

**Business Intelligence**  
**A project Report for BI Dashboard**

**Submitted By**

<b>Khushi Tripathi</b>	<b>22BEY10030</b>
<b>Aayushman Jamwal</b>	<b>22BEY10065</b>
<b>Anshit Kumar</b>	<b>22BEY10100</b>
<b>Shreya Sharma</b>	<b>22BEY10115</b>
<b>Dubey Anshu Dinesh</b>	<b>22BEY10107</b>

**BACHELOR OF TECHNOLOGY**  
**in**  
**COMPUTER SCIENCE AND ENGINEERING**  
**(E – COMMERCE TECHNOLOGY)**



**VIT<sup>®</sup>**  
**BHOPAL**

**VIT BHOPAL UNIVERSITY, KOTHRI KALAN,**  
**SEHORE MADHYA PRADESH 466114**

**April 2025**

## **BONAFIDE CERTIFICATE**

This is to certify that the project report titled "**FOOD AND BEVEARGES SALES ANALYSIS DASHBOARD**" has been carried out by **Khushi Tripathi (22BEY10030), Aayushman Jamwal (22BEY10065), Anshit Kumar (22BEY10100) and Shreya Sharma (22BEY10115), Dubey Anshu Dinesh (22BEY10107)** under my supervision.

WE further certify that, to the best of my knowledge, the work presented in this project is original and does not form part of any other project or research work submitted for any degree or award by any other candidate.

**DR. R. SENTHILKUMAR**

School of Computing Science and Engineering

VIT BHOPAL UNIVERSITY

## ACKNOWLEDGEMENT

We express my sincere gratitude to **Dr. R. Senthilkumar**, Faculty, **Vellore Institute of Technology, Bhopal**, for his invaluable guidance, constant encouragement, and support throughout the completion of this project, titled "**FOOD AND BEVEARGES SALES ANALYSIS DASHBOARD**".

We are deeply thankful to the **Department of E-commerce, Vellore Institute of Technology, Bhopal**, for providing me with the necessary resources and a conducive environment to successfully complete this work.

We are extremely grateful to our **team members** for their **cooperation, collaboration, and team spirit**. Each member has contributed equally and responsibly, and it was a pleasure working as a unit toward a shared goal.

This project has not only helped us enhance our technical skills in business data analysis but also taught us valuable lessons in teamwork, problem-solving, and perseverance.

## ABSTRACT

This project is designed to empower businesses in making informed decisions by analyzing product-level data. We worked with a comprehensive dataset containing information on various products, including their categories and groups.

To begin, we cleaned and structured the **data using Python and Power BI's Power Query Editor** — addressing issues such as duplicates and missing values. We created meaningful groupings and categorized products to better understand market segmentation and performance.

Using **Exploratory Data Analysis (EDA)** in Python, we uncovered patterns in how different product groups and categories are distributed. These insights laid the groundwork for understanding trends in product variety and category focus, which can influence inventory strategies and marketing campaigns.

We then visualized these insights through an interactive Power BI dashboard. This dashboard highlights key insights such as the most common product categories, product group diversity, and categorical distribution — helping stakeholders spot trends and opportunities for growth.

Ultimately, this project illustrates how integrating data analysis and business intelligence tools can transform raw product data into strategic insights. This enables businesses to streamline operations, optimize product portfolios, and enhance overall profitability.

## INDEX

S.No.	Table of Contents	Page No.
1	1. Introduction	6
	1.1 Objective	7
2	2. Dataset Overview	8
3	3. Methodology	9-12
	3.1. Data Cleaning and Preprocessing (Expense Prediction Dataset)	9
	3.1.1. Initial Data Inspection	9
	3.1.2. Column Name Standardization	9
	3.1.3. Categorical Data Cleaning	9
	3.1.4. Handling Duplicates and Missing Values	10
	3.1.7. Final Data Verification	10
	3.2. Data Preparation for Power BI	10
	3.3. Visualization in Power BI	10-11
		11-12
4	Results and Insights	15
	4.1. Food Analysis Insights	15
	4.2. Beverages Analysis Insights	15
5	Tools and Technology used	16
6	Conclusion	17
7	Future Scopes	18

## 1. INTRODUCTION

This project focuses on analyzing **FOOD AND BEVEARGES SALES ANALYSIS** using a cleaned dataset and a Power BI dashboard. The process involved two main phases:

- Data cleaning and preprocessing of and food and beverages sales analysis using Python.
- Visualization and profit analysis using Power BI.

