

AI for Data Analysis & Image Generation

July 17, 2023



ChatGPT as Data Analyst



Bakery Sales Data.csv



Finished working

Show work ▾

The dataset contains 20,507 rows and 5 columns. Here are some additional details

1. **TransactionNo**: There are 9,465 unique transaction numbers. This indicates that some transactions included the purchase of multiple items.
2. **Items**: There are 94 unique items that were purchased.
3. **DateTime**: There are 9,465 unique dates and times, which matches with the number of unique transactions. This suggests that each transaction has a unique timestamp.
4. **Daypart**: There are 4 unique dayparts. These could be categories like "Morning", "Afternoon", "Evening", etc.
5. **DayType**: There are 2 unique day types, likely "Weekday" and "Weekend".

Importantly, there are no missing values in the dataset, which is good as it means we won't have to deal with imputation or dropping rows.

Z1

Let's look into the frequent items and the best sellers



Finished working

Show work ▼

Here are the top 10 most frequently purchased items:

1. Coffee: 5471 transactions
2. Bread: 3325 transactions
3. Tea: 1435 transactions
4. Cake: 1025 transactions
5. Pastry: 856 transactions
6. Sandwich: 771 transactions
7. Medialuna: 616 transactions
8. Hot chocolate: 590 transactions
9. Cookies: 540 transactions
10. Brownie: 379 transactions

This list shows the best-selling items in the bakery, with Coffee being the most popular, followed by Bread and Tea.

Z1

Please put in table format with a heading for each column



Finished working

Show work ▼

Here are the top 10 most frequently purchased items presented in a table format:

	Item	Transactions
0	Coffee	5471
1	Bread	3325
2	Tea	1435
3	Cake	1025
4	Pastry	856
5	Sandwich	771
6	Medialuna	616
7	Hot chocolate	590
8	Cookies	540
9	Brownie	379

Z1

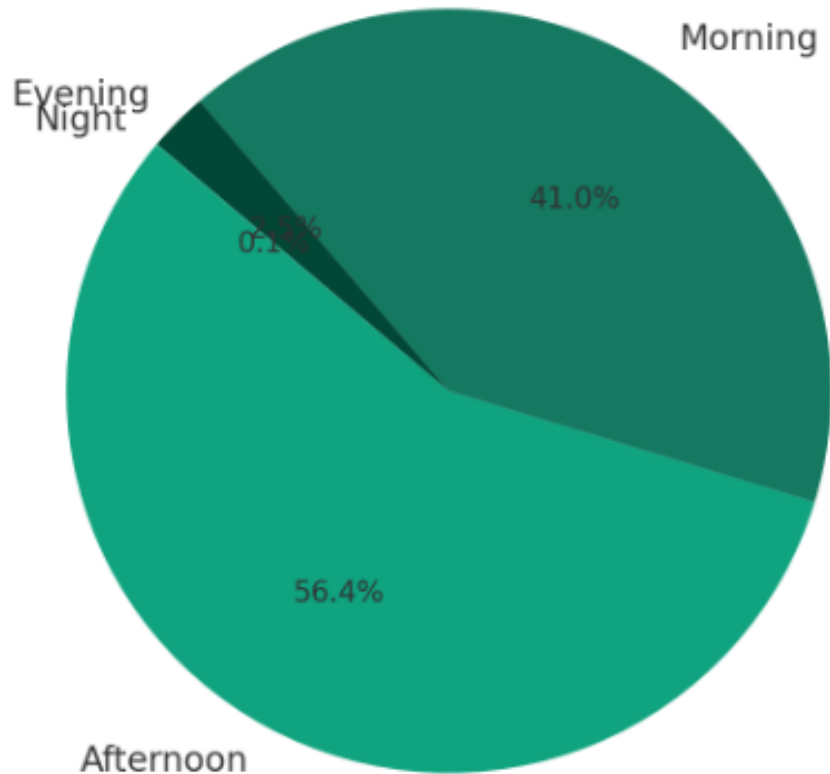
Graph in a pie chart the peak hours of sales



Finished working

Show work ▼

Sales Distribution by Daypart



From the chart, we can see that:

- The afternoon is the peak time for sales, accounting for 56.4% of all transactions.
- The morning is the second busiest period, with 30.7% of sales.
- The evening has a smaller proportion of sales at 12.4%.
- The night period has the least sales, contributing only 0.5% of total transactions.

