

Coffee Shop Sales Project

Nehal Ansari

17th December, 2022

Index

| | |
|----------------------------|----|
| Objective | 2 |
| Questions | 2 |
| Data Cleaning and Analysis | 3 |
| Answering the Questions | 10 |
| Measures | 17 |
| Slicres | 17 |
| Final Dashboard | 18 |
| Insights and Conclusion | 19 |
| | |

Objective

The main objective of this project is to analyze retail sales to gain actionable insights that will enhance the performance of the Coffee Shop.

Questions

We will analyze the data and try to extract some useful insights by answering some questions.

- 1) How do sales vary by day of the week and hour of the day?
- 2) Are there any peak times for sales activity?
- 3) What is the total sales revenue for each month?
- 4) How do sales vary across different store locations?
- 5) What is the average price/order per person?
- 6) Which products are the best selling in terms of quantity and revenue?
- 7) How do sales vary by product category and price?

Data Cleaning

We will check for any error in data using filters

Data Analysis

We will analyze data, we will add some more columns into our data which will help us to extract insights easily.

So we use Power Query Terminal to Clean and Analyse data.

From the product_details column we will insert a column having Size Lg, Rg and Sm. Which will make more sense.

Add Conditional Column

Add a conditional column that is computed from the other columns or values.

New column name

Size

| | Column Name | Operator | Value ① | | Output ① |
|---------|----------------|----------|------------|------|-----------------|
| If | product_detail | contains | ABC 123 Lg | Then | ABC 123 Large |
| Else If | product_detail | contains | ABC 123 Rg | Then | ABC 123 Regular |
| Else If | product_detail | contains | ABC 123 Sm | Then | ABC 123 Small |

Add Clause

Else ①

ABC 123 Not Defined

OK

Cancel

Now we will remove Lg,Rg and Sm from the product_detail column because we want a clean product_detail column. So, to remove this we will transform the column product_detail.

Replace Values

Replace one value with another in the selected columns.

Value To Find

Replace With

▸ Advanced options

Replace Values

Replace one value with another in the selected columns.

Value To Find

Replace With

▸ Advanced options

Replace Values

Replace one value with another in the selected columns.

Value To Find

Replace With

▸ Advanced options

Now we will remove white spaces from the product_detail column by using 'Trim' functionality.

| | product_id | unit_price | product_type | product_detail | Size |
|----|------------|------------|-----------------------|-----------------------|-------------|
| 1 | 32 | | Gourmet brewed coffee | Ethiopia | Regular |
| 2 | 57 | | Brewed Chai tea | Spicy Eye Opener Chai | Large |
| 3 | 59 | | Hot chocolate | Dark chocolate | Large |
| 4 | 22 | 2 | Coffee | Drip coffee | Small |
| 5 | 57 | 3.1 | Tea | Brewed Chai tea | Large |
| 6 | 77 | 3 | Bakery | Scone | Not Defined |
| 7 | 22 | 2 | Coffee | Drip coffee | Small |
| 8 | 28 | 2 | Coffee | Gourmet brewed coffee | Small |
| 9 | 39 | 4.25 | Coffee | Barista Espresso | Regular |
| 10 | 58 | 3.5 | Drinking Chocolate | Hot chocolate | Regular |
| 11 | 56 | 2.55 | Tea | Brewed Chai tea | Regular |

Now we have column transaction_qty and unit_price, so we will insert a column for 'total_bill' which will multiply both the columns and then we will change the datatype to currency of total_bill column.

Custom Column

Add a column that is computed from the other columns.

New column name

Total_bill

Custom column formula

= [unit_price]*[transaction_qty]

Available columns

store_location
product_id
transaction_qty
unit_price
product_category
product_type
product_detail
Size

<< Insert

[Learn about Power Query formulas](#)

✓ No syntax errors have been detected.

OK

Cancel

Now we will transform the transaction_time column by removing date and having only time value, so to do that we will use “Text After Delimiter” functionality. We will pass white space in the Delimiter Box.

| | transaction_id | transaction_date | transaction_time | | |
|---|----------------|------------------|---------------------|---|-------|
| 1 | 1 | 01-01-2023 | 31-12-1899 07:14:04 | 5 | Lower |
| 2 | 2 | 01-01-2023 | 31-12-1899 07:20:24 | 5 | Lower |
| 3 | 3 | 01-01-2023 | 31-12-1899 07:22:41 | 5 | Lower |
| 4 | 4 | 01-01-2023 | 31-12-1899 07:22:41 | 5 | Lower |
| 5 | 5 | 01-01-2023 | 31-12-1899 07:25:49 | 5 | Lower |
| 6 | 6 | 01-01-2023 | | 5 | Lower |
| 7 | 7 | 01-01-2023 | | 5 | Lower |

Text After Delimiter

Enter the delimiter that marks the beginning of what you would like to extract.