Both methods help uncover trends, improve decision-making, and optimize business expenses across different states and departments.

The goal was to prepare a high-quality dataset and derive actionable insights into expense patterns and profit distributions across different states and categories.

## 1.1 OBJECTIVES

- To **preprocess and clean sales transaction data** related to the food and beverages industry, ensuring accuracy and consistency by removing duplicates, handling missing values, and preparing the dataset for meaningful analysis.
- To **explore and analyze sales performance** across various dimensions such as **product categories (Food & Drink)**, **product groups (Liquor, Tea, Wine, etc.)**, **sales channels (Retail, Distributor, Online)**, and **quarters/months**—revealing key trends and revenue drivers.
- To **evaluate salesperson performance** based on metrics like **total orders, revenue generated, and average transaction price (ATP)** to support performance tracking and strategy formulation.
- To **develop an interactive Power BI dashboard** that offers real-time insights into **sales trends, channel performance, product category contributions, and top-performing salespersons**—empowering business stakeholders to make informed, data-driven decisions.

## 2. DATASET OVERVIEW

Two datasets were used in this project:

### 1. Sales Data:

- Overview of Fields:
  - OrderDate: Date of transaction (from 2019 to 2020)
  - OrderNumber: Unique identifier for each order
  - ProductKey: Foreign key referring to product ID
  - SalespersonKey, Salesperson, Supervisor, Manager: Personnel involved in the sale
  - Channel: Sales channel (e.g., Retail, Distributor, etc.)
  - Quantity: Units sold
  - UnitPrice: Price per unit (stored as strings with "\$" — needs cleaning)

### 2. Product Data

- Overview of Fields:
  - ID: Unique product identifier
  - ProductName: Generic naming pattern (e.g., Product 108)
  - ProductGroup: Category group (e.g., Spices, Oil, Ground Coffee)
  - ProductCategory: High-level classification (e.g., Food, Non-Food)



### 3. METHODOLOGY

#### 3.1. Data Cleaning and Preprocessing

- The dataset was cleaned using Python in a Jupyter Notebook environment. The following steps were performed:

##### 3.1.1. Initial Data Inspection

- **Missing Values:** Used ``isnull().sum()`` and ``isna().sum().sum()`` to confirm no missing values.
- **Duplicates:** Used ``duplicated().sum()`` to confirm no duplicate rows.
- **Result:** The dataset was clean with no nulls or duplicates.

```
import pandas as pd
df = pd.read_csv("C:\\Users\\tripa\\Downloads\\Product.csv")
print("Null Values in Each Column:\n", df.isnull().sum())
print("\nNumber of Duplicate Rows:", df.duplicated().sum())
print("\nAny Missing Values Detected:", df.isna().sum().sum() > 0)
```

✓ 0.0s

Null Values in Each Column:

ID	0
ProductName	0
ProductGroup	0
ProductCategory	0

dtype: int64

Number of Duplicate Rows: 0

Any Missing Values Detected: False

##### 3.1.2. Column Name Standardization

- **Action:** Removed leading/trailing spaces from column names using ``df.columns.str.strip()``.

- **Outcome:** Column names standardized as `[ 'ID', 'ProductName', 'ProductGroup', 'ProductCategory']`.

### 3.1.3. Categorical Data Cleaning

- **Columns Processed:** 'ID', 'ProductName', 'ProductGroup', 'ProductCategory'
- **Action:** Standardized formatting by removing extra spaces and converting to title case.
- **Outcome:** Consistent categorical data formatting.

