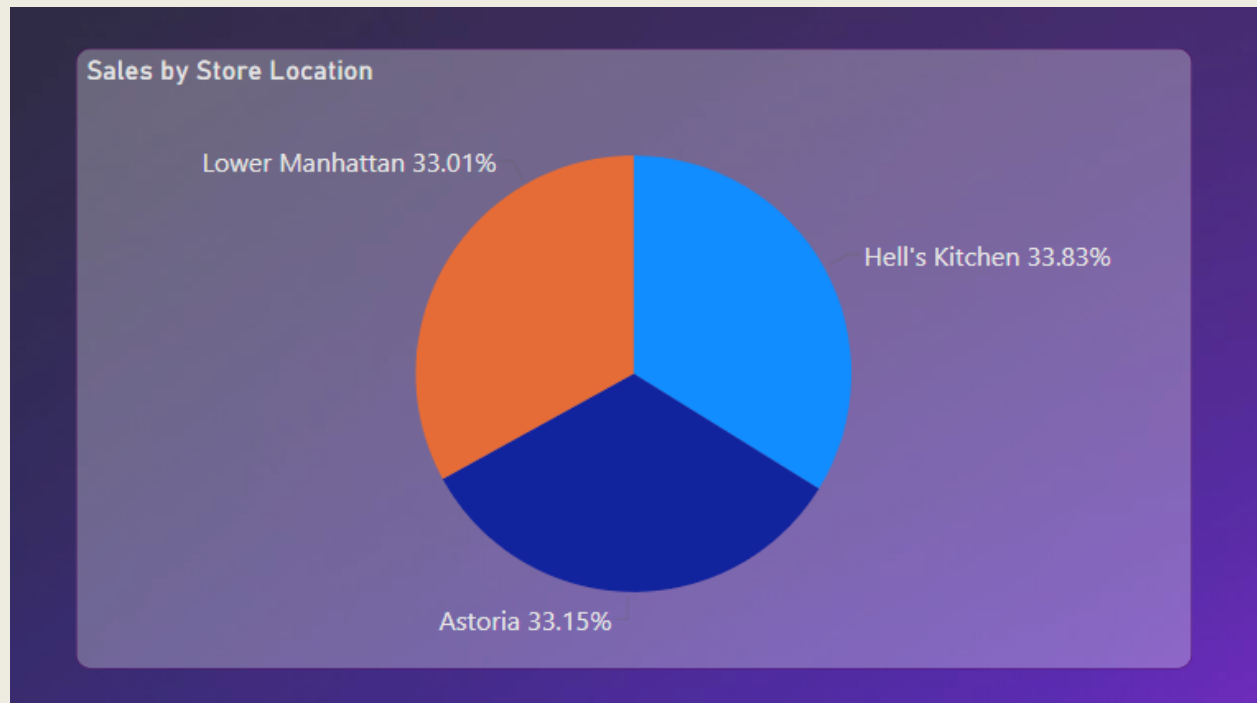
**Store Comparison:** Comparing sales performance between different store locations.

1. **Balanced Distribution**: Sales are almost evenly split among the three locations, with Hell's Kitchen having a slightly higher percentage of sales.

2. **Minor Variations**: The small differences in sales percentages suggest that all locations are performing similarly, with no single store significantly outperforming or underperforming.

Product Category

Bar charts visualize sales distribution by product categories.

