Refrigerate the Risk: Optimizing Inventory and Operations for Profit Enhancement in Dairy Retail

A Mid-Term report for the BDM capstone Project

Submitted by

Name: Mili Parashar

Roll number: 23f2004291

IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

Contents

1 Executive Summary and Title **2**

2 Proof of Originality of Data **2**

2.1 Shop Photos: 3

2.2 Letter Head: 3

2.3 Video Interaction with the Owner: 3

3 Metadata and Descriptive Analysis **3**

3.1 Metadata: 3

3.2 Descriptive Analysis: 7

4 Detailed Explanation of Analysis Process/Method **9**

4.1 Data Cleaning and Preprocessing: 9

4.2 Analysis Process / Method 10

4.2.1 Descriptive Statistics 10

4.2.2 ABC Analysis 10

4.2.3 Time Series Forecasting 10

4.2.4 Correlation Analysis 10

4.2.5 Seasonality and Peak Analysis 10

5 Results and Findings **11**

5.1 Three Varieties of Milk Vs Total Sales 11

5.2 Three Ice Cream Brands Vs Total Sales 11

5.3 Monthly Sales Quantity Trend 12

5.4 Cumulative Sales Contribution (ABC Analysis) 12

5.5 Temperature and Sales Correlation 13

1 Executive Summary and Title

This study focuses on ‘Kribhco Karmachari Sahkari Dhiran Purvarsh Society’, a small-scale dairy and ice cream retail business operating under a B2C model. The shop caters to over 420 regular milk customers, supported by a team of seven delivery personnel. The business faces recurring challenges in inventory management due to demand fluctuations, temperature-sensitive spoilage during extended power outages, and delivery inefficiencies, which contribute to lower revenue generation and customer dissatisfaction.

Data was collected for the period April 2024 to March 2025 from raw Excel files and a compiled supplier invoice register maintained by the owner. The datasets included milk sales and purchases, ice cream sales and purchases, and rough environmental data (temperature and power outages). Key metadata fields comprised transaction dates, bill numbers, product names, quantities, rates, and supplier agency names. Descriptive statistics and cumulative contribution analysis classified products by their business impact.

In order to mitigate these challenges, these datasets were modelled and structured into a 5-star schema with well-defined fact tables (milk sales, milk purchases, ice cream sales, ice cream purchases) and dimension tables (product info, customers, suppliers). Preliminary insights showed that Amul Gold Milk accounted for 57% of milk revenue (Class A product), while Amul brand dominated ice cream sales with chocolate-flavored sticks being most preferred. Tools like ABC analysis and trend charts highlighted peak months—April for ice cream and July for milk. Upcoming phases will focus on demand forecasting, spoilage reduction during outages, and delivery optimization to support data-driven decision-making and profit enhancement.

2 Proof of Originality of Data

Shop Details:

Shop Name	Kribhco Karmachari Sahkari Dhiran Purvarsh Society
Address	Kribhco Town Shopping Center, Fertilizer Complex, Kribhconagar, Hazira INA, Hazira, Limla, Gujarat 394515
Shop Owner’s Name	Mr. Ashok Jana

Type of Organization	B2C Organization
Number of Employees	1 (Owner) + 7 (Delivery Boys)
Working Time	5:00 AM to 12:00 PM (Morning) 4:00 PM to 11:00 PM (Evening)

2.1 Shop Photos:

Figure 1: Front View of the Shop

Other shop photos: Shop Images and Daily-Bills Register Book

2.2 Letter Head:

Here is the link to the duly signed letter head signed by the owner of the shop.

2.3 Video Interaction with the Owner:

The 5-minute video interaction with the owner was conducted in Hindi, and to ensure clarity, accurate subtitles (in English) have been added to the video. Here is the Video Link.

3 Metadata and Descriptive Analysis

3.1 Metadata:

It describes the essential attributes of the datasets gathered over a year of transaction data, transforming raw handwritten and system-generated records into a digital format. The compiled dataset includes 869 milk sales records, 37 milk purchase records, 107

ice cream purchase invoices, and 88 ice cream sales invoices. Temperature and precipitation data were sourced from Visual Crossing Weather Data, ensuring reliable and location-specific weather records.

