**Coffee shop**

Maven Roasters is a coffee shop with three branches located in **Hell's Kitchen**, **Lower Manhattan**, and **Astoria**, NYC. Over the past year, the shop has been serving a variety of coffee products, striving to provide customers with high-quality beverages and a delightful experience. This dataset, encompassing 6 month of transaction records, provides insights into sales trends, customer behavior, and operational performance across the branches.

**Data Understanding**

**Dataset summery**

* **transaction\_id** : Unique sequential ID representing an individual transaction
* **transaction\_date** : Date of the transaction (DD/MM/YY)
* **transaction\_time** : Timestamp of the transaction (HH:MM:SS)
* **transaction\_qty** : Quantity of items sold
* **store\_id** : Unique ID of the coffee shop where the transaction took place
* **store\_location** : Location of the coffee shop where the transaction took place
* **product\_id** : Unique ID of the product sold
* **unit\_price** : Retail price of the product sold
* **product\_category** : Description of the product category
* **product\_type** : Description of the product type
* **product\_detail** : Description of the product detail

**About data :**

* **Transaction\_id** : Contain 149116 unique transaction record
* **Transaction\_date** : Contain 6 month data in 2023
* **Transaction\_time**  : shop opeining time 6:00 AM to 10:00 PM
* **Store\_id** : Hear 3 store in New York City
* **Store\_location** : store location is Hell's Kitchen, Lower Manhattan, and Astoria
* **Product\_category**  : Here sale 9 different category in our shop
* **Product\_type** : Here 87 type of different product available

**Data Identify:**

* **Category column :** 3 (store\_location , product\_category , product\_type )
* **Numeric column :** 5 (transaction\_id ,transaction\_qty ,store\_id , product\_id ,unit\_price )
* **Text column :** 1 (product\_detail )
* **Time series column :** 2 (transaction\_date ,transaction\_time )

**Data Cleaning**

* **Missing value** : Here 3 data is missing in product type
* **Filling missing value** : We are trying to identify the product type using the Product ID for "Gourmet brewed coffee." However, there is an issue: the same product type, "Gourmet brewed coffee," is associated with two different Product IDs (30 and 32). To resolve this, we will compare the product type and Product ID with the product details, and then correct any missing or inconsistent values.
* **Miss spall data** : The store location “Hell's Kitchen, Lower Manhattan” is misspelled as “Hell's Kitcen, Lower Manattan.” To address this issue, we need to correct the spelling errors and clean the data to ensure consistency.
* **Data Types**: Confirm that transaction\_date and transaction\_time are in the correct date and time formats for analysis.
* **Duplicates**: transaction\_id is unique and there are no duplicate rows.

**Basic statistics Analysis:**

Before starting the statistical analysis, we need to calculate the total sales by multiplying the Quantity column and the Unit\_Price column.

* **Total row :** 149116
* **Total column :** 11
* **Total Sale :** 698812.33
* **Min Sale :** 0 .8
* **Avarage Sale :** 4.67
* **Max Sale :** 360

**Visual Analysis:**

**1)Sales trend :**

Before analysis create month and week day column for sales trend analysis

**Month analysis:**

-There is a clear upward trend in sales from January to June, indicating strong growth over time.

- June recorded the highest sales, showing peak performance.

- February had the lowest sales , suggesting a possible dip in demand or performance during this month.

- A steep increase is noticeable between April and May

- Except for February, all months show consistent growth

**2) Week day and time analysis:**

- Morning has the highest sales, contributing significantly to the total revenue.

- Monday and Friday show the highest sales.

- Peak sales occur during the Morning across most weekdays, particularly on Fridays and Thursdays.

- Evenings and Nights contribute the least, indicating these might not be ideal times for promotional efforts.

-Offer discounts or promotions to boost evening and night sales.

**3) Location analysis:**

-The distribution of sales is nearly equal across the three locations, with less than 1% variation between them.

-This balance suggests consistent sales strategies and customer demand across the locations.

**4) Category:**

-Coffee dominates sales, contributing the most revenue. This indicates strong consumer demand and preference for coffee-related products.

- Bakery and Drinking Chocolate can benefit from combo deals and seasonal promotions.

-Branded, Flavours, and Packaged Chocolate need better marketing or product adjustments.

-Use promotions and loyalty programs to drive sales for coffee, tea, and bakery items.

**5) Top 10 product:**

-Maximize Revenue from Top Performers

-Pair Scones with top-selling drinks (e.g., "Espresso + Scone" combo deals).

-Highlight the health benefits and unique qualities of Organic Brewed Coffee in marketing campaigns.

-Focus on unique flavors or premium offerings in brewed coffee and tea categories to stand out from competitors.