Identification of Top Selling Products

- **By Product Category**
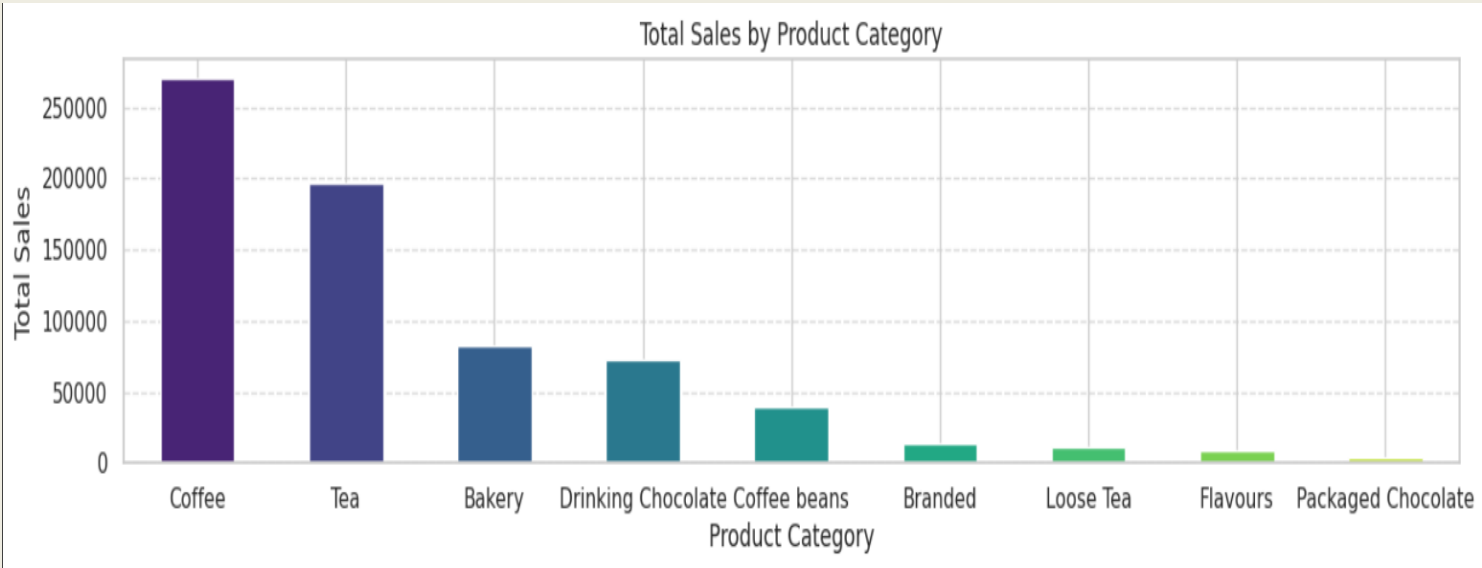
1. **Highest Sales**: The Coffee category leads in sales, significantly outperforming other categories.

2. **Moderate Sales**: The Tea category follows, with moderate sales figures, while Bakery and Drinking Chocolate also contribute notably.

3. **Lower Sales**: Categories such as Coffee beans, Branded items, Loose Tea, Flavours, and Packaged Chocolate show comparatively low sales.

- **By Product Type**

Top 10 Soled Product Types
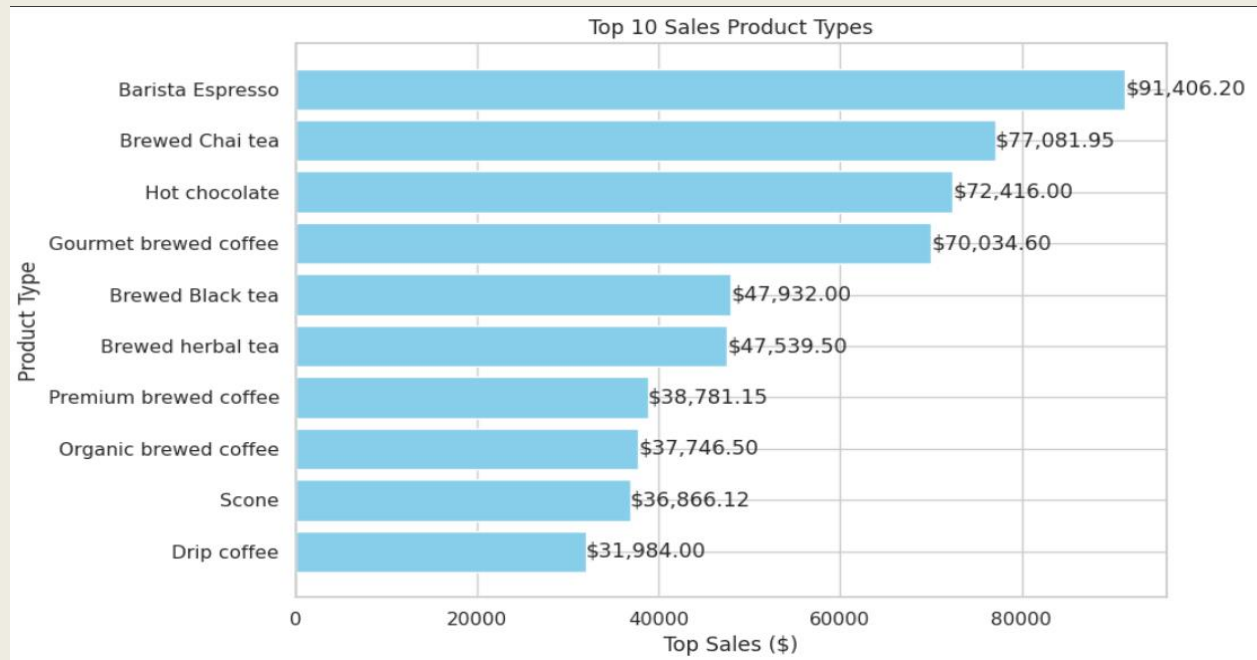
1. **Highest Sales**: Barista Espresso, Brewed Coffee, and Hot Chocolate are the top-performing products.