- **Clean Dataset Link:** BDM Project Data
- **Data collection duration:** 12 months
- **Data collection dates:** 01-04-2024 to 31-03-2025

The datasets, their variables, and their relevance to the analysis are as follows:

- **Milk Sales Dataset:**
    - **Sale_ID:** Unique identifier for each monthly aggregated sale record.
    - **Sales Date:** Month of recorded sales (not daily transactions), reflecting cumulative quantities sold per SKU.
    - **Bill:** Reference bill number
    - **Customer ID:** Unique identifier of each milk delivery customer (linked to the Customer Data).
    - **Milk Name:** Type of milk sold (Amul Gold, Amul Shakti, Cow Milk).
    - **Quantity (in Litres)**: Total liters sold in the month (essential for inventory forecasting).

  Justification: Supports demand forecasting, ABC classification, and seasonality analysis to reduce overstocking and stockouts.

- **Milk Purchase Dataset:**
    - **Purchase ID:** Unique identifier of each purchase record.
    - **Purchase Date:** Date when purchase was recorded.
    - **Milk Name:** SKU purchased (Amul Gold, Amul Shakti, Cow Milk)
    - **Quantity (in Litres):** Quantity procured.
    - **Bill Amount:** Total procurement cost.

  Justification: Enables calculation of gross margins, replenishment cycles, and cost dynamics influencing profitability.

- **Milk Data Table:**
    - **Milk ID:** Unique SKU identifier.
    - **Milk Name:** Name of the milk product.
    - **Milk Type:** Toned or Full Cream classification

- **Supplier Name:** Single supplier identity

Justification: Essential for linking sales and purchase data to product characteristics.

- **Customer Data Table:**
  - **Customer ID:** Unique identifier of each customer.
  - **Customer Name:** Name of the milk delivery customer.

Justification: Enables customer-level profitability analysis and segmentation.

- **External Factor (Temperature) Data Table:**
  - **Date:** Calendar date.
  - **Temperature (°F):** Daily ambient temperature.
  - **Preciptype:** Precipitation type, if any. (rain/no rain)
  - **Power Outage:** Whether a power outage occurred (Yes/No).

Justification: Critical to analyze correlations between spoilage events, delivery disruptions, and external conditions

- **Ice Cream Sales Dataset:**
  - **Date:** Transaction date.
  - **Bill:** Invoice number.
  - **Quality Code:** Code representing ice cream SKU.
  - **Quantity:** Quantity sold.
  - **Rate:** Unit selling price.
  - **Bill Amount:** Revenue per transaction.

Justification: Used for ABC analysis, demand trends, and brand-wise sales performance.

- **Ice Cream Purchase Dataset:**
  - **Date:** Purchase date.
  - **Bill:** Purchase Invoice number given by supplier agency.
  - **Quality Code:** Code representing ice cream SKU.
  - **Quantity:** Quantity purchased.
  - **Rate:** Procurement cost.
  - **Bill Amount:** Total purchase value

Justification: Enables purchase planning and cost optimization.

- **Ice Cream Data Table:**
    - **IC ID:** Unique identifier for each particular ice cream.
    - **Brand Name:** Amul, Havmor, or Vadilal.
    - **Ice Cream Name:** Name of ice cream under a particular quality code
    - **Quality Code:** Code representing ice cream SKU.
    - **Supplier Name:** Supplier providing each brand.
    - **Ice Cream Type:** Format such as Stick, Cone, Tub, or Family Pack.
    - **Flavour:** Variant (e.g., Chocolate, Butterscotch).

Justification: Helps with SKU-level profitability, product segmentation, and forecasting.

- **IC Quality Code Count Table:**
    - **Count of IC ID:** Numer of ice creams under each quality code
    - **Quality Code:** Code representing ice cream SKU, identifier to match transactional records to master data.

Justification: Supports accurate mapping of ice cream SKUs in transactional analysis.