Delimiter

► Advanced options

OK

Cancel

Page 10 of 10

[illegible]

Transform **Add Column** **View**

Invoke Custom Function

General

Conditional Column

Index Column

Duplicate Column

Format

From Text

Merge Columns

Extract

Parse

Statistics

Standard Scientific

Rounding

Information

Date

Time

Duration

| | transaction_id | transaction_date | transaction_time |
|----|----------------|------------------|------------------|
| 1 | 1 | 01-01-2023 | 07:06:11 |
| 2 | 2 | 01-01-2023 | 07:08:56 |
| 3 | 3 | 01-01-2023 | 07:14:04 |
| 4 | 4 | 01-01-2023 | 07:20:24 |
| 5 | 5 | 01-01-2023 | 07:22:41 |
| 6 | 6 | 01-01-2023 | 07:22:41 |
| 7 | 7 | 01-01-2023 | 07:25:49 |
| 8 | 8 | 01-01-2023 | 07:33:34 |
| 9 | 9 | 01-01-2023 | 07:39:13 |
| 10 | 10 | 01-01-2023 | 07:39:34 |
| 11 | 11 | 01-01-2023 | 07:43:05 |

Also we will extract Hours from the transaction_time column.

The screenshot shows the Tableau Desktop interface. The 'Time' menu is open, and 'Hour' is selected. The background table has the following data:

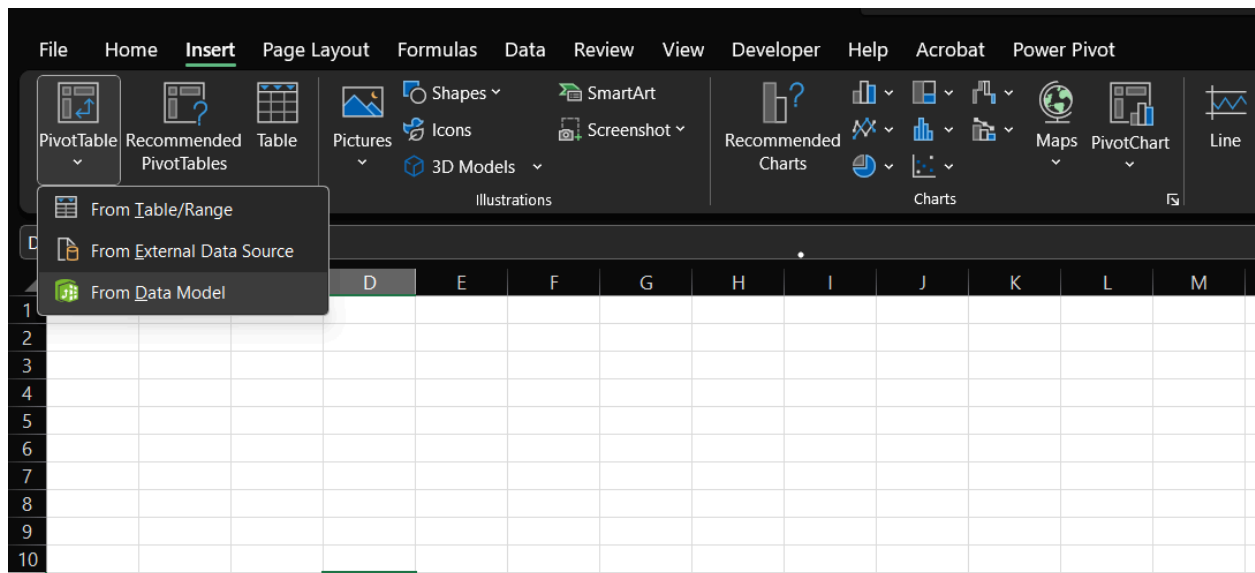
| transaction_id | transaction_date | transaction_time | store_name |
|----------------|------------------|------------------|-----------------|
| 1 | 01-01-2023 | 07:06:11 | Lower Manhattan |
| 2 | 01-01-2023 | 07:08:56 | Lower Manhattan |
| 3 | 01-01-2023 | 07:14:04 | Lower Manhattan |
| 4 | 01-01-2023 | 07:20:24 | Lower Manhattan |
| 5 | 01-01-2023 | 07:22:41 | Lower Manhattan |
| 6 | 01-01-2023 | 07:22:41 | Lower Manhattan |
| 7 | 01-01-2023 | 07:25:49 | Lower Manhattan |
| 8 | 01-01-2023 | 07:33:34 | Lower Manhattan |