These insights can help the bakery plan their operations, such as staff scheduling and inventory management, based on the peak sales hours.

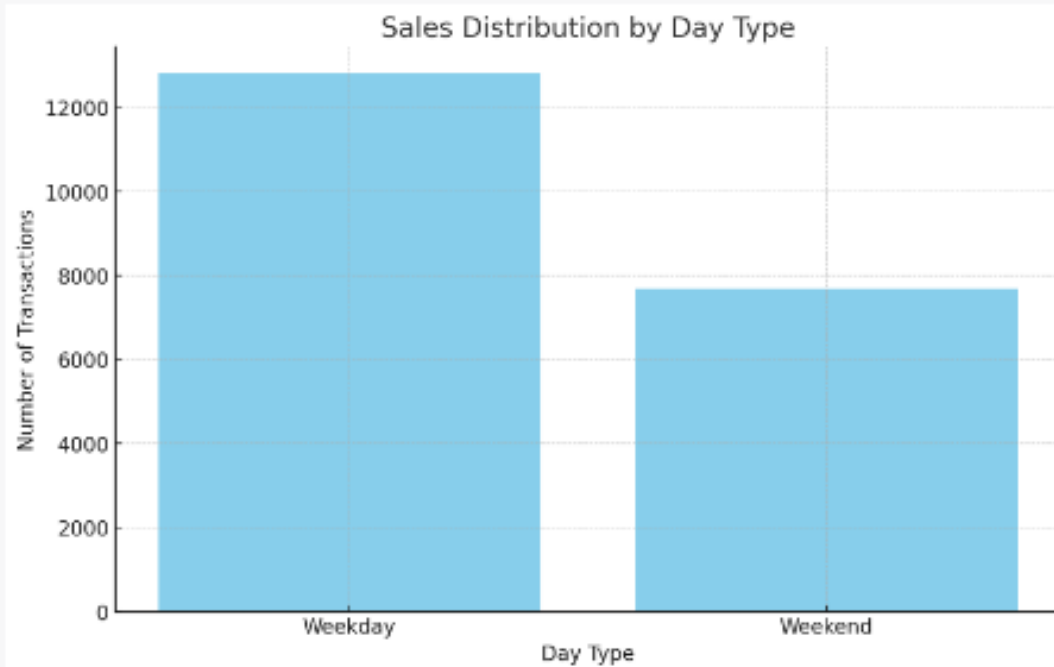
Z1

In a bar chart, ordered decending, list the most productive sales days



Finished working

Show work ▼



The bar chart displays the sales distribution by day type:

