**DAILY SALES TRANSACTION FOR COFFEE ACROSS NEWYORK CITY (2023)**

This dataset captures detailed sales transactions from multiple locations in New York City, specifically Astoria, Lower Manhattan and Hell's Kitchen, during January-June 2023. It provides a comprehensive record of purchases across various product categories, including beverages like coffee, tea, and drinking chocolate, as well as bakery items such as scones and biscotti.

**DATA SOURCE**

The dataset was sourced through a structured survey conducted across selected store locations in New York City, including Astoria, Lower Manhattan and Hell's Kitchen. The survey results were documented in an Excel spreadsheet before it was imported into Power BI for advanced data cleaning, transformation, visualization and analysis. The transition from Excel to Power BI allowed for more dynamic insights, enabling the identification of sales trends, customer behavior patterns, and performance differences between store locations.

**Tools:**

1. **Data Sourcing and Importing:**
   * Successfully importing raw data from Excel into Power BI, ensuring a seamless transition for further analysis.
2. **Data Cleaning and Transformation using DAX:**
   * Utilizing DAX (Data Analysis Expressions) functions within Power BI to perform essential data cleaning and transformation tasks.
   * Adding new calculated columns for metrics such as revenue, month, day, and time extraction, which are crucial for detailed analysis.
   * Ensuring data accuracy and consistency through these transformations, preparing the dataset for insightful reporting.
3. **Measure Creation and Calculation:**
   * Developing new measures using DAX to calculate key performance indicators, including total revenue and total quantity sold.
   * Demonstrating the ability to distinguish between explicit measures (calculated fields) and implicit measures (automatically generated by Power BI) to tailor analysis to specific needs.
4. **Data Visualization:**
   * Designing and implementing visualizations that effectively communicate insights derived from the dataset.
   * Using both explicit and implicit measures within visualizations to provide a comprehensive view of the data, making it easier to identify trends and patterns.
5. **Analytical Skills:**
   * Applying advanced analytical techniques to interpret the visualized data, uncovering trends, and generating actionable insights.
   * Demonstrating the ability to make data-driven decisions based on the analysis performed in Power BI.
6. **Proficiency in Power BI:**
   * Showcasing deep proficiency in Power BI, from data importation and transformation to visualization and analysis.
   * Effectively leveraging Power BI’s capabilities to handle complex datasets and extract meaningful insights.

#upload analysis#

### Data Cleaning and Transformation

The process of data cleaning and transformation is crucial to ensure that the dataset is accurate, consistent, and ready for analysis. For the given dataset, which was imported from Excel into Power BI, the following steps outline the cleaning and transformation process:

1. **Data Cleaning:**

* **Handling Missing Values:**
  + Review the dataset for any missing or null values in key columns, such as transaction date, transaction qty, product id, and unit price.
* **Data Type Correction:**
  + Ensure that each column has the correct data type. For instance, transaction date should be formatted as a date, transaction qty as an integer, and unit price as a decimal. This is essential for accurate calculations and analyses.
* **Removing Duplicates:**
  + Check for and remove any duplicate records that may distort the analysis. Duplicate entries can occur due to errors in data entry or processing.
* **Consistency Checks:**
  + Ensure consistency in data entries, especially in categorical columns like store location and product type. For example, ensure that all instances of "Lower Manhattan" are spelled the same way and have the same case format.

2. **Data Transformation:**

* **Adding Calculated Columns:**
  + Use DAX functions in Power BI to create new columns that provide additional insights:
    - **Revenue Calculation:** Create a Revenue column that multiplies transaction\_qty by unit\_price to determine the total revenue for each transaction.

Revenue = [transaction\_qty] \* [unit\_price]

* **Extracting Date Components:** Create new columns for Month, Day, and Time by extracting these components from the transaction\_date column:

Month Number = MONTH([transaction\_date])

Day of the week = WEEKDAY([transaction\_date])

Transaction hour = Hour([transaction\_time])

* **Converting Month Number to Month Name:**

Convert the MonthNumber into a MonthName using the FORMAT() function:

MonthName = FORMAT([transaction\_date], "MMMM")

* **Creating New Measures:**

Develop new measures to calculate key metrics such as Total Revenue and Total Quantity sold:

Total Revenue = SUMX(Sales, [transaction\_qty] \* [unit\_price])

Total Quantity = SUM(Sales[transaction\_qty])

* **Standardizing Formats and Units:**
  + Ensure that all currency values are consistently formatted and that units (such as quantity) are standardized across the dataset.

**3. Implementation of Explicit and Implicit Measures:**

* **Explicit Measures:**
  + Define explicit measures using DAX to perform specific calculations that will be used across various visualizations. These measures provide consistent results across different contexts within Power BI.
* **Implicit Measures:**
  + Use implicit measures, which are automatically generated by Power BI when dragging and dropping fields into visualizations. These are useful for quick calculations but should be verified for consistency.