2. **Descending Trend**: Sales figures show a descending trend from Barista Espresso to Drip Coffee.

3. **Least Popular**: Drip Coffee has the lowest sales among the listed products.

Lowest 10 Sold Product Types

Top 10 Products with Lowest Total Transaction Quantity

The chart titled "Lowest 10 Sales Product Types" provides insights into the sales performance of various product types with the lowest sales figures. Key observations include:

1. **Least Popular**: Green Beans have the lowest sales among the listed products.

2. **Descending Trend**: There is a clear descending trend in sales from House Blend Beans to Green Tea.

3. **Highest Sales among the Lowest**: Gourmet Beans and Black Tea and Herbal Tea have the highest sales within this category.

Profit/Loss Analysis

In conducting a comprehensive profit and loss analysis, it is essential to have data on both revenues and costs to accurately calculate profitability. However, the provided dataset for the "Coffee Shop Sales" project does not include a column for "profit_loss" or detailed "cost" data, which are critical for direct profit and loss calculations.

Sales as a Proxy for Profitability

Despite this limitation, we can leverage the available "sales" data as a proxy to gain insights into profitability. In most cases, higher sales are directly proportional to higher profits, assuming cost structures are consistent across products. Therefore, analyzing sales data can still provide valuable indications of product performance and potential profitability.

Analysis Approach

Visualization of Sales Data: By visualizing sales data, we can identify trends and patterns that suggest which products are likely contributing positively to overall profitability.

Sales Trends: Time series analysis of sales to observe peak periods and seasonal trends.

Product Performance: Bar charts or pie charts to display topselling products and categories, indicating which items are likely more profitable.

Identification of Top Selling Products:

Identifying products with the highest sales can provide insights into which items are generating the most revenue and, by extension, likely the most profit.

Top Performers: Ranking products by sales volume or revenue to highlight the best performers.

Comparative Analysis: Conducting comparative analyses across different stores, product categories, and time periods to identify variations in sales performance, which can help infer potential profitability differences.

Store Comparisons: Analyzing sales data across various store locations to determine which locations are performing better.

Category Comparisons: Comparing sales across different product categories to identify high revenue generating categories.

Conclusion

While the absence of explicit profit and cost data limits the ability to perform a detailed profit and loss analysis, the available sales data still offers significant insights. By focusing on sales as a proxy for profitability, we can identify key products and trends that are likely contributing to the coffee shop's financial performance. Future analyses would benefit greatly from the inclusion of cost data to enable more precise profitability calculations.

Strategies for Profit Increase

## Recommendations Based on Sales Analysis

The chart "Sum of Sales by Day" provides valuable insights into daily sales data from January 2021 to June 2021.

- **Dynamic Pricing Strategies**: Implement dynamic pricing models to capitalize on peak sales periods, adjusting prices based on demand fluctuations to maximize revenue.

- **Promotional Campaigns**: Schedule promotional activities and discounts during lowsales periods to boost customer engagement and drive additional sales.

- **Customer Feedback Integration**: Collect and analyze customer feedback to understand preferences and trends. Use this information to refine product offerings, introducing new items that align with customer demands and phasing out less popular ones.