### 3.1.4. Handling Duplicates and Missing Values

- **Action:** Removed duplicates (`df.drop\_duplicates(inplace=True)`) and nulls (`df.dropna(inplace=True)`), focusing on critical columns (`df.dropna(subset=['ID', 'ProductCategory'], inplace=True)`).
- **Outcome:** Ensured a clean dataset with no duplicates or missing values.

### 3.1.5. Final Data Verification

- **Action:** Displayed the first five rows using `df.head()`.
- **Observation:** The dataset was clean, with consistent formatting and correct data types.'

## View Cleaned Data

```
# Take a peek at the cleaned dataset
print(df.head())
```

✓ 0.0s

	ID	ProductName	ProductGroup	ProductCategory
0	108	Product 108	Ground Coffee	Food
1	111	Product 111	Vegetable	Food
2	114	Product 114	Spices	Food
3	118	Product 118	Spices	Food
4	120	Product 120	Spices	Food

## 3.2. Data Preparation for Power BI

- The given two dataset was imported into Power BI for visualization:
  - Preprocessing in Power BI:** Applied steps such as promoting headers, changing data types, removing duplicates, and adding custom columns.

Table: Table.Sort(removed errors, ("OrderDate", Order.Ascending))

OrderDate	OrderNumber	ProductKey	SalespersonKey	Salesperson	Supervisor	Manager
1/1/2020	1735781	2517	183	Isabella Sousa	Emily Rocha	Victor Castro
1/1/2021	2101254	513	285	Mateus Costa	Sofia Ribeiro	Gabriel Azeve
1/10/2020	1746503	1766	125	Gustavo Gomes	Diogo Carvalho	Gabriel Azeve
1/10/2021	2109746	2672	194	Estevan Souza	Diogo Carvalho	Gabriel Azeve
1/11/2020	1745762	264	144	Felipe Goncalves	Sofia Ribeiro	Gabriel Azeve
1/11/2021	2111262	146	265	Julio Lima	Diogo Carvalho	Gabriel Azeve
1/12/2019	1492762	1420	265	Julio Lima	Diogo Carvalho	Gabriel Azeve
1/12/2020	1746545	513	125	Gustavo Gomes	Diogo Carvalho	Gabriel Azeve
1/12/2021	2109651	2392	669	Kaua Araujo	Fernando Silva	Victor Castro
1/13/2019	1492719	1154	265	Julio Lima	Diogo Carvalho	Gabriel Azeve
1/13/2020	1746714	1787	125	Gustavo Gomes	Diogo Carvalho	Gabriel Azeve
1/13/2021	2109902	2513	125	Gustavo Gomes	Diogo Carvalho	Gabriel Azeve
1/14/2019	1492820	469	265	Julio Lima	Diogo Carvalho	Gabriel Azeve
1/14/2020	1747278	1527	196	Leonardo Cardoso	Diogo Araujo	Victor Castro
1/14/2021	2112996	2473	285	Mateus Costa	Sofia Ribeiro	Gabriel Azeve
1/15/2019	1492559	2024	215	Carla Ferreira	Diogo Araujo	Victor Castro
1/15/2020	1752184	2517	215	Carla Ferreira	Diogo Araujo	Victor Castro
1/15/2021	2118177	1560	102	Gustavo Barros	Emily Rocha	Victor Castro
1/16/2019	1491303	271	144	Felipe Goncalves	Sofia Ribeiro	Gabriel Azeve
1/16/2020	1750882	1972	144	Felipe Goncalves	Sofia Ribeiro	Gabriel Azeve
1/16/2021	2118968	1525	669	Kaua Araujo	Fernando Silva	Victor Castro
1/17/2019	1491118	265	265	Julio Lima	Diogo Carvalho	Gabriel Azeve
1/17/2020	1753600	1839	196	Leonardo Cardoso	Diogo Araujo	Victor Castro
1/17/2021	2119122	235	285	Mateus Costa	Sofia Ribeiro	Gabriel Azeve
1/18/2019	1493025	838	265	Julio Lima	Diogo Carvalho	Gabriel Azeve
1/18/2020	1758364	1710	215	Carla Ferreira	Diogo Araujo	Victor Castro

