## **About Dataset**

The following dataset is generated using Python library named Faker and contains sales data of Soft Drink sold in the year 2022-2023.

The data can be used for a comprehensive analysis on overall performance and trends in soft drink sales, including total sales, number of transaction, average bill per person, and average orders per person? What are the sales distributions by product type, product category, store location, time of day, and weekdays, and how can these insights be analyzed for specific months and days

# **Dataset Columns**

|  |  |
| --- | --- |
| Fields | Description |
| transaction\_id | Unique sequential ID representing an individual transaction |
| transaction\_date | Date of the transaction (MM:DD:YY) |
| transaction\_time | Timestamp of the transaction(HH:MM:SS) |
| transaction\_qty | Quantity of items sold |
| store\_id | Unique ID of the Soft Drink shop where the transaction took place |
| store\_location | Location of the Soft drink 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\_name | Description of the product name |

# **Problem Statement**

**Sales Trends:** Analyze the trend of soft drink sales over time. Are there certain months or seasons with higher sales? Are there any noticeable trends in specific hours or days?

**Popular Products and Brands:** Determine the most popular soft drink products and brands by sales. Which products have the highest sales volume? Which brands dominate the market?

**Demand by Product Type:** Identify the most in-demand product types across different categories of soft drinks. Are there any emerging product types that are increasingly sought after by consumers?

**Geographical Analysis:** Explore the distribution of soft drink sales across various store locations. Which locations have the highest number of transactions? Are there any regional disparities in sales?

**Time of Purchase Analysis:** Investigate the time of purchase patterns for soft drinks. Are there specific hours of the day or days of the week with higher sales? How do these patterns vary across different products?

**Category Sales Distribution:** Compare the distribution of sales across different soft drink categories (e.g., Carbonated Soft Drinks, Dairy and Plant-Based Beverages, Energy and Sports Drinks). Are there any categories that are particularly popular?

**Order Quantity Analysis:** Examine the factors influencing the average order quantity per person. Are there specific products, categories, or store locations that result in larger orders?

**Recent Sales Analysis:** Explore the characteristics of recent sales. Do they differ from older sales in terms of product preferences, purchase time, or order quantity? How does recent sales data compare to historical trends?