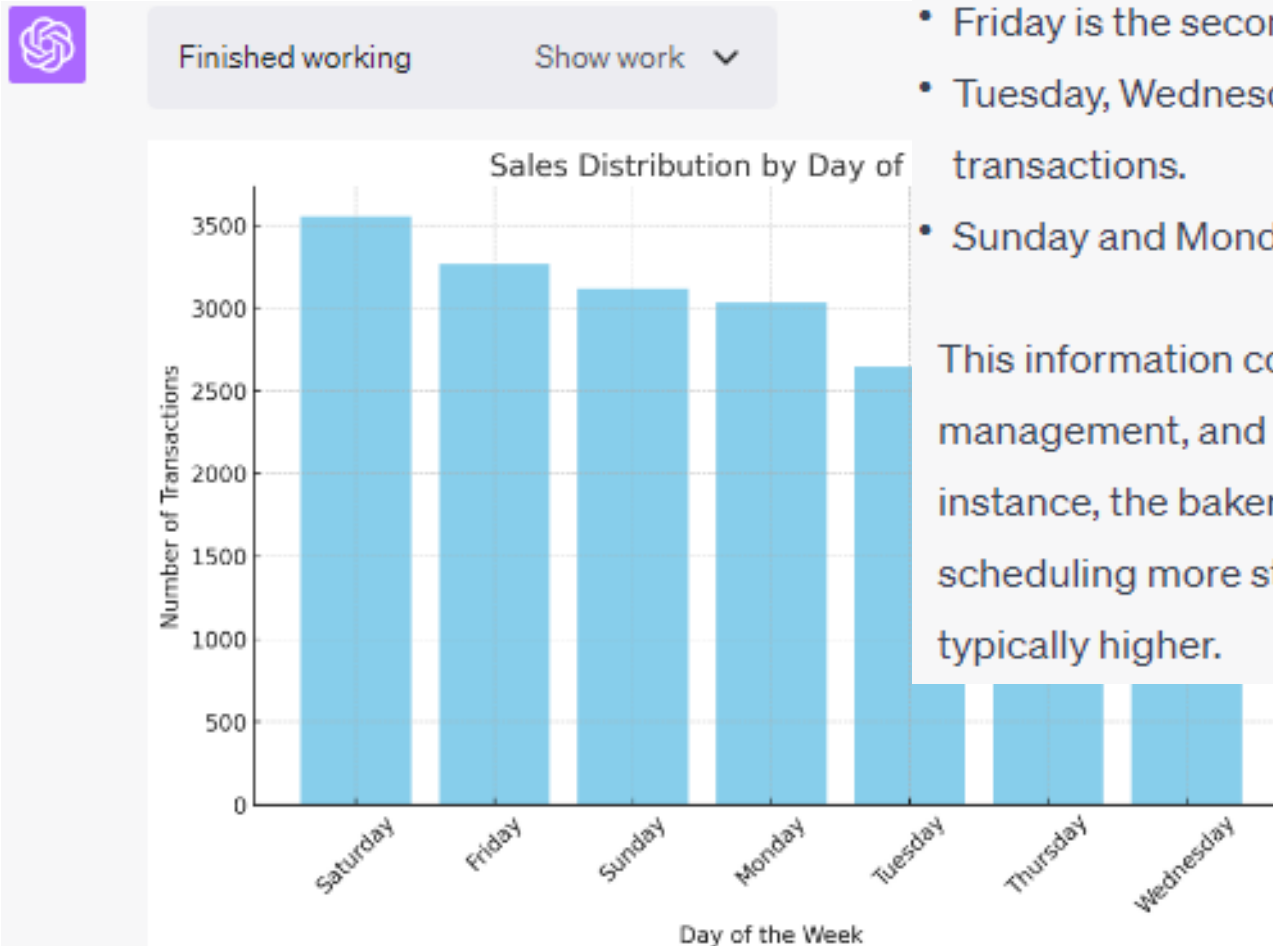
- Weekdays are more productive for sales with a significantly higher number of transactions compared to weekends.

This information could be useful for staffing, inventory management, and promotional activities. For example, the bakery might consider stocking more goods or scheduling more staff on weekdays when sales are typically higher.

Using this new DayofWeek variable, sorted in descending order, list the most productive days of the week.

The bar chart displays the sales distribution by the day of the week:

- Saturday has the highest number of transactions, making it the most productive day for sales.
- Friday is the second most productive day.
- Tuesday, Wednesday, and Thursday have a similar number of transactions.
- Sunday and Monday have the fewest transactions.



This information could be useful for staffing, inventory management, and planning promotional activities. For instance, the bakery might consider stocking more goods or scheduling more staff on Saturdays and Fridays when sales are typically higher.