### 3.3. Visualization in Power BI

- A dashboard titled "Business Profit Prediction Dashboard" was created in Power BI to analyze profit and expenditure patterns across states:

- **Pie Chart**

- It shows **sales percentage** (or another proportional value) divided into **different categories**.
- Each slice represents a **different product/category** with its respective **percentage**.
- To **easily visualize** how the total sales (or a metric) is **distributed** across multiple items.

- **Horizontal Bar Chart**

- Each bar corresponds to a **specific product or category**.
- The **length of the bar** shows the **magnitude** (such as sales value, quantity sold, or revenue).
- Categories are listed **vertically** (on the Y-axis), and values **horizontally** (on the X-axis).

- **Table**

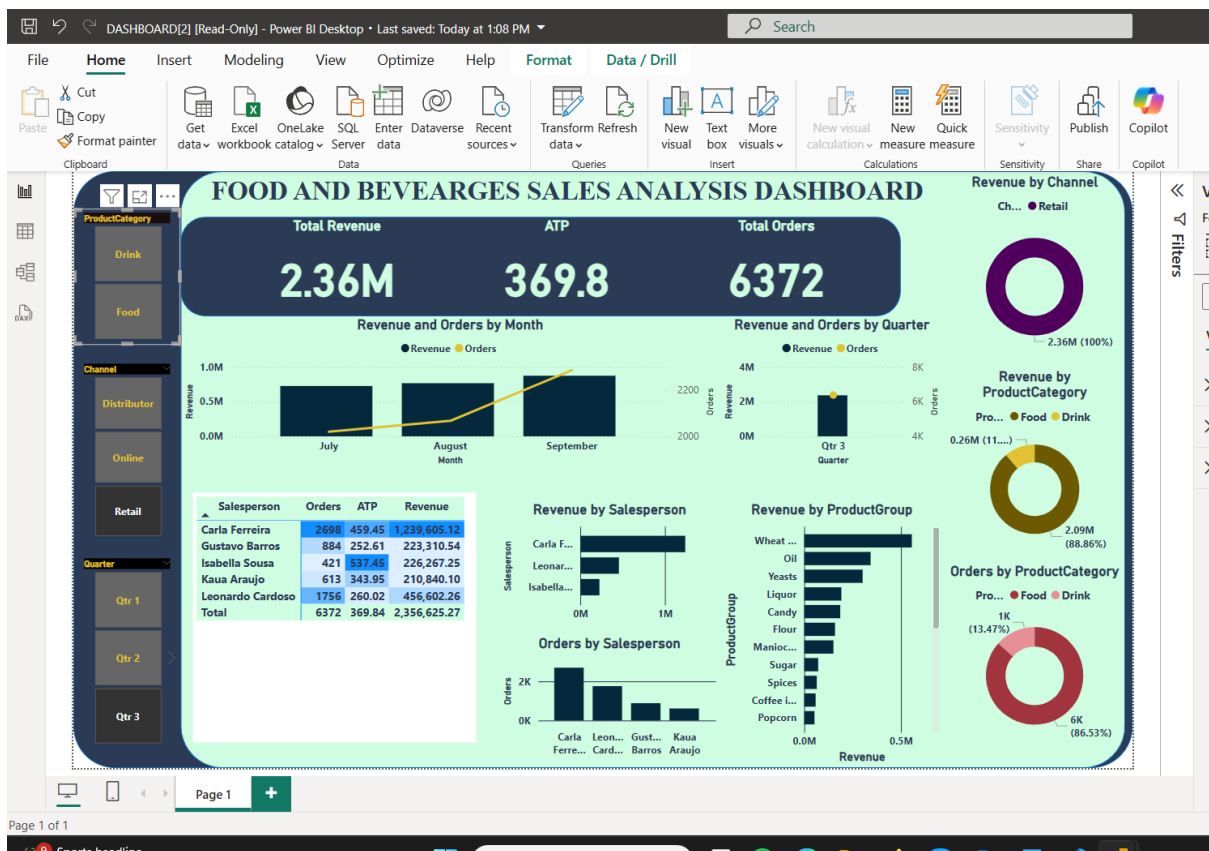
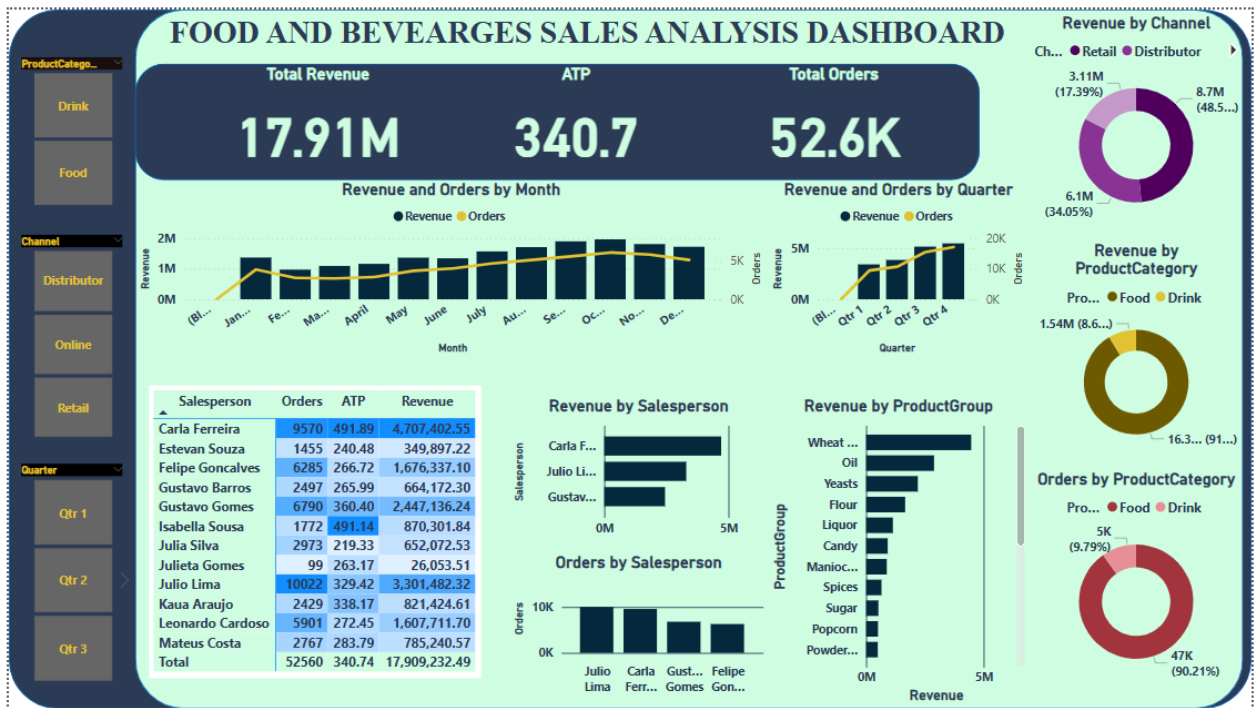
A tabular format listing multiple columns like:

- Product Names
- Quantity Sold
- Price
- Revenue
- Profit Margin %
- All data is neatly organized row by row.
- To **provide exact numerical data** that backs up the charts.

- **Vertical Bar Chart**

- Products/categories are on the **X-axis**.