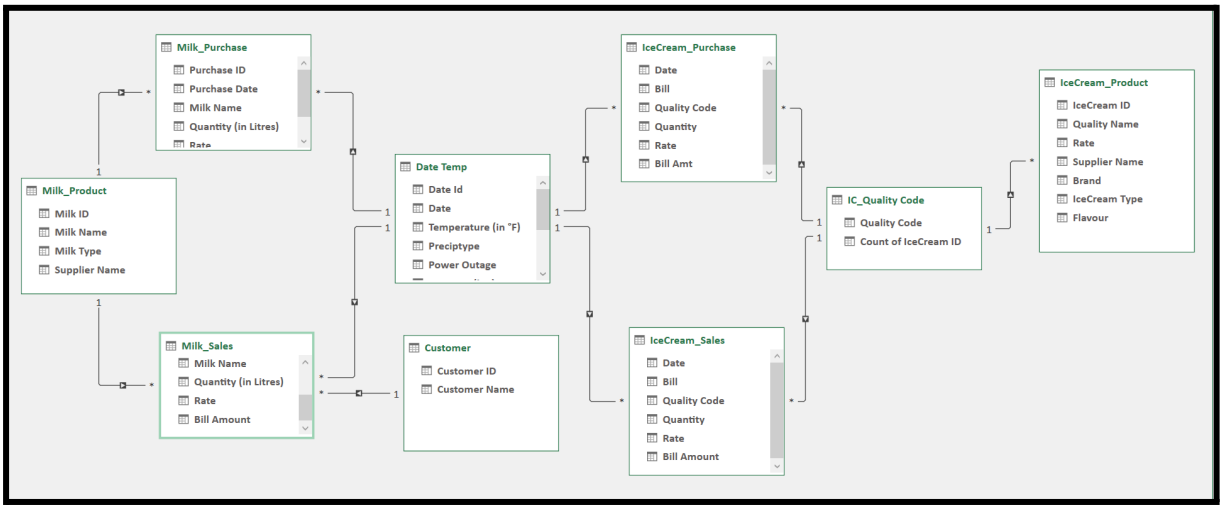


Figure 4: Entity-Relationship Dagram

Together, these integrated datasets and their structured relationships enable focused analysis of demand forecasting, cold-chain risks, and delivery reliability. This foundation ensures each problem statement is addressed through relevant, connected, and actionable data.

3.2 Descriptive Analysis:

To gain quantitative insights into inventory movement, sales performance, and external risk factors, descriptive statistics were computed for both milk and ice cream product categories, as well as for environmental conditions affecting perishability.

- **Milk Sales Data**

STATISTICAL MEASURES (Sales Value in ₹)	MILK NAME		
	AMUL GOLD MILK	AMUL SHAKTI MILK	COW MILK
MEAN	2242.16	1873.21	2107.79
MEDIAN	1604.00	1344.00	1131.00
STANDARD DEVIATION	7990.20	4586.70	3308.46
MAXIMUM	1,44,025.00	68,510.00	21,504.00
MINIMUM	0.00	0.00	0.00
Total Sales	54,61,907.00	36,95,840.00	5,33,272.00

STATISTICAL MEASURES (Quantities Sold in L)	MILK NAME		
	AMUL GOLD MILK	AMUL SHAKTI MILK	COW MILK
MEAN	32.90	29.45	36.55
MEDIAN	23	21	19.5
STANDARD DEVIATION	114.65	72.15	57.22
MAXIMUM	2057.5	1105	363
MINIMUM	0	0	0
Total Quantity	78,448	58,908	9,246

- **Ice Cream Sales Data**

STATISTICAL MEASURES	ICE CREAM BRAND

(Sales Value in ₹)	AMUL	VADILAL	HAVMOR
MEAN	1179.72	1407.44	1075.99
MEDIAN	864.00	864.00	1080.00
STANDARD DEVIATION	1028.35	1168.22	906.20
MAXIMUM	8820.00	11,880.00	5900.00
MINIMUM	81.00	153.00	180.00
Total Sales	8,77,709.00	4,51,788.00	2,19,501.00

STATISTICAL MEASURES (Quantities Sold in Packets)	ICE CREAM BRAND		
	AMUL	VADILAL	HAVMOR
MEAN	4.21	4.60	3.97
MEDIAN	3	3	3
STANDARD DEVIATION	3.27	4.07	3.16
MAXIMUM	24	36	20
MINIMUM	1	1	1
Total Sales	3130	1478	810

Justification:

1. **Mean:** Helps estimate typical daily or monthly demand, which is essential for forecasting inventory.
2. **Median:** Shows the central tendency, reducing the impact of extreme sales or quantity spikes.
3. **Standard Deviation:** High fluctuations signal greater demand uncertainty, requiring higher safety stock levels and flexible delivery planning.
4. **Maximum**: Identifies the largest sales or quantity event during occasional large orders.