By performing these data cleaning and transformation steps, the dataset becomes more reliable and ready for insightful analysis. The clean and well-structured data facilitates the creation of accurate visualizations and supports informed decision-making based on the analysis conducted in Power BI. This preparation also ensures that subsequent analysis, whether it be sales trends, customer behavior, or performance metrics, is built on a solid foundation.

#Upload the raw file and transformed file#

**DATA DESCRPTION**

The dataset represents detailed transaction records from three store locations, Astoria, Lower Manhattan and Hell's Kitchen, in New York City. It captures sales data for various beverage and bakery products over a specific period, including critical information such as transaction details, product characteristics, and pricing.

Here's an overview of the key columns in the dataset:

1. **transaction\_id**:
   * Description*: A unique identifier for each transaction.*
   * Type*: Integer*
   * Example*: 1, 2, 3*
2. **transaction\_date**:
   * Description: The date when the transaction occurred.
   * Type: Date
   * Example: 1/1/2023, 1/2/2023
3. **transaction\_time**:
   * Description: The specific time at which the transaction was made.
   * Type: Time
   * Example: 7:06:11, 8:17:27
4. **transaction\_qty**:
   * Description: The quantity of items purchased in the transaction.
   * Type: Integer
   * Example: 1, 2, 3
5. **store\_id**:
   * Description: A unique identifier for the store where the transaction occurred.
   * Type: Integer
   * Example: 5, 8
6. **store\_location**:
   * Description: The geographical location of the store where the transaction took place.
   * Type: Text
   * Example: Astoria, Lower Manhattan, Hell's Kitchen
7. **product\_id**:
   * Description: A unique identifier for each product sold.
   * Type: Integer
   * Example: 32, 57, 59
8. **unit\_price**:
   * Description: The price per unit of the product sold.
   * Type: Decimal
   * Example: 3.00, 4.50, 2.55
9. **product\_category**:
   * Description: The category to which the product belongs, such as Coffee, Tea, Drinking Chocolate, or Bakery.
   * Type: Text
   * Example: Coffee, Tea, Bakery
10. **product\_type**:
    * Description: A more specific classification of the product within its category.
    * Type: Text
    * Example: Gourmet brewed coffee, Brewed Chai tea, Hot chocolate, Scone
11. **product\_detail**:
    * Description: Additional details about the product, such as specific blend names or size.
    * Type: Text
    * Example: Ethiopia Rg, Spicy Eye Opener Chai Lg, Dark chocolate Lg

**Problem Statements**

1. **Sales Performance Analysis:**
   * How do sales trends for different product categories (e.g., coffee, tea, bakery items) vary across different store locations in New York City during January-June 2023?
   * Which store location—Astoria, Lower Manhattan or Hell's Kitchen—generated the highest revenue, and what factors contributed to this performance?
2. **Customer Behavior Insights:**
   * What are the peak transaction times for each store location, and how does this influence inventory management and staffing needs?
   * What is the average spending per transaction for different product categories, and how does it vary between locations?
3. **Product Popularity and Demand Forecasting:**
   * Which specific products (e.g., Ethiopia Lg, Spicy Eye Opener Chai Lg) are the most popular among customers, and how can this data be used to forecast future demand?
   * How does the pricing of various products affect their sales volume across different store locations?
4. **Operational Efficiency:**
   * Are there inefficiencies in transaction processing times that could be addressed to improve customer experience and store throughput?
   * How can the store locations optimize their product offerings based on the observed sales data to better meet customer demand?
5. **Revenue Optimization:**
   * What is the impact of transaction quantity on overall revenue, and how can this information be used to design promotional strategies that encourage bulk purchasing?
   * How does the pricing strategy for different product types affect the store's profitability, and what adjustments could be made to maximize revenue?

### Findings/Results

Analyzing this dataset could reveal several key insights about sales trends, customer behavior, and store performance. Here are some potential findings and results:

1. **Top-Selling Products:**
   * Identify which products are the most popular based on total units sold.
   * finding: "Gourmet brewed coffee" and "Hot chocolate" are among the top-selling products across both store locations.
2. **Revenue Distribution by Product Category:**
   * Analyze the revenue contribution of different product categories such as Coffee, Tea, Drinking Chocolate, and Bakery.
   * Possible finding: "Coffee" generates the highest revenue, accounting for 45% of total sales, followed by "Tea".
3. **Peak Sales Times:**
   * Determine the times of day when sales are highest by analyzing the transaction\_time data.
   * Possible finding: Sales peak during the morning hours between 7:00 AM and 9:00 AM, indicating a strong morning rush.
4. **Store Performance Comparison:**
   * Compare the performance of the three store locations (Astoria, Lower Manhattan and Hell's Kitchen) in terms of total revenue and units sold.
   * Possible finding: Hell's Kitchen have the highest performance followed by Astoria while Lower Manhattan performed lest.
5. **Customer Preferences by Location:**
   * Examine product preferences based on store location to identify any regional differences in consumer behavior.
   * Possible finding: Customers prefer coffee more in the three locations.
6. **Seasonal Sales Trends:**
   * If the data covers multiple months, analyze sales trends to identify seasonal patterns or monthly fluctuations in sales.
   * Possible finding: Sales increase by 15% during the winter months, with "Hot chocolate" and "Gourmet brewed coffee" being particularly popular.
7. **Product Pairing Trends:**
   * Investigate common product pairings to identify bundles or cross-selling opportunities.
   * Possible finding: "Scone" is frequently purchased with "Brewed Chai tea", suggesting a strong product pairing trend.
8. **Impact of Discounts or Promotions:**
   * If applicable, assess the impact of any discounts or promotions on sales volume and revenue.
   * Possible finding: Promotional pricing on "Brewed herbal tea" led to a 25% increase in units sold, but with a slight reduction in overall profit margin.
9. **Customer Loyalty and Repeat Purchases:**
   * Analyze whether certain products lead to repeat purchases or higher customer loyalty.
   * Possible finding: Customers who purchase "Barista Espresso" tend to return more frequently, indicating a potential loyalty driver.

**RECOMMENDATION**

Based on the analysis of the dataset, several strategic recommendations can be made to improve sales, enhance customer experience, and optimize store performance.

1. **Enhance Product Offerings:**
   * **Recommendation:** Focus on expanding the range of top-selling products, such as "Gourmet brewed coffee" and "Hot chocolate," to meet high demand. Consider introducing new flavors or variations of these popular items to attract repeat customers.
   * **Rationale:** Capitalizing on popular products can drive increased revenue and customer satisfaction.
2. **Optimize Store Operations During Peak Hours:**
   * **Recommendation:** Increase staff and resources during peak sales hours, particularly in the morning between 7:00 AM and 9:00 AM, to ensure quick service and reduce wait times.
   * **Rationale:** Efficient operations during peak times can enhance customer experience and potentially boost sales.
3. **Tailor Marketing Strategies by Location:**
   * **Recommendation:** Implement location-specific marketing strategies that cater to the preferences of customers in each area. For example, promote "Tea" and "Bakery" items more heavily in Lower Manhattan, while emphasizing "Coffee" offerings in Hell's Kitchen.
   * **Rationale:** Customized marketing can better resonate with local customers and drive higher engagement and sales.
4. **Bundle Popular Product Pairs:**
   * **Recommendation:** Introduce bundled deals or combo offers for popular product pairings, such as "Brewed Chai tea" and "Scone." This could include discounts for purchasing these items together.
   * **Rationale:** Bundling can increase average transaction value and encourage customers to try additional products.
5. **Leverage Seasonal Trends:**
   * **Recommendation:** Prepare for seasonal demand spikes, particularly in the winter months, by stocking up on popular items like "Hot chocolate" and "Gourmet brewed coffee." Additionally, consider launching seasonal promotions or limited-time offers to attract customers.
   * **Rationale:** Aligning product availability and marketing with seasonal trends can maximize revenue during peak periods.
6. **Focus on Customer Retention:**
   * **Recommendation:** Implement a loyalty program or reward system to encourage repeat purchases, especially for high-loyalty products like "Barista Espresso." Offering discounts or free items after a certain number of purchases could incentivize return visits.
   * **Rationale:** Retaining existing customers is often more cost-effective than acquiring new ones, and loyalty programs can build long-term customer relationships.
7. **Explore New Product Categories:**
   * **Recommendation:** Consider expanding into related product categories or offering new items based on customer preferences and market trends. For example, introducing premium or organic options could cater to health-conscious consumers.
   * **Rationale:** Diversifying the product portfolio can attract a broader customer base and create additional revenue streams.
8. **Refine Pricing Strategies:**
   * **Recommendation:** Regularly review and adjust pricing strategies to remain competitive while maintaining healthy profit margins. Consider dynamic pricing models or promotional discounts during off-peak times to boost sales.
   * **Rationale:** A well-balanced pricing strategy can attract price-sensitive customers without sacrificing profitability.
9. **Invest in Training for Staff:**
   * **Recommendation:** Provide ongoing training for staff to enhance their product knowledge, customer service skills, and efficiency, especially during peak times. Well-informed and courteous staff can significantly enhance the customer experience.
   * **Rationale:** Improved staff performance can lead to better customer satisfaction and loyalty, ultimately driving repeat business.
10. **Monitor and Adapt to Customer Feedback:**
    * **Recommendation:** Actively collect and analyze customer feedback to identify areas for improvement in product offerings, service, and store environment. Use this feedback to make data-driven decisions and continuously adapt to customer needs.
    * **Rationale:** Responsive and adaptive business practices can foster strong customer relationships and keep the business aligned with market demands.

### Conclusion

Implementing these recommendations can help optimize operations, boost sales, and enhance overall customer satisfaction. By aligning strategies with data-driven insights, the business can achieve sustainable growth and maintain a competitive edge in the market.