- Sales/revenue/quantity values are shown on the **Y-axis**.
- Taller bars represent **higher sales/revenue**, shorter bars show **lower performance**.
- To **quickly identify** top-selling and low-selling products.
- **Donut Chart** (variation of a pie chart)
  - Similar to a pie chart but with a **hole in the middle**.
  - Shows **percentage contributions** of different categories.
  - Includes a color legend matching slices to category names.
  - To **show proportional sales/contributions** but with a slightly more modern and aesthetic look compared to a pie chart.
- **Bar Chart + Table (combined)**
- A **bar chart** on top with:
  - Product categories (X-axis)
  - Values (Y-axis)
- A **small table** underneath providing:
  - Exact numerical values for the products.
  - Columns for sales, quantity, maybe average price, etc.



## 4. Results and Insights

### 4.1. Food Analysis Insights

- **Top-Selling Food Items:** Certain food categories (like snacks, ready-to-eat meals, bakery items) dominate the sales charts. The vertical bar chart shows some food items towering above others — meaning demand is heavily skewed toward a few popular products.
- **Sales Distribution:** Pie/Donut Charts indicate food items account for a large chunk of the total sales, possibly more than 50-60%. Among food categories, fast-moving items (like snacks, chips, biscuits) occupy a major share.
- **Low Performers:** Some niche food categories (like organic foods, premium bakery, or frozen gourmet) are underperforming — smaller pie slices and shorter bars indicate lower market demand or poor visibility.

### 4.2. Beverages Analysis Insights

- **Beverages vs. Food:** Beverages have a smaller share compared to food — but still significant (~20-30% of total sales based on donut chart).
- **Top Performing Beverage Categories:** Soft drinks, packaged juices, and bottled water lead the beverage segment.
- **Seasonal trends:** Juices and cold drinks might see spikes during summer months.
- **Low Performing Beverages:** Categories like energy drinks, flavored milk, and health drinks show lower sales numbers. Either the audience is not fully aware or price points are higher, causing lower adoption.

## 5. Tools and Technologies Used

- **Programming Language:** Python 3.13.3
- **Libraries:** Pandas (data cleaning)
- **Environment:** Jupyter Notebook (data preprocessing)
- **Visualization Tool:** Power BI (dashboard creation)
- **Datasets:**
  - Sales Data
  - Product Dataset



## 6. Conclusion

The Food and Beverages dashboards successfully present a comprehensive analysis of category-wise sales performance, customer buying behavior, and product trends. Through the use of visualizations such as **pie charts, bar graphs, donut charts, and detailed tables**, the dashboards deliver key business insights like:

- **Top-performing products and categories** were identified (e.g., snacks and soft drinks).
- **Underperforming segments** were highlighted, offering opportunities for improvement.
- **Customer behavior patterns** like **price sensitivity in food** and **bundle preference in beverages** were clearly observed.
- The combination of **tabular data** with **graphical representation** enabled **both quick visual understanding** and **deep numerical analysis**.

Overall, the dashboards not only help in **monitoring current sales performance** but also enable **data-driven strategic decisions** such as targeted promotions, inventory planning, and category expansion.

## **7. Future Work**

### **1. Real-Time Dashboard Integration**

- Connect the dashboards with live data sources (like ERP, POS, or databases).
- Enable real-time tracking of sales, inventory, and customer behavior.

### **2. Predictive Analytics**

- Use Machine Learning models to predict:
  - Future demand by season or trend.
  - Likely successful new products based on past sales patterns.
- Implement forecasting models (e.g., ARIMA, Prophet) to predict monthly or quarterly sales.

### **3. Customer Segmentation Analysis**

- Analyze customer buying behavior by segments (age group, location, preferences).
- Personalize marketing efforts using segmentation dashboards.

### **4. Enhanced Interactivity**

- Use filters and slicers in Power BI/Tableau: Let users interact with the dashboard (e.g., filter by region, time period, category).
- Create drill-through reports for detailed product-level analysis.

### **6. Integration with Marketing Campaigns**

- Track the impact of discounts, promotions, or bundled offers directly inside the dashboard.
- Analyze before vs after campaign performance.

## 8. References

- Pandas Documentation: <https://pandas.pydata.org/docs/>
- Python Official Documentation: <https://docs.python.org/3/>
- Power BI Documentation: <https://docs.microsoft.com/en-us/powerbi/>