5. **Minimum:** Highlights the lowest or zero-sales days, useful to understand low-demand periods, aiding cash flow management and seasonality analysis.
6. **Total Sales and Total Quantity:** Show overall business contribution of each product.

- **Temperature Logs**

STATISTICAL MEASURES (in °F)	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
MEAN	73.6	77.6	83.0	86.0	89.1	87.5	83.3	82.7	82.7	84.4	80.3	73.9
MAX	76.6	82.5	87.0	90.9	92.7	90.2	85.8	85.6	85.6	87.2	84.0	81.3
MIN	70.8	72.4	79.4	82.7	86.2	82.5	81.1	80.2	79.7	82.4	74.9	70.8
POWER OUTAGE (days per month)	4	3	2	4	5	7	12	8	4	4	4	1
PRECIPITATION (days per month)	0	0	0	2	2	21	31	31	27	20	0	0

Justification:

High-temperature variance, frequent outages, and a large number of rainy days align with cold storage failure risk and delivery disruptions. Correlation between temperature spikes and spoilage can support actionable cold-chain improvements and plan delivery workload.

4 Detailed Explanation of Analysis Process/Method

4.1 Data Cleaning and Preprocessing:

Raw data was obtained from sales registers and system records. The cleaning process involved:

- Data deduplication was performed, removing redundancy.
- Missing records were identified and handled by cross-verifying bills.
- Creating consistent name and date formats (eg, Amul, AMUL, amul, A-200 all are the same, and for date: DD-MM-YYYY)
- Implementing a 5-star schema for modeling data tables.

- Linking products to descriptive attributes (eg, milk type, ice cream flavor)

  Justification: Clean data guarantees valid statistical measures, better predictive models, and thereby results in the prevention of under- or overestimating true demand and trustworthy insights to inform decisions.

4.2 Analysis Process / Method

4.2.1 Descriptive Statistics

- **Objective**: Understand demand patterns, variability, and seasonality.
- **Methods Applied:** Calculate mean (to estimate baseline sales/quantities), median (to detect outliers), standard deviation (to measure volatility), and max/min (to identify peaks and dips)
- **Justification:** Baseline demand and product prioritization

4.2.2 ABC Analysis

- **Objective:** Prioritize products contributing to revenue and volume
- **Methods Applied:** Rank SKUs by cumulative sales value, calculate their contribution to total sales, and finally classify them into class A (High), class B (Moderate), or class C (Low) products.
- **Justification:** Inventory optimization

4.2.3 Time Series Forecasting

- **Objective:** Predict future demand, precise reorder point and EOQ planning
- **Methods Applied:** Plot monthly sale trends to model seasonality
- **Justification:** Predictive stocking to prevent stockouts/spoilage

4.2.4 Correlation Analysis

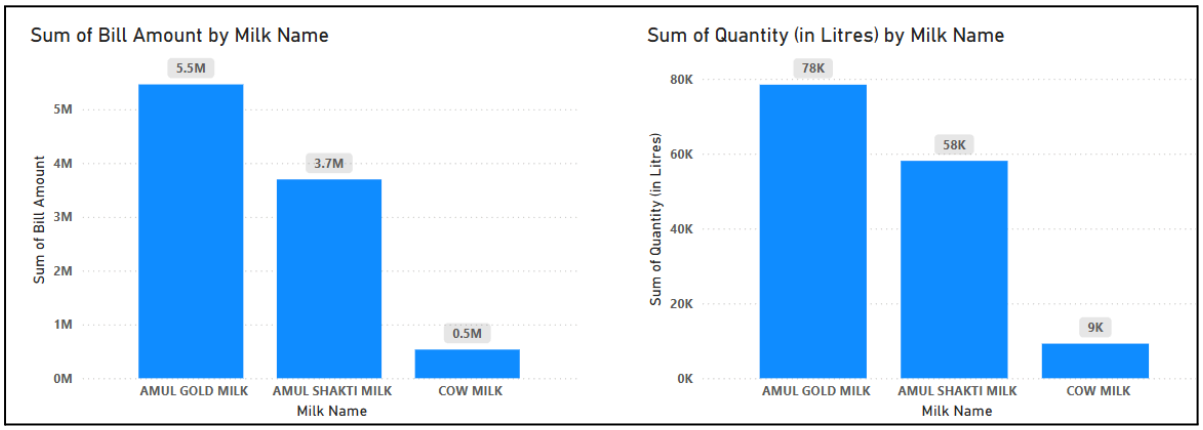
- **Objective:** Assess relationship between temperature, rain, power outage, and lower sales
- **Methods Applied:** Analyze temperature records versus monthly sales variations to detect spoilage vulnerabilities.
- **Justification:** Cold-chain risk quantification