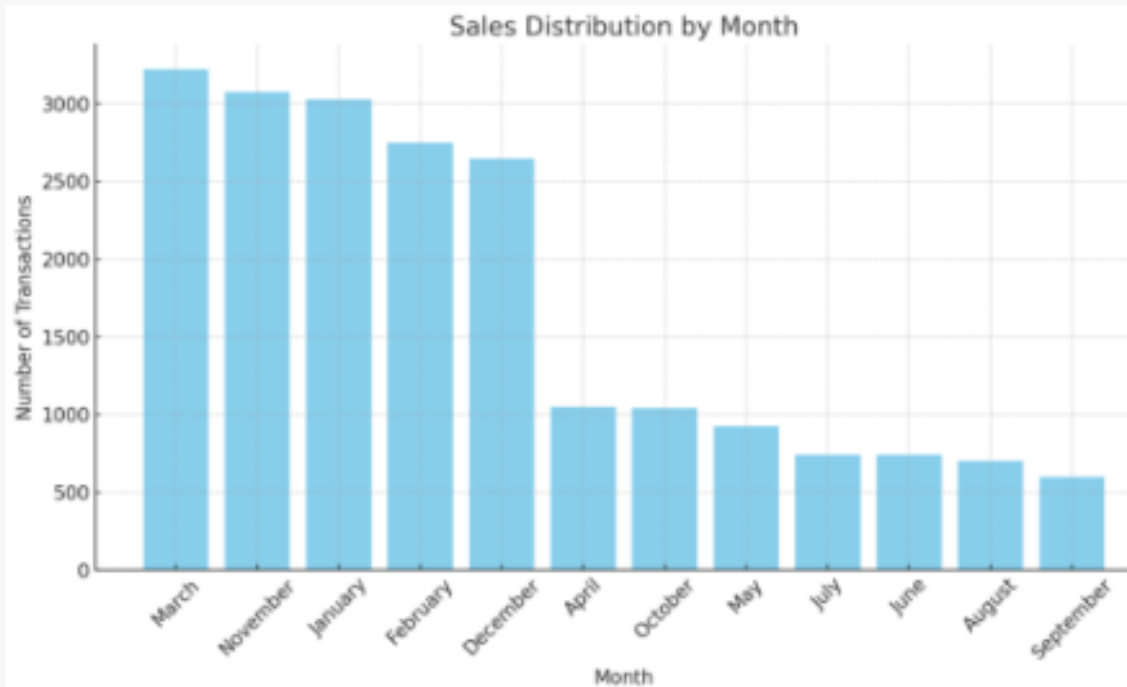
Z1

Create a bar chart in descending order of the most productive months of sales



Finished working

Show work ▼



The bar chart displays the sales distribution by month:

- March is the most productive month in terms of sales with the highest number of transactions.
- February and November also have a high number of transactions.
- January, May, and June have the least number of transactions.

This information could be useful for planning promotional activities, inventory management, and staffing. The bakery might consider stocking more goods or scheduling more staff during March, February, and November when sales are typically higher.

Z1

Conduct a market basket analysis



I've grouped the data by transaction number and created a list of items for each transaction. Here are the first few transactions:

TransactionNo	Items
1	[Bread]
2	[Scandinavian, Scandinavian]
3	[Hot chocolate, Jam, Cookies]
4	[Muffin]
5	[Coffee, Pastry, Bread]

Next, we'll transform this data into a boolean matrix where the rows represent transactions and the columns represent items, with each cell indicating whether a specific item was included in a specific transaction. Then, we'll conduct the market basket analysis using the Apriori algorithm.

However, before we start, we need to transform our data into a format that can be used by the algorithm. In this case, we need a DataFrame where each row represents a transaction and each column represents an item, with boolean values indicating whether the item was included in the transaction.

Let's start by transforming the data.

P |

I apologize for the confusion, but it appears there was an error because the `mlxtend` library is not available in the current environment. Unfortunately, as of my last training cut-off in September 2021, this environment does not support the installation of additional packages.