After Adding 3 columns, Month Name, Day Name and Hour.

| | product_detail | Size | Total_bill | Month Name | Day Name | Hour |
|----|---------------------------|-------------|------------|------------|----------|------|
| 1 | Ethiopia | Regular | 6.00 | January | Sunday | 7 |
| 2 | Spicy Eye Opener Chai | Large | 6.20 | January | Sunday | 7 |
| 3 | Dark chocolate | Large | 9.00 | January | Sunday | 7 |
| 4 | Our Old Time Diner Blend | Small | 2.00 | January | Sunday | 7 |
| 5 | Spicy Eye Opener Chai | Large | 6.20 | January | Sunday | 7 |
| 6 | Oatmeal Scone | Not Defined | 3.00 | January | Sunday | 7 |
| 7 | Our Old Time Diner Blend | Small | 2.00 | January | Sunday | 7 |
| 8 | Columbian Medium Roast | Small | 4.00 | January | Sunday | 7 |
| 9 | Latte | Regular | 4.25 | January | Sunday | 7 |
| 10 | Dark chocolate | Regular | 7.00 | January | Sunday | 7 |
| 11 | Spicy Eye Opener Chai | Regular | 2.55 | January | Sunday | 7 |
| 12 | Ethiopia | Large | 7.00 | January | Sunday | 7 |
| 13 | Earl Grey | Large | 3.00 | January | Sunday | 7 |
| 14 | Spicy Eye Opener Chai | Large | 3.10 | January | Sunday | 7 |
| 15 | Ouro Brasileiro shot | Not Defined | 6.00 | January | Sunday | 7 |
| 16 | Serenity Green Tea | Large | 6.00 | January | Sunday | 7 |
| 17 | Jumbo Savory Scone | Not Defined | 3.75 | January | Sunday | 7 |
| 18 | Lemon Grass | Regular | 2.50 | January | Sunday | 8 |
| 19 | Dark chocolate | Large | 9.00 | January | Sunday | 8 |
| 20 | Sustainably Grown Organic | Large | 4.75 | January | Sunday | 8 |
| 21 | Ethiopia | Large | 7.00 | January | Sunday | 8 |
| 22 | Spicy Eye Opener Chai | Regular | 5.10 | January | Sunday | 8 |
| 23 | Hazelnut Biscotti | Not Defined | 3.25 | January | Sunday | 8 |
| 24 | Spicy Eye Opener Chai | Regular | 2.55 | January | Sunday | 8 |
| 25 | Cappuccino | Not Defined | 3.75 | January | Sunday | 8 |
| 26 | Lemon Grass | Large | 3.00 | January | Sunday | 8 |
| 27 | Chocolate Chip Biscotti | Not Defined | 3.50 | January | Sunday | 8 |
| 28 | Peppermint | Large | 6.00 | January | Sunday | 8 |
| 29 | Dark chocolate | Not Defined | 3.75 | January | Sunday | 8 |

Now we will Load our data to the Worksheet and we are done with data cleaning and analysing.

We have loaded data in the worksheet as a 'data model', so when we create pivot tables from our data set we have to select 'From Data Model'.



Answering Questions

Now let's jump into data visualisation and insight extraction.