4.2.5 Seasonality and Peak Analysis

- **Objective:** Identify periods of high delivery workload
- **Methods Applied:** Calculate the frequency distribution of order quantities by month and analyze the cumulative percentage
- **Justification:** Delivery planning

5 Results and Findings

5.1 Three Varieties of Milk Vs Total Sales

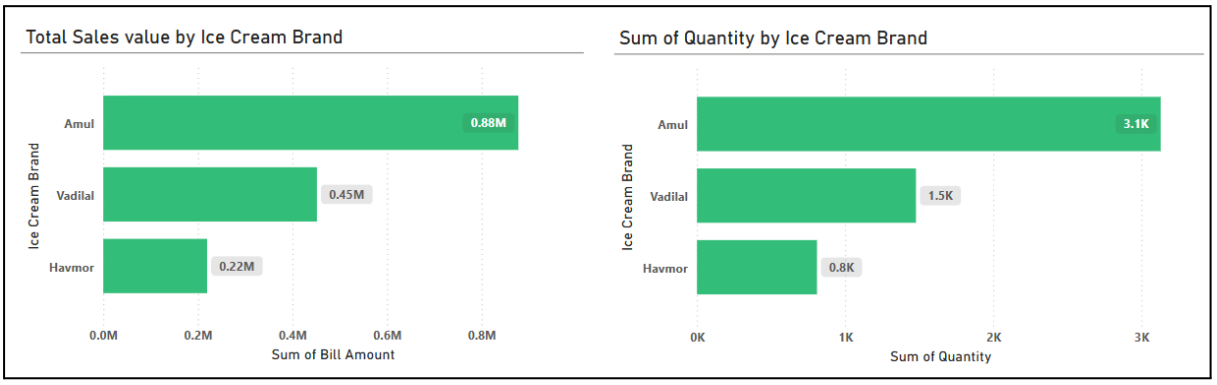


Graph 1&2(Column Chart): Total sales value vs Milk variety (left), Total Quantity (in L) sold vs Milk Variety (right)

Insights:

Amul Gold contributed the **highest revenue** and **highest volume** sold among milk products, followed by Amul Shakti and Cow Milk. This indicates Amul Gold is the most critical SKU for profitability and should be prioritized in forecasting and stock management.