Recommendations Based on Location Analysis
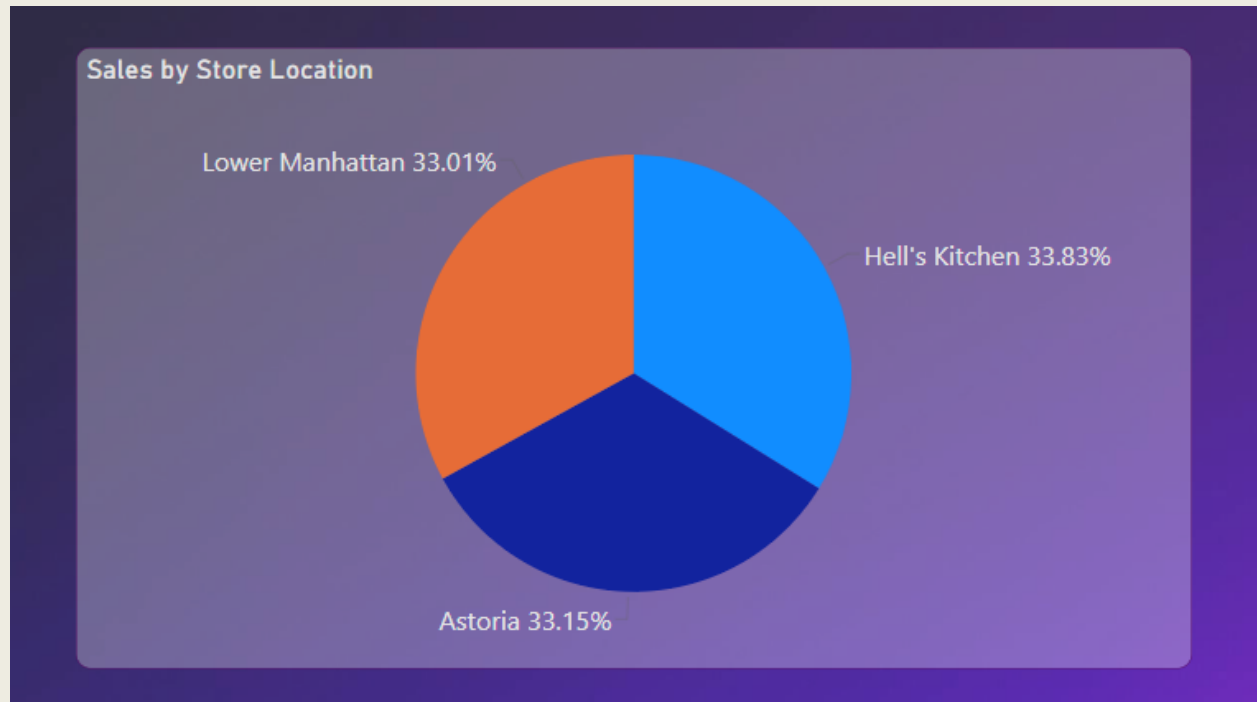
The chart titled "Sales by Store Location" provides insights into the distribution of sales across three different store locations: Lower Manhattan, Hell's Kitchen, and Astoria.

- **Localized Marketing Strategies**: Develop marketing strategies tailored to each store location's customer base to maximize engagement and drive sales. For instance, Hell's Kitchen could benefit from promoting its slightly higher sales performance to attract more customers.

- **Uniform Customer Experience**: Ensure a consistent and high-quality customer experience across all locations to maintain balanced sales distribution and foster customer loyalty.

- **Cost Optimization**: Analyze operating costs for each location and identify opportunities to reduce expenses without compromising customer service or product quality.

- **Location-Specific Preferences**: Understand and cater to the unique preferences of customers at each location. Offer tailored product assortments that align with local tastes and preferences.

- **Product Promotion**: Highlight and promote high-margin products across all locations to increase overall profitability. Ensure that these products are readily available and well-marketed.