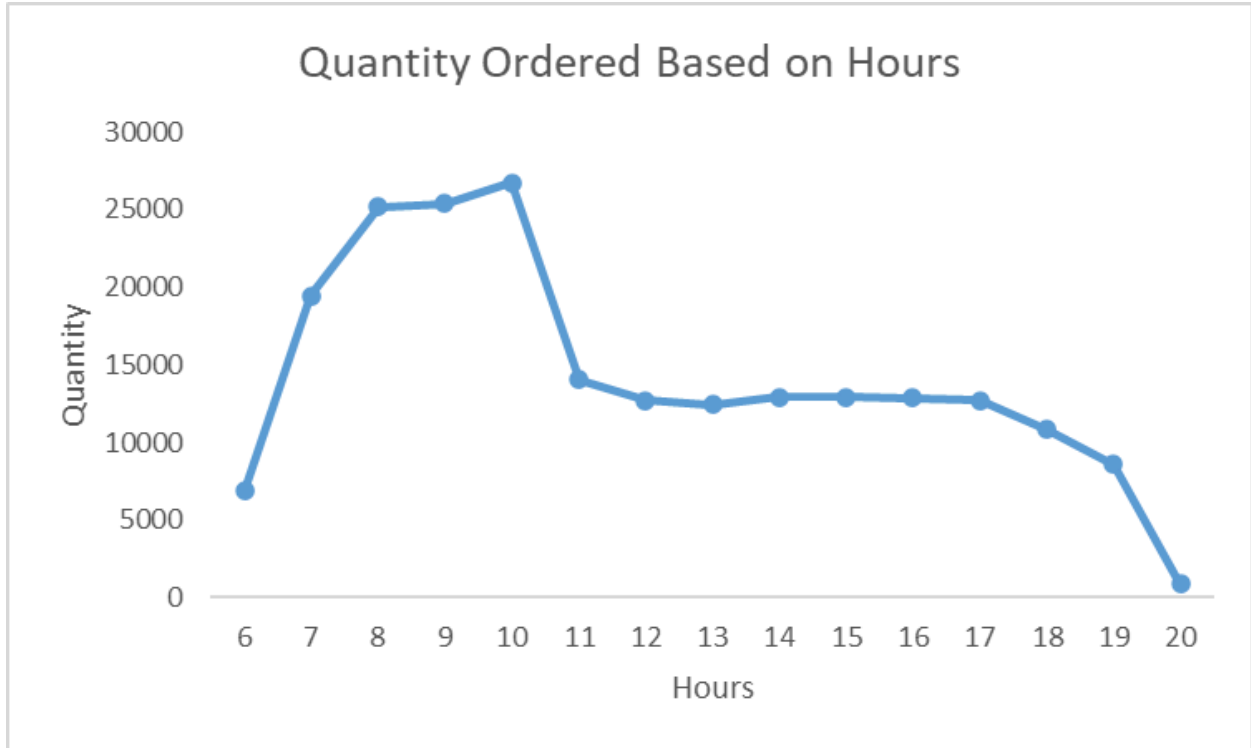
Question 1) How do sales vary by day of the week and hour of the day?

| Row Labels | Sum of transaction_qty |
|--------------------|------------------------|
| 6 | 6865 |
| 7 | 19449 |
| 8 | 25197 |
| 9 | 25370 |
| 10 | 26713 |
| 11 | 14035 |
| 12 | 12690 |
| 13 | 12439 |
| 14 | 12907 |
| 15 | 12923 |
| 16 | 12881 |
| 17 | 12700 |
| 18 | 10826 |
| 19 | 8595 |
| 20 | 880 |
| Grand Total | 214470 |

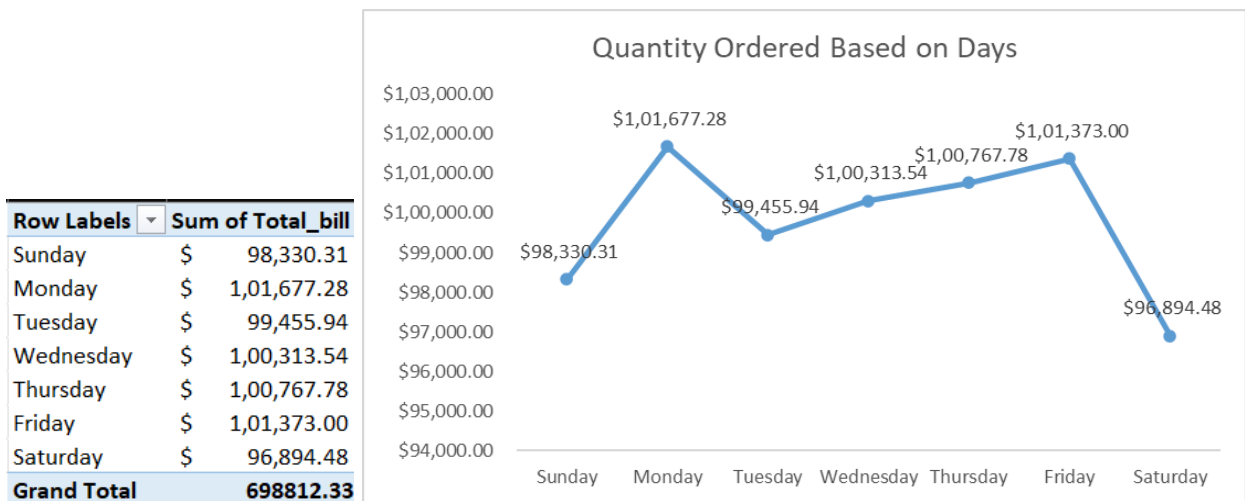
transaction_date
transaction_time
store_id
store_location
product_id
☒ transaction_qty
unit_price
product_category

Drag fields between areas below:

| | |
|----------------|------------------------|
| Filters | Columns |
| | |
| Rows | Values |
| Hour | Sum of transaction_qty |