However, I can guide you through the process so you can perform the analysis in your own local Python environment. Here's how you can proceed:

1. Install the necessary library, if not already installed: You can do this by running `!pip install mlxtend` in your local Jupyter notebook or `pip install mlxtend` in your command line.
2. Use the `TransactionEncoder` function from `mlxtend.preprocessing` to transform your data into the right format:

python

Copy code

```
from mlxtend.preprocessing import TransactionEncoder

te = TransactionEncoder()
te_array = te.fit(transactions).transform(transactions)
transactions_matrix = pd.DataFrame(te_array, columns=
```

1. Use the `apriori` function from `mlxtend.frequent_patterns` to find frequent itemsets:


Z1

What menu items should be expanded with different variations, including catchy names for each variation? What items seem to be under-performing and what should be considered to increase performance?



To determine which items should be expanded with different variations and which are underperforming, we should look at the sales of individual items. Let's first identify the top performing and bottom performing items.

Finished working

Show work 

Here are the top 5 best-selling items:

1. Coffee: 5471 transactions
2. Bread: 3325 transactions
3. Tea: 1435 transactions
4. Cake: 1025 transactions
5. Pastry: 856 transactions

These items are very popular and could be expanded with different variations. For example:

- Coffee: Could introduce variations like "Caramel Dream Coffee", "Vanilla Sunrise Coffee", or "Spicy Mocha Delight".
- Bread: Could introduce variations like "Sundried Tomato Focaccia", "Multi-Seed Health Loaf", or "Caramelized Onion Brioche".
- Tea: Could introduce variations like "Lavender Bliss Tea", "Spicy Chai Warmth", or "Minty Green Revive".
- Cake: Could introduce variations like "Decadent Dark Chocolate Cake", "Summer Berry Delight Cake", or "Caramel Crunch Cake".
- Pastry: Could introduce variations like "Apple Cinnamon Turnover", "Chocolate Almond Croissant", or "Cherry Cream Cheese Danish".

Z1

What menu items should be expanded with different variations, including catchy names for each variation? What items seem to be under-performing and what should be considered to increase performance?

On the other hand, here are the 5 least popular items, each with only 1 transaction:

- Bacon
- Gift voucher
- Olum & polenta
- Raw bars
- Polenta

These items seem to be underperforming. To increase their performance, the bakery could consider the following strategies:

- **Promotion:** Increase the visibility of these items through in-store promotion, discounts, or bundling them with popular items.
- **Product Improvement:** Evaluate the quality and appeal of these items. Is there something about them that customers don't like? Could the recipe be improved?
- **Customer Feedback:** Gather customer feedback to understand why these items are not selling. Are they priced too high? Are they not to the customers' taste?
- **Discontinuation:** If despite efforts, these items continue to underperform, it might be more profitable to discontinue them and replace them with more appealing items.

Please note that these suggestions are based solely on the number of transactions, and other factors such as cost, pricing, and profitability of each item should be taken into account in the decision-making process.