Recommendations Based on Product Category Analysis
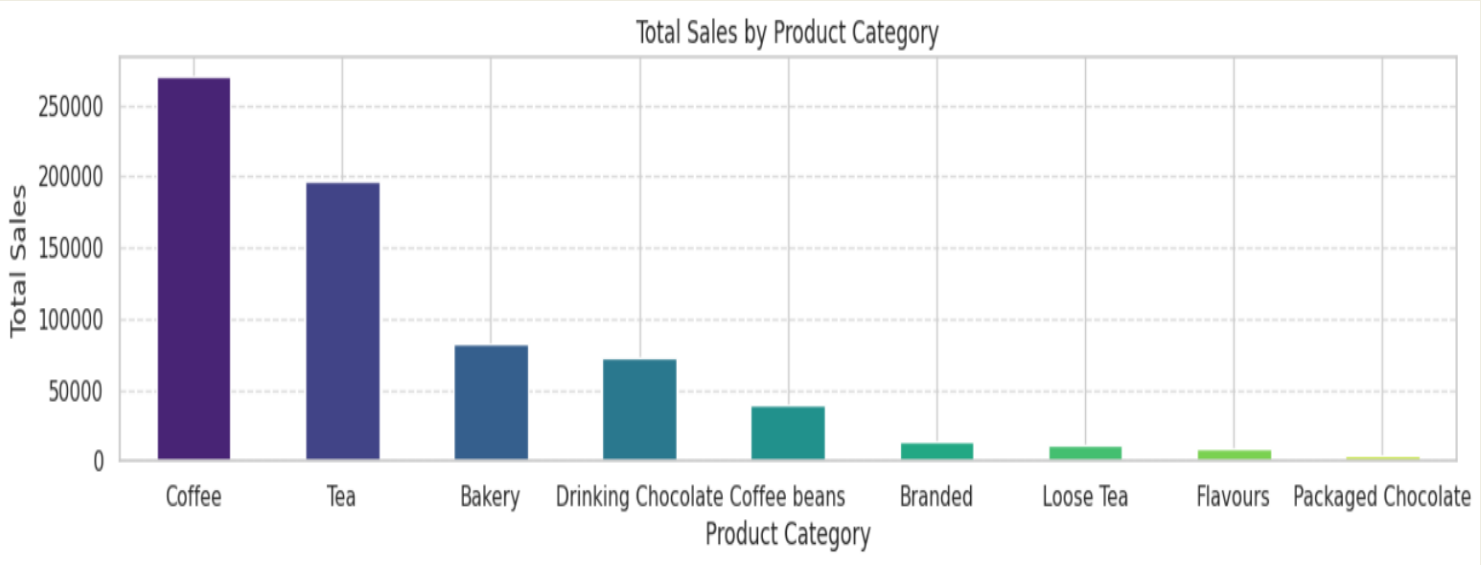
The chart titled "Total Sales by Product Category" provides valuable insights into the sales performance of various product categories.

- **Focus on High Performers**: Given the high sales of Coffee, further investment in this category could yield substantial returns. Introduce new coffee variants, promote premium offerings, and enhance customer engagement through loyalty programs focused on coffee purchases.

- **Boost Moderate Performers**: The Tea, Bakery, and Drinking Chocolate categories show potential for growth. Targeted marketing campaigns and promotions can help increase their sales. Consider bundling these products with high-selling items to boost their visibility and appeal.

- **Revitalize Low Performers**: For categories with lower sales, such as Coffee beans, Branded items, Loose Tea, Flavours, and Packaged Chocolate, explore reasons for their underperformance. Adjust product offerings, refine marketing strategies, and consider product repositioning to improve their sales.

- **Cost Reduction**: Review supply chain and production processes to identify cost-saving opportunities. Negotiate better terms with suppliers and streamline operations to reduce costs without compromising quality.

- **Pricing Strategies**: Implement dynamic pricing strategies, offering discounts during off-peak periods and premium pricing for high-demand items. Regularly review and adjust prices based on market trends and competition.