Here is the - sales ordered based on days

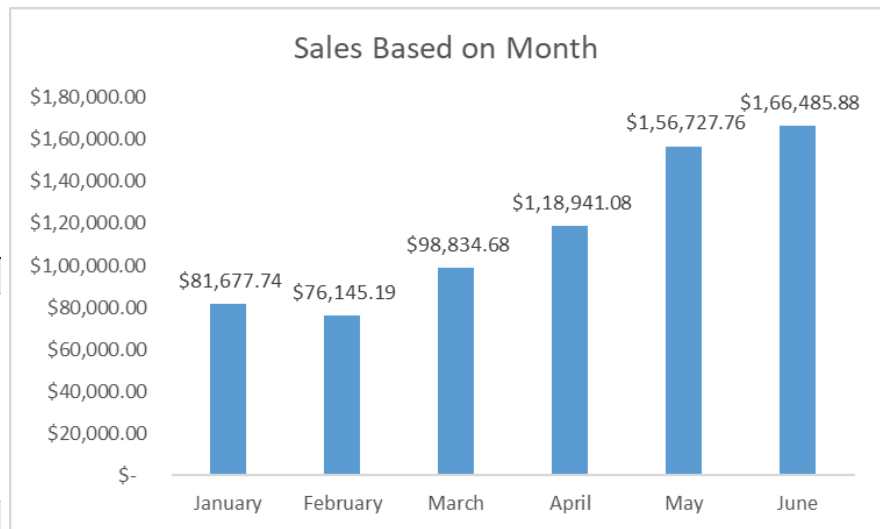


Question 2) Are there any peak times for sales activity?

So, in the chart 'quantity ordered based on Hours' , we can say that peak time for sales are from 9:00 AM to 10:00 AM in the morning.

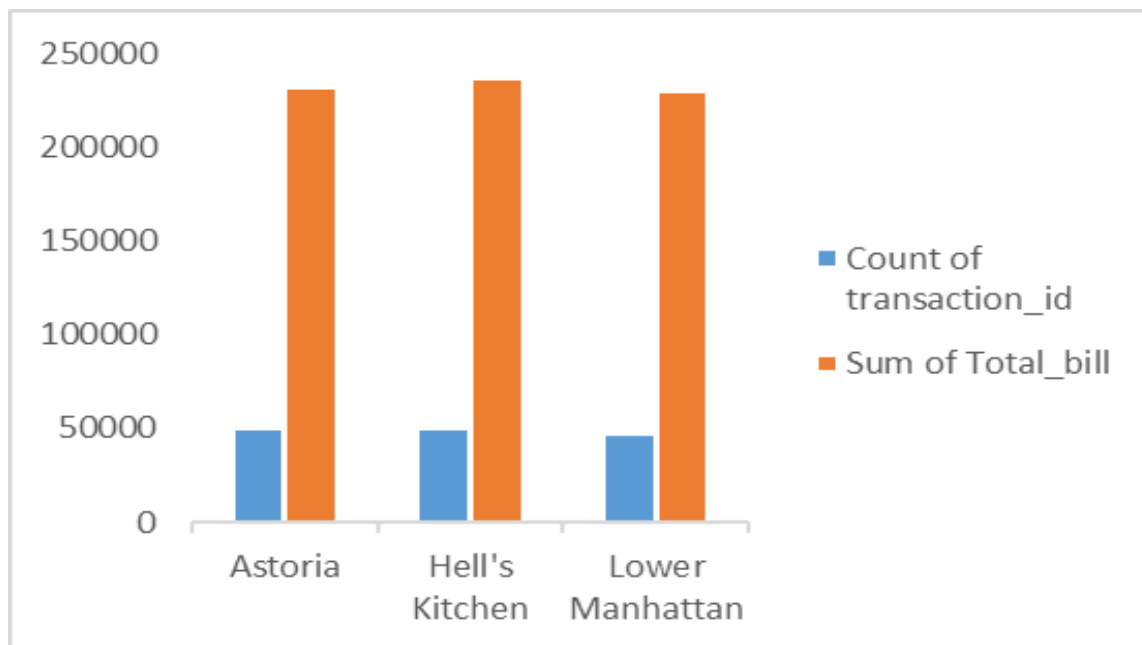
Question 3) What is the total sales revenue for each month?

| Month | Sum of Total_bill |
|--------------------|-------------------|
| January | \$ 81,677.74 |
| February | \$ 76,145.19 |
| March | \$ 98,834.68 |
| April | \$ 1,18,941.08 |
| May | \$ 1,56,727.76 |
| June | \$ 1,66,485.88 |
| Grand Total | 698812.33 |



Question 4) How do sales vary across different store locations?