5.2 Three Ice Cream Brands Vs Total Sales

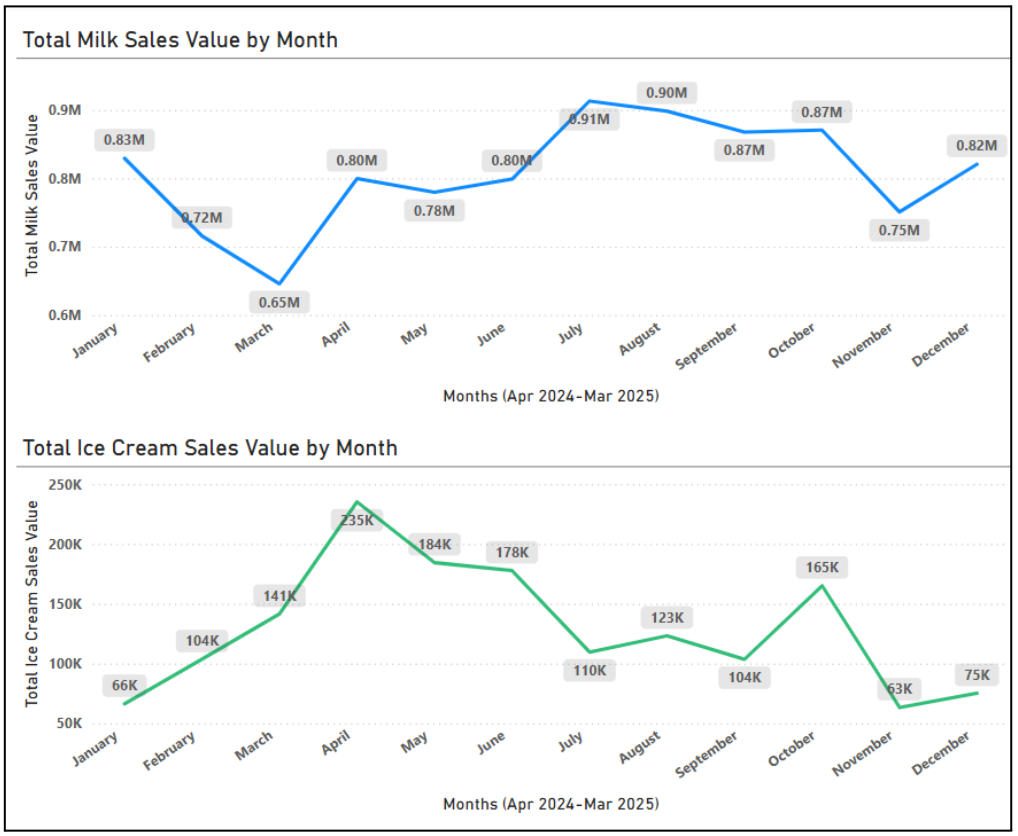


Graph 3&4(Bar Chart): Total sales value vs Ice Cream Brand (left), Total Quantity sold vs Ice Cream Brand (right)

Insights:

Among ice cream brands, **Amul** generated the **largest** share of sales revenue, followed by Vadilal and Havmor. The quantity trends aligned with sales value, and maintaining a higher safety stock of Amul Ice cream is essential to avoid lost sales.

5.3 Monthly Sales Quantity Trend

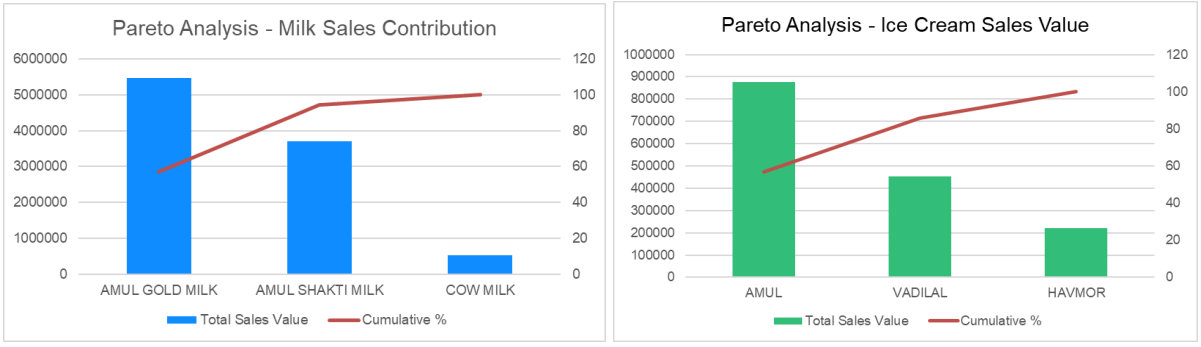


Graph 5&6(Line Chart): Monthly Milk Sales Value Trend (up), Monthly Ice Cream Sales Value Trend (down)

Insights:

Ice cream sales peaked in **April**, with additional highs in **June** and **October**, and dropped significantly in **January**, **November**, **July**, and **September**, indicating strong seasonality likely tied to weather and festivals. Milk sales reached their highest levels in **July**, followed by **September** and **October**, with noticeable dips in **March** and a slight decline in **November**. These trends emphasize the importance of tailoring inventory and delivery planning to seasonal demand fluctuations for both product categories.