Z1

Specifying Graph

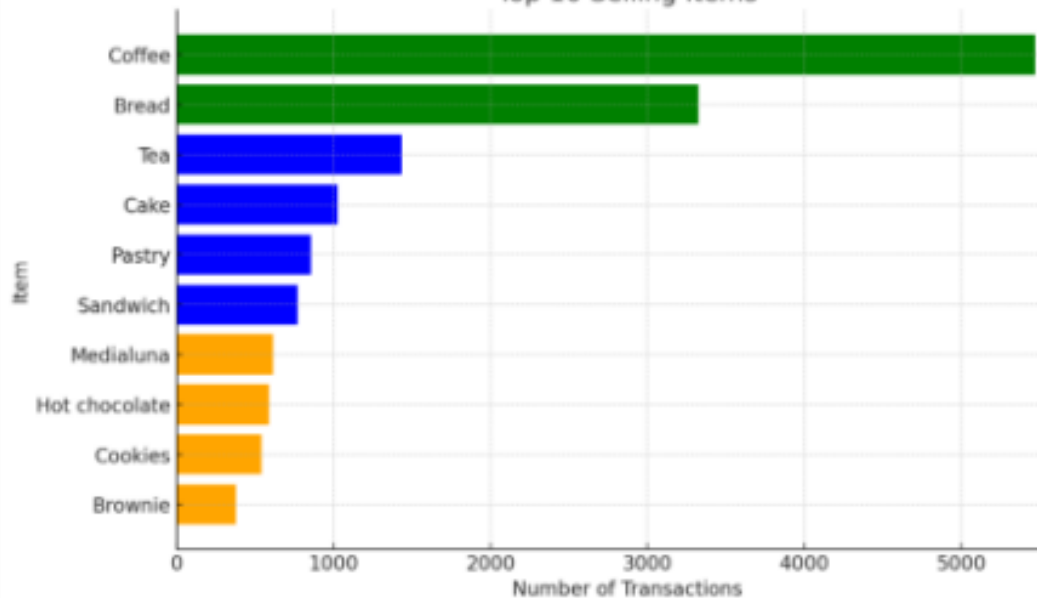
Regroup chart listing drinks first descending by transactions, then desserts. Use the same color schema



Finished working

Show work ▼

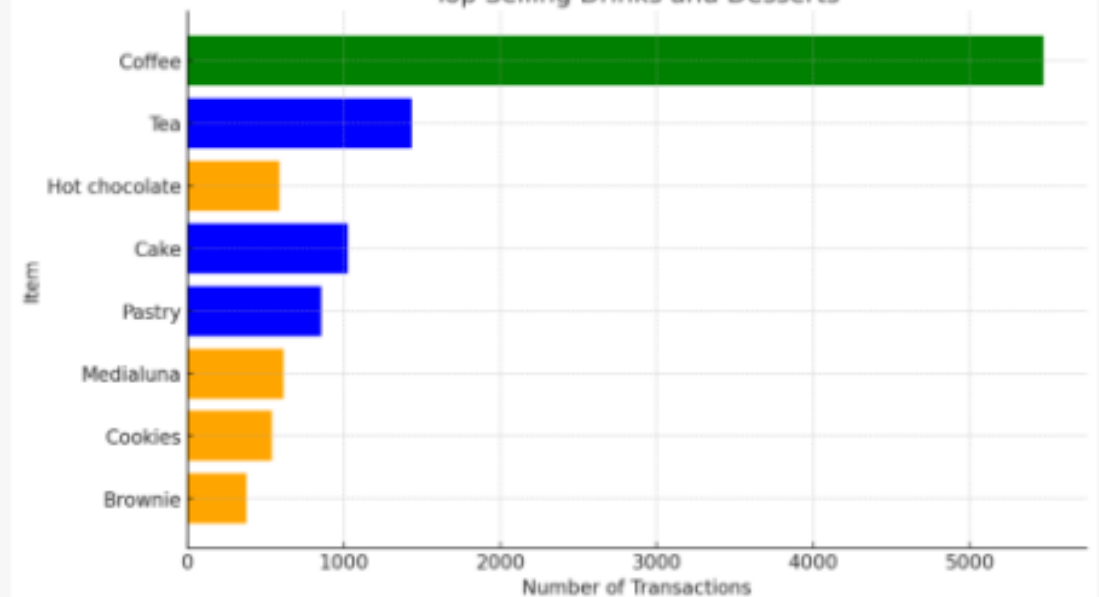
Top 10 Selling Items



Finished working

Show work ▼

Top Selling Drinks and Desserts



Questions?



Jeff Zupancic

Business Consultant, Coach, Facilitator
& Speaker



Scott Lovejoy

AI Consultant, Instructor
& Speaker

248-798-3440

PlaidBridge.com

Making Sense of AI for Small Business™