| Location | Count of transaction_id | Sum of Total_bill |
|--------------------|-------------------------|----------------------|
| Astoria | 50599 | \$ 2,32,243.91 |
| Hell's Kitchen | 50735 | \$ 2,36,511.17 |
| Lower Manhattan | 47782 | \$ 2,30,057.25 |
| Grand Total | 149116 | \$6,98,812.33 |

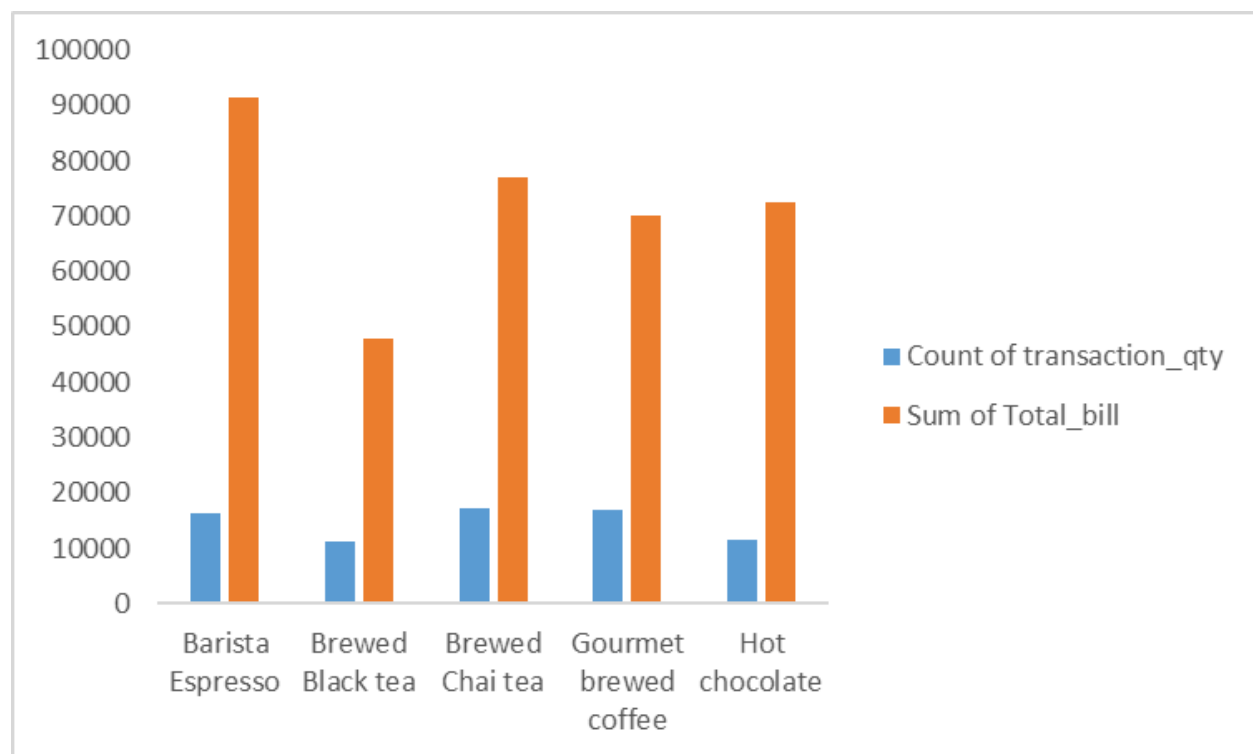


Question 5) What is the average price/order per person?

It will be done by Measures and displayed on the dashboard.