- **Promotion and Up-selling**: Develop targeted promotions and up-sell strategies to encourage higher spending per transaction. Offer bundle deals, loyalty rewards, and limited-time offers to drive sales.
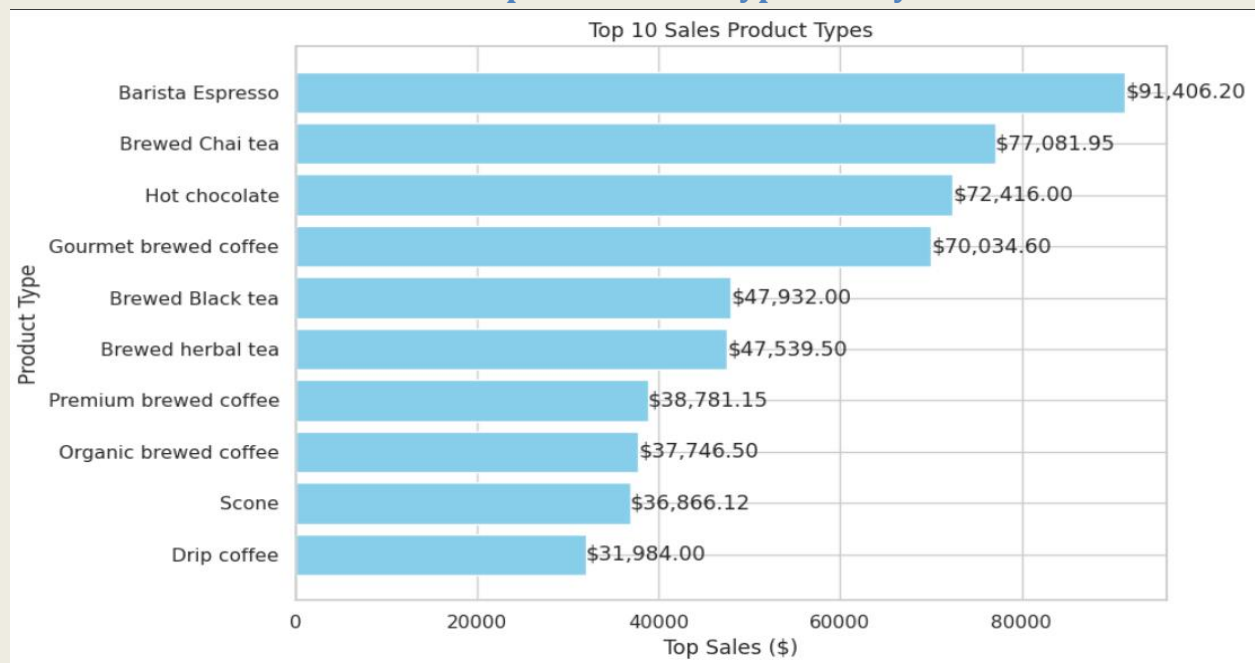
Product-Specific Recommendations

Based on the sales data, specific recommendations for each product category are:

1. Coffee: Continue to innovate and introduce new flavors or limited-edition blends. Leverage customer feedback to enhance offerings and maintain the high sales momentum.

2. Tea: Expand the range of tea products, including specialty and organic options. Promote tea during colder seasons and explore tea-based beverages to attract a broader customer base.

3. Bakery and Drinking Chocolate: Enhance product presentation and offer complementary items to increase sales. Introduce seasonal bakery items and premium drinking chocolate options to attract customers.

4. Coffee Beans: Educate customers on the benefits of home brewing and offer brewing equipment or starter kits to boost sales of coffee beans.

5. Branded Items: Reevaluate the product line and focus on items that align with the coffee shop's brand image. Consider exclusive merchandise or collaborations to increase appeal.

6. Loose Tea: Highlight the freshness and quality of loose tea. Offer sample packs and provide information on the health benefits of loose tea to attract health-conscious customers.

7. Flavours: Promote flavored syrups and additives as customizable options for beverages. Encourage customers to personalize their drinks with these add-ons.

8. Packaged Chocolate: Improve product placement and consider partnerships with local chocolatiers for exclusive offerings. Promote packaged chocolate as a gift option during festive seasons.

By implementing these strategies, the coffee shop can effectively increase profit margins and enhance overall profitability. The "Total Sales by Product Category" chart serves as a critical tool in understanding sales distribution and guiding strategic initiatives for each product category.

Recommendations Based on Top 10 Product Types Analysis

Top 10 Sales Product Types

Product Type	Top Sales ($)
Barista Espresso	$91,406.20
Brewed Chai tea	$77,081.95
Hot chocolate	$72,416.00
Gourmet brewed coffee	$70,034.60
Brewed Black tea	$47,932.00
Brewed herbal tea	$47,539.50
Premium brewed coffee	$38,781.15
Organic brewed coffee	$37,746.50
Scone	$36,866.12
Drip coffee	$31,984.00

Strategies for Profit Increase

- **Maximize High Performers:** Focus on promoting and expanding the top-selling products such as Barista Espresso, Brewed Coffee, and Hot Chocolate. Develop marketing campaigns around these products to boost sales further.

- **Targeted Promotions**: Create promotions and discounts for moderate and lower-performing products to increase their visibility and sales. Highlight the unique features and benefits of these products to attract customers.

