# UNVEILING SALES INSIGHTS THROUGH DATA ANALYTICS

### By

**ABIRAMI M**

# ABSTRACT

This project focuses on leveraging Excel for data analytics, emphasizing its powerful capabilities in organizing, processing, and analyzing data. The primary objective was to explore the application of various Excel functions and tools, including data cleaning, visualization, and statistical analysis, to gain meaningful insights from raw datasets. The project included importing data, conducting data preprocessing steps such as handling missing values and normalizing data, and then using Excel’s advanced features like pivot tables, charts, and functions to perform detailed data analysis.The analysis also involved creating dynamic dashboards and reports that provide a clear and interactive way to present data- driven insights. Key performance indicators (KPIs) were calculated and visualized to track trends, correlations, and outliers within the dataset. The results of this analysis were utilized to inform decision-making processes, demonstrating the importance of data-driven strategies in various business scenarios.This project aims to analyze supermarket sales data using Excel to identify trends, optimize inventory, and enhance sales strategies. The analysis focused on key metrics such as total sales, average transaction value, sales by product category, and sales by time period. The data was meticulously cleaned and organized within Excel, ensuring accuracy and reliability.

Advanced Excel tools, including pivot tables, charts, and data filtering, were utilized to break down sales data into actionable insights. Through the use of pivot tables, sales were segmented by various dimensions such as product category, location, and customer demographics, enabling a deeper understanding of customer behavior and preferences.Visualizations, including bar charts, line graphs, and Pie graphs, were created to highlight sales patterns and anomalies over time. This facilitated the identification of peak sales periods, high-performing product categories, and underperforming items, allowing for data-driven decisions in inventory management and marketing strategies.

**TABLE OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| CHAPTER NO: | TITLE | PAGE NO: |
|  | ABSTRACT | Ⅴ |
| 1 | INTRODUCTION | 1 |
| 2 | LITERATURE SURVEY | 5 |
| 3 | AIM AND OBJECTIVES | 6 |
| 4 | SUPERMARKET SALES DATA | 8 |
| 5 | CASE STUDY AND ANALYSIS | 11 |
| 6 | CONCLUSION | 12 |
| 7 | INTERNSHI[P CERTIFICATE | 18 |
| 8 | REFERENCE | 20 |

**CHAPTER 1 INTRODUCTION**

Data analytics in Excel typically begins with data cleaning and preparation, ensuring that datasets are free of errors and inconsistencies. Once the data is prepared, Excel's analytical tools are used to uncover patterns and correlations within the data, which can inform decision-making and strategy. Visualization features, such as charts and graphs, play a critical role in presenting data findings in a clear and impactful manner, making it easier for stakeholders to understand and act on the insights derived.Excel remains a powerful and accessible tool for data analytics, providing the necessary tools to transform raw data into actionable insights that can drive business success**.** In the context of business, Excel's data analytics capabilities are particularly valuable for tasks such as sales forecasting, financial analysis, market trend analysis, and operational performance reviews. The ability to quickly analyze and interpret data within Excel helps organizations stay agile and competitive in their respective markets.

### What Is Data Analytics?

Data analytics is the process of examining, cleaning, transforming, and interpreting data to discover meaningful patterns, trends, and insights that can inform decision- making. It involves using statistical and computational techniques to analyze large datasets, identify correlations, and generate predictive models that can be used to forecast outcomes or optimize processes..

## The field of data analytics encompasses several types, including:

1. **Descriptive Analytics**: Focuses on summarizing historical data to understand what has happened in the past.
2. **Diagnostic Analytics**: Examines data to determine the cause of past outcomes. iii.**Predictive Analytics**: Uses statistical models and machine learning algorithms to

forecast future events or trends.

1

iv.**Prescriptive Analytics**: Suggests actions based on data insights to achieve desired outcomes.

Data analytics is applied in various industries such as finance, healthcare, marketing, and retail to drive business decisions, enhance customer experiences, improve operational efficiency, and increase profitability. The process typically involves data collection, data processing, data analysis, and the presentation of findings through reports or visualizations.

### How Excel can be used effectively in data analytics?

Excel is a powerful tool for data analytics, and it can be used for a variety of tasks. Here are some key areas where Excel is commonly used

### Data Entry and Cleaning

* + Data Import: Import data from various sources (CSV, text files, databases).
  + Data Cleaning: Remove duplicates, handle missing values, and correct data inconsistencies.

### Data Analysis

* + Formulas and Functions: Use built-in functions like `SUM`, `AVERAGE`,

`VLOOKUP`, `INDEX`, `MATCH`, `IF`, and `COUNTIF` for basic data analysis.

* + Pivot Tables: Summarize and analyze data by creating pivot tables to aggregate, filter, and cross-tabulate data.

### Data Visualization

* + Charts and Graphs: Create various types of charts (bar, line, pie, scatter, etc.) to visualize data trends and distributions.
  + Conditional Formatting: Highlight data trends and patterns using color scales, data bars, and icon sets.

### Statistical Analysis

* + Descriptive Statistics: Calculate measures like mean, median, standard deviation, and variance.
  + Regression Analysis: Perform simple and multiple regression analysis to explore relationships between variables.
  + Hypothesis Testing:Use tools like the Analysis ToolPak for more advanced statistical tests.