5.4 Cumulative Sales Contribution (ABC Analysis)

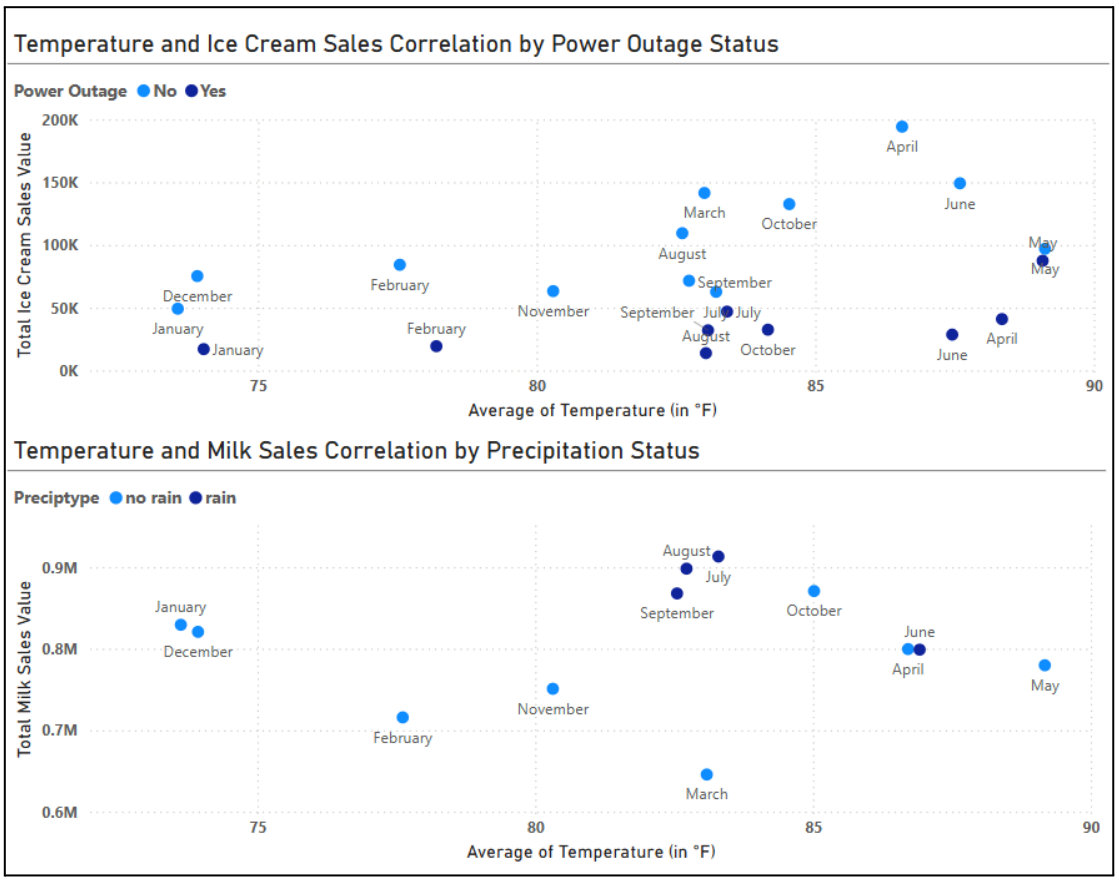


Graph 7& 8(Pareto Chart): ABC Analysis of Milk Sales(left), ABC Analysis of Ice Cream Sales (right)

Insights:

The analysis revealed that **Amul Gold Milk** falls into **Class A**, contributing the majority of milk revenue, while **Amul Shakti** and **Cow Milk** belong to **Class C** due to their lower contribution. In the ice cream category, the **Amul brand** emerged as **Class A**, dominating both sales value and volume. **Vadilal** was categorized as **Class B**, and **Havmor** as **Class C**. This confirms the **80/20 principle**, highlighting the need to prioritize **Amul Gold Milk** and **Amul Ice Creams.**

5.5 Temperature and Sales Correlation



Graph 9&10(Scatter Plot): Monthly Avg. Temp and Ice Cream Sales Correlation (up), Monthly Avg. Temp and Milk Sales Correlation (down)

Insights:

In Ice Cream Sales Scatter Plot, Higher temperatures drive more ice cream sales (April–June peak) and Power outages reduce sales even at high temperatures. However, in Milk Sales Scatter Plot, milk sales remain steady across temperatures, rain has minimal impact on milk demand, and slight dips are seen in March and June. Milk demand is **stable** and less weather-dependent, indicating a consistent baseline demand useful for forecasting. Ice cream is **highly sensitive to temperature** and **power outages**, suggesting cold-chain reinforcement during hot months and contingency planning for outages.