- **Customer Feedback Integration**: Collect and analyze customer feedback to understand preferences and improve product offerings, ensuring they meet customer demands and expectations.

Loss Mitigation

Analysis of Factors Contributing to Losses

In conducting a comprehensive loss analysis, it is essential to have data on both revenues and costs to accurately identify products or services contributing to losses. However, the provided dataset for the "Coffee Shop Sales" project does not include a column for "profit_loss" or detailed "cost" data, which are critical for direct loss calculations.

Sales as a Proxy for Identifying Losses

Despite this limitation, we can leverage the available "sales" data as a proxy to gain insights into potential losses. Products with lower sales figures are likely to be contributing less to overall revenue, which can be an indicator of potential loss-making items if their associated costs are significant.

Analysis Approach

1. **Visualization of Sales Data:**
2. **Sales Trends:** By visualizing sales data, we can identify trends and patterns that suggest which products are underperforming. Time series analysis of sales helps observe periods of low sales and identify persistent low-performing products.
3. **Product Performance:** Using bar charts or pie charts to display products with the lowest sales, indicating which items are likely underperforming and potentially contributing to losses.

Identification of Low Selling Products:

Underperformers: Identifying products with the lowest sales can provide insights into which items are generating the least revenue and, by extension, potentially incurring losses. Ranking products by sales volume highlights the underperformers.

Strategies for Loss Mitigation

Analysis of Factors Contributing to Losses

The analysis of transaction quantities for various products reveals insights into factors contributing to potential losses. Products with lower transaction quantities are indicative of lower demand, which can lead to underutilization of resources and potential loss-making scenarios for the coffee shop.

Recommendations Based on Lowest 10 Sold Product Types Analysis
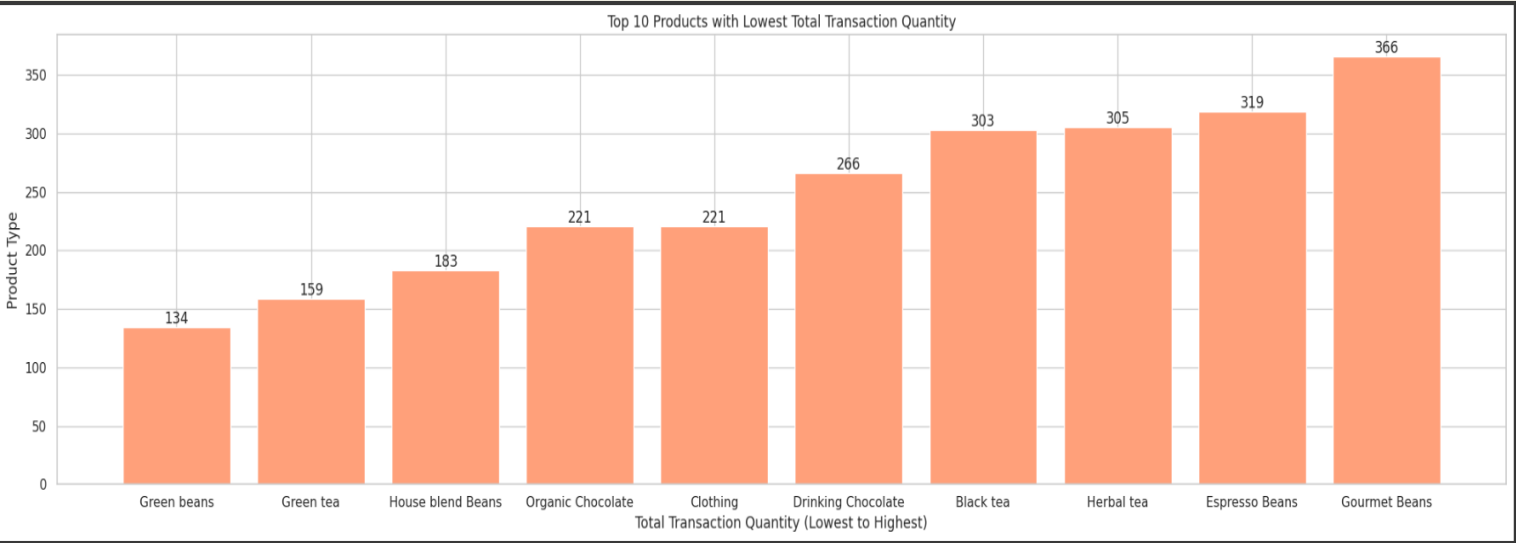
Strategies to Minimize or Eliminate Losses

- **Product Review and Adjustment:**
    1. Evaluate Demand: Conduct further market research or customer surveys to understand why these products have low transaction quantities.

