**Project - Phase I: Planning**

**Student Name:**Ashish Bhaskar Bandarkar

**Section:** - IFT 598 Data Visualization & Reporting for IT

**Section:** 2023Spring-P-IFT598-29280

**Instructor Name**: Prof. Asmaa Elbadrawy

**Due Date**: 9th April 2023

Supermarket expansion and market competitiveness are both on the rise in the majority of populous cities. The dataset contains historical sales information from a grocery chain that was captured over a three-month period at 3 distinct branches. The dataset is separated into 17 characteristics and gives information on each transaction performed by clients. For additional information, please see the table below.

|  |  |  |  |
| --- | --- | --- | --- |
| Attribute | Data type | Description | Domain |
| Invoice id | Categorical | An identification number assigned to an invoice after a successful sale | Range(number) |
| Branch | Categorical | The branch of the supercenter where the sale was made | List (string) |
| City | Categorical | The location of the supercenter where the sale was made | List (string) |
| Customer type | Categorical | The type of customer who made the purchase, recorded as either "Member" for customers using a member card, or "Normal" for customers without a member card | List (string) |
| Gender | Categorical | The gender of the customer who made the purchase | List (string) |
| Product line | Categorical | The general category of the item(s) purchased, such as "Electronic accessories," "Fashion accessories," etc. | List (string) |
| Unit price | Ratio | The price of each product in dollars | Range(number) |
| Quantity | Ratio | The number of products purchased by the customer | Range(number) |
| Tax | Ratio | The tax fee for the customer's order, calculated as 5% of the total price | Range(number) |
| Total | Ratio | The total price of the order, including tax | Range(number) |
| Date | Interval | The date of the purchase | Range(number) |
| Time | Interval | The time of the purchase during working hours | Range(number) |
| Payment | Categorical | The payment method used by the customer for the transaction, such as "Cash," "Credit card," etc. | List (string) |
| COGS | Ratio | The cost of goods sold (COGS) for the order, excluding tax | Range(number) |
| Gross margin percentage | Ratio | The gross margin percentage on the order | Range(number) |
| Gross income | Ratio | The gross income made on the order | Range(number) |
| Rating | Ordinal | The customer's total purchasing experience rated on a scale of 1 to 10, with different categories based on their degree of satisfaction | Range(number) |

* Based on the given dataset, the prospective users for the dashboard and how they could use it are:
  + Sales team: The dashboard may be used by the sales team to track the sales performance of various items across different areas.
  + Marketing team: The dashboard may be used by the marketing team to determine the target demographic for each product and the channels via which they can contact them.
  + Product development team: The dashboard may be used by the product development team to determine the features that are most popular among consumers, as well as the locations where those features are in high demand.
  + Executive team: The dashboard provides executives with an overview of the company's sales success as well as the performance of individual items.
  + Finance team: The dashboard may be used by the finance team to track the revenue and profit margins of various products and areas.
* Set of questions that the users might ask about this dataset.

1. What is the frequency of payment methods for invoice attributes?
2. What would be the total sales of a branch per month?
3. What would be the total sales of a product line per month?
4. Checking the number of visitors for each customer type.
5. Product line average rating per city?
6. How much were the total sales per day?
7. The number of customers that are members in a specific city?
8. Gross profit made per product line in each branch?
9. What is the frequency of Gender purchasing over all the product line product line?
10. Average quantity sale for all Product line?
11. How much was the total tax revenue generated across all product lines per day?
12. Average sale made in a specific time period with respect to the branch?

Mural link:  
  
<https://app.mural.co/t/mili4400/m/mili4400/1680789780520/bbef0b79bab3d701c5952e8657876f31829f090d?sender=ueae0fdde06df7d3aeafc9985>

Dataset Link:

<https://www.kaggle.com/datasets/aungpyaeap/supermarket-sales>

Color for individuals in Mural:

Purple: Mili   
Brown: Ashish

Yellow: Monish