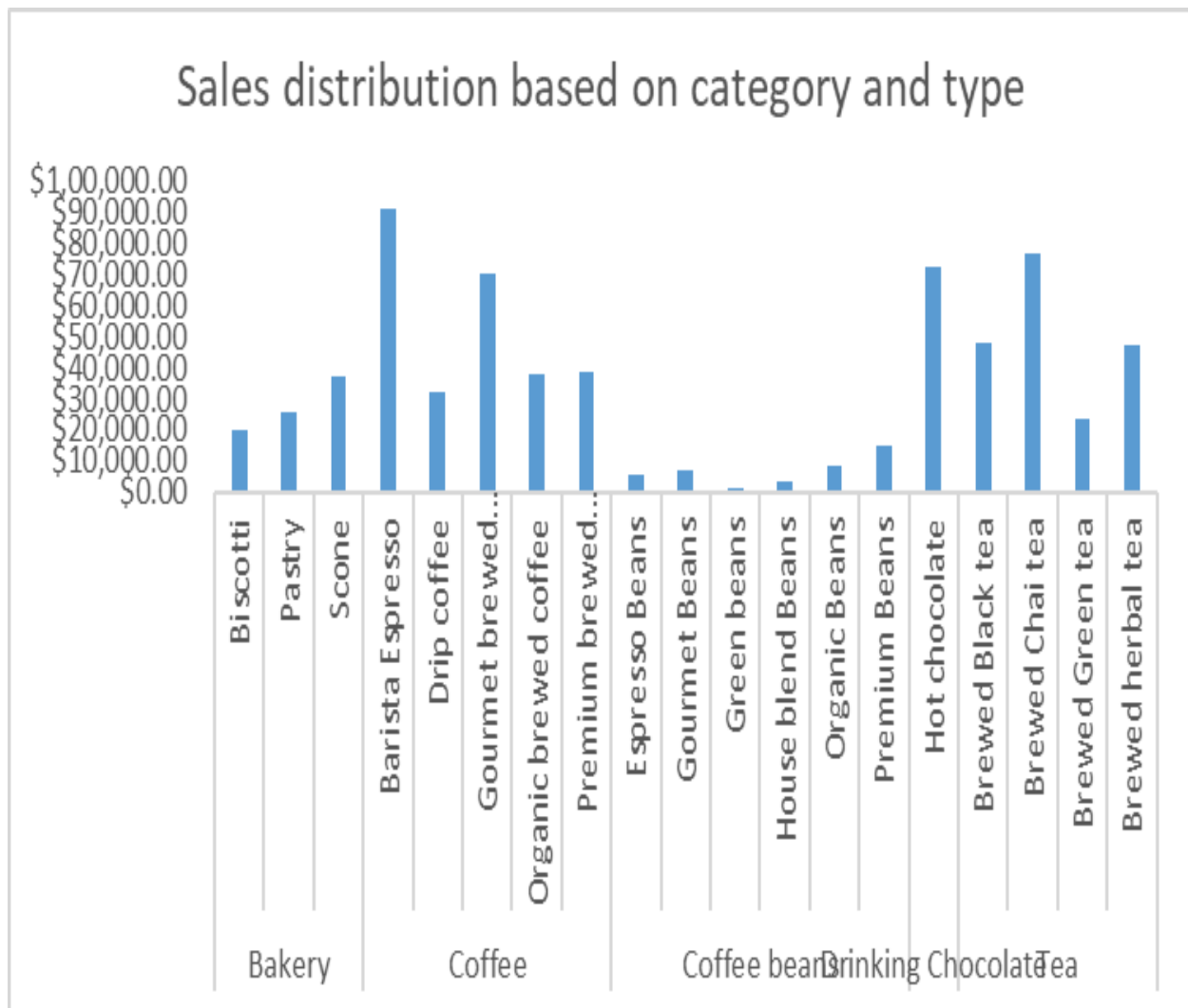
Question 6) Which products are best selling in terms of quantity and revenue? (Top 5)

| Top 5 Product | Count of transaction_qty | Sum of Total_bill |
|-----------------------|--------------------------|----------------------|
| Barista Espresso | 16403 | \$ 91,406.20 |
| Brewed Black tea | 11350 | \$ 47,932.00 |
| Brewed Chai tea | 17183 | \$ 77,081.95 |
| Gourmet brewed coffee | 16912 | \$ 70,034.60 |
| Hot chocolate | 11468 | \$ 72,416.00 |
| Grand Total | 73316 | \$3,58,870.75 |



Question 7) How do sales vary by product category and type?

| product_category(Top 5) | product_type | Sum of Total_bill |
|---------------------------|-----------------------|----------------------|
| Bakery | | |
| | Biscotti | \$19,793.53 |
| | Pastry | \$25,655.99 |
| | Scone | \$36,866.12 |
| Coffee | | |
| | Barista Espresso | \$91,406.20 |
| | Drip coffee | \$31,984.00 |
| | Gourmet brewed coffee | \$70,034.60 |
| | Organic brewed coffee | \$37,746.50 |
| | Premium brewed coffee | \$38,781.15 |
| Coffee beans | | |
| | Espresso Beans | \$5,560.25 |
| | Gourmet Beans | \$6,798.00 |
| | Green beans | \$1,340.00 |
| | House blend Beans | \$3,294.00 |
| | Organic Beans | \$8,509.50 |
| | Premium Beans | \$14,583.50 |
| Drinking Chocolate | | |
| | Hot chocolate | \$72,416.00 |
| Tea | | |
| | Brewed Black tea | \$47,932.00 |
| | Brewed Chai tea | \$77,081.95 |
| | Brewed Green tea | \$23,852.50 |
| | Brewed herbal tea | \$47,539.50 |
| Grand Total | | \$6,61,175.29 |