    2. **Product Mix Optimization**: Consider reducing the variety of lowdemand products or adjusting their offerings to better align with customer preferences.

- **Promotional Strategies:**
  1. **Promotions and Discounts**: Offer promotions or discounts specifically targeted at increasing sales of low-demand products.

  2. **Bundle Offers**: Create bundles with popular items to encourage sales of less popular products.

- **Operational Efficiency:**
  1. **Inventory Management**: Implement better inventory management practices to reduce costs associated with maintaining low-selling products.

  2. **Production Optimization**: Adjust production schedules or quantities to minimize waste and maximize efficiency.

Actionable Steps to Convert Loss-Making Products into Profitable Ones

- **Market Positioning:**
  1. **Repositioning**: Consider repositioning low-demand products to target niche markets or specific customer segments that may value these products more.

  2. **Brand Perception**: Enhance the perceived value of products through branding or packaging improvements.

- **Customer Engagement:**
  1. **Customer Education:** Educate customers about the benefits or unique qualities of low-demand products through marketing campaigns or tastings.

  2. **Feedback Integration**: Use customer feedback to continuously improve and refine product offerings.

By focusing on products with low transaction quantities, the coffee shop can proactively mitigate potential losses and transform underperforming products into profitable assets. Implementing targeted strategies based on detailed analysis and customer insights will be crucial in achieving sustainable profitability and enhancing overall business performance.

Project Summary and Conclusion

Introduction

The "Coffee Shop Sales" project delves into the analysis of transactional data from a coffee shop to uncover insights into product profitability and operational strategies. This dataset comprises detailed records capturing various aspects of customer interactions and product sales across multiple locations.

Project Overview

The project aims to analyze transactional data to understand sales trends, identify top-performing products, and develop strategic recommendations to optimize profitability and operational efficiency.

Objectives

1. **Data Cleaning and Integrity**: Ensure data accuracy by addressing inconsistencies, missing values, and duplicates.

2. **Sales Analysis**: Identify high-performing products and categories through comprehensive sales analysis.

3. **Strategic Recommendations**: Develop actionable strategies to increase sales and enhance overall business performance.

Scope

The scope includes analyzing transactional data over a specified period to uncover sales patterns, identifying factors influencing product sales, and recommending optimizations to improve financial outcomes.

Data Overview

Description of the Dataset

The dataset includes transactional details such as transaction ID, time, quantity sold, store location, product details, and more, sourced from the coffee shop's operational systems.

Data Sources

Data were sourced from transactional systems within the coffee shop, offering insights into customer preferences and sales trends crucial for strategic decision-making.

Data Cleaning

The data cleaning process involved handling missing values, addressing duplicates, and ensuring data integrity through standardization of data types and format.

Sales Analysis

Visualization of Sales Data

Visualizations such as time series analyses and location-based comparisons provided insights into sales trends and performance metrics across different store locations and product categories.

Profit/Loss Analysis

Despite the absence of explicit profit and cost data, sales data served as a proxy to analyze profitability trends, identify top-selling products, and infer potential profitability differences across products and categories.

Recommendations

Strategies for Profit Increase

- Based on detailed sales analysis:

Dynamic Pricing Strategies: Implement dynamic pricing models to capitalize on peak sales periods.

Promotional Campaigns: Schedule targeted promotions to boost sales during low-demand periods.

Customer Feedback Integration: Utilize customer feedback to refine product offerings and enhance customer satisfaction.

Strategies for Loss Mitigation

Analysis of Factors Contributing to Losses: Identify underperforming products using sales data as indicators.

Product Review and Adjustment: Evaluate and optimize product offerings to minimize losses and enhance profitability.

Operational Efficiency: Improve inventory management and production processes to reduce costs associated with low-selling products.

This summary integrates the key aspects of your project, from data handling and analysis to strategic recommendations, providing a cohesive overview of your findings and proposed actions.