### Data Modeling:

* + Scenario Analysis: Use tools like `What-If Analysis`, `Data Tables`, and `Goal Seek` to model different scenarios and forecast outcomes.
  + Solver: Optimize complex problems by finding the best solution for constraints and objectives.

### Automation and Advanced Techniques

* + Macros: Automate repetitive tasks using VBA (Visual Basic for Applications) to write custom scripts.
  + Power Query: Import, transform, and combine data from multiple sources with more advanced data transformation capabilities.
  + Power Pivot: Create complex data models, relationships, and perform advanced calculations using DAX (Data Analysis Expressions).

### Reporting

* + Dashboards: Build interactive dashboards to display key metrics and performance indicators.
  + Custom Reports: Generate customized reports to communicate findings effectively.

### MATHEMATICS IN EXCEL SPREADSHEET

Excel is quite versatile and can be used for various mathematical applications. Here are some common ways Excel is used in mathematics:

### Arithmetic Calculations

* + Basic Operations: Perform simple calculations like addition, subtraction, multiplication, and division.
  + Complex Formulas: Use nested formulas to handle more complex arithmetic operations.

### Algebra

* + Solving Equations: Use Excel to solve linear equations and systems of equations by leveraging built-in functions and Solver for more complex problems.
  + Polynomial Fitting: Use trendlines and polynomial functions to fit data to polynomial equations.

### Statistics

* + Descriptive Statistics: Calculate measures such as mean, median, mode,

standard deviation, and variance.

* Probability Distributions: Use functions like NORM.DIST, BINOM.DIST, and POISSON.DIST to work with probability distributions.
* Correlation and Regression Analysis: Use functions like CORREL and perform linear or multiple regression analysis.

### Matrix Operations

* + Matrix Multiplication: Perform matrix multiplication using the MMULT function.
  + Determinants and Inverses: Use the MDETERM and MINVERSE functions for matrix determinants and inverses, respectively.

### Calculus

* + Numerical Differentiation: Estimate derivatives by calculating differences in function values.
  + Numerical Integration: Approximate integrals using methods like the trapezoidal rule.

### Linear Algebra

* + Solving Linear Systems: Use Excel's Solver or matrix functions to solve systems of linear equations.
  + Eigenvalues and Eigenvectors: While more advanced, you can use Excel to approximate eigenvalues and eigenvectors with some manual work or add-ins.

### Optimization

* + Solver Add-In: Optimize functions subject to constraints, such as finding maximum or minimum values.

### Graphing and Visualization

* + Function Graphs: Plot functions and equations to visualize mathematical relationships.
  + Curve Fitting: Use scatter plots and trendlines to fit curves to data.

### Financial Mathematics

* + Compound Interest: Use functions like FV and PV to calculate future and present values.
  + Loan Calculations: Calculate loan payments, interest rates, and amortization schedules.

### Advanced Mathematics

* + Fourier Analysis: Use Excel to perform basic Fourier transforms for signal processing applications.

## CHAPTER 2 LITERATURE SURVEY

A literature survey on data analysis in Excel spreadsheets typically involves exploring various academic papers, articles, and books that discuss the methodologies, tools, and techniques used for analyzing data within Excel. Key areas of focus include the application of formulas and functions, the use of pivot tables for summarizing data, and the integration of statistical tools within Excel. The survey would also examine case studies where Excel has been utilized for data-driven decision-making, highlighting its strengths and limitations compared to more advanced data analysis software. Additionally, the literature may cover recent advancements in Excel, such as the incorporation of Power Query and Power Pivot, which enhance its data manipulation capabilities. This review provides a comprehensive understanding of how Excel continues to be a valuable tool for both basic and complex data analysis tasks in various fields.

## CHAPTER 3

**AIM AND OBJECTIVE OF SUPERMARKET DATA IN EXCEL SPREEDSHEET**

## Aim

The aim of analyzing supermarket data in an Excel spreadsheet is to systematically organize, process, and interpret the data to uncover insights that can improve business decision-making, enhance operational efficiency, and increase profitability**.**

## Objectives

1. **Data Organization:** Structure raw data in an Excel spreadsheet to facilitate easy access and analysis, ensuring data is clean, accurate, and properly formatted.
2. **Sales Analysis**: Track and analyze sales data to identify trends, peak sales periods, and best-selling products. This helps in optimizing inventory and sales strategies.
3. **Customer Insights:** Analyze customer data to understand purchasing behavior, preferences, and demographics, enabling targeted marketing and personalized customer experiences**.**
4. **Inventory Management**: Monitor stock levels, track inventory turnover, and identify slow-moving products to improve inventory management and reduce holding costs.
5. **Revenue and Profit Analysis:** Calculate and visualize key financial metrics like total revenue, profit margins, and cost of goods sold to assess the supermarket's financial health.
6. **Market Basket Analysis:** Perform market basket analysis to identify product combinations frequently purchased together, which can inform cross-selling strategies and promotional campaigns.
7. **Supplier Performance:** Evaluate supplier performance by tracking delivery times, order accuracy, and cost, helping to streamline procurement processes.
8. **Forecasting and Budgeting:** Use historical data to forecast future sales, demand, and expenses, aiding in budget planning and resource