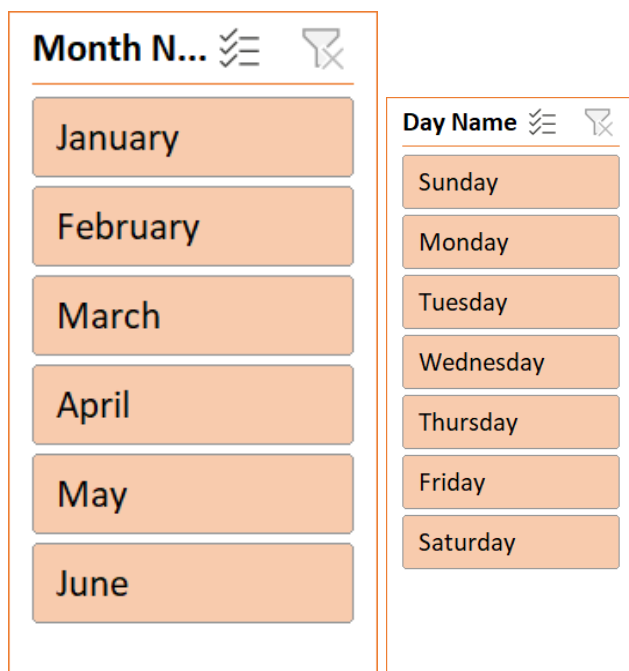
Measures

We need to display total sales, total footfall etc in the dashboard. So, to do it we will create 'measures'.

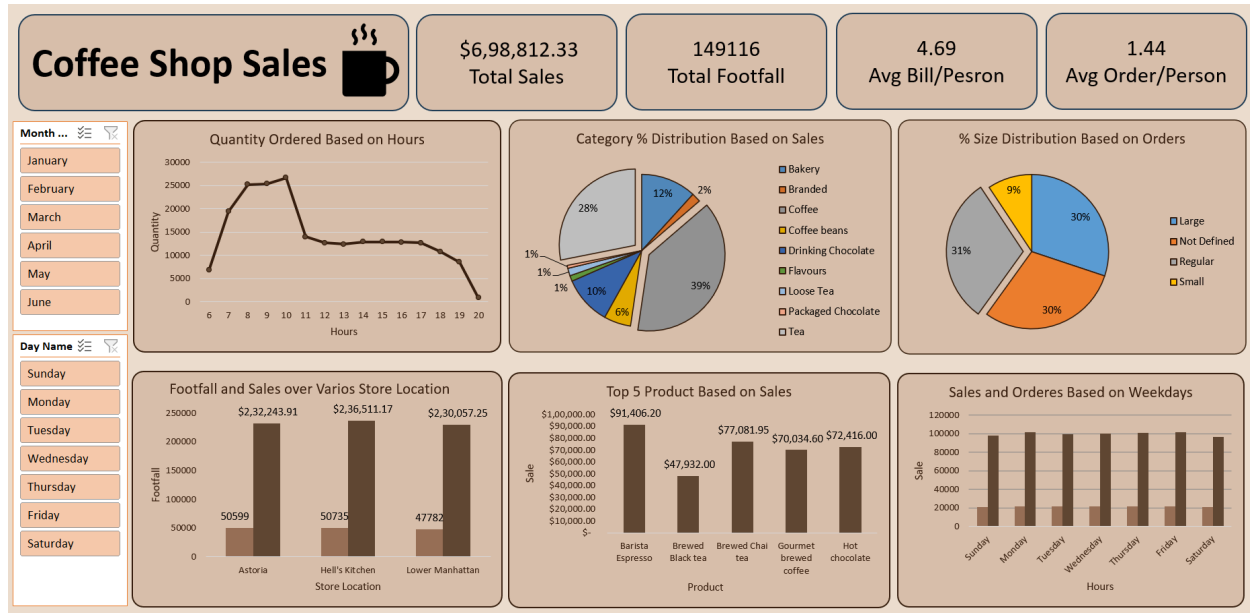


Slicers

We will add some slicers to make our dashboard more interactive, we can sort the charts based on Months and Days.




Final Dashboard



Insights

So we are done with the final report, we have cleaned, analysed and visualised the data, now we are ready to extract some useful insights and answer to our questions.

- 1) There are more sales from 9:00 AM to 10:00 AM. So we can focus more at this time in terms of product quality and service.
- 2) In terms of category, we have more sales of Coffee followed by Tea.
- 3) If we talk about size of product, then Regular size is sold more and Small size is Less.
- 4) Store, Hell's Kitchen, has generated the highest revenue.
- 5) Barista Espresso is the highest selling product among all the products.



6) Monday and Friday are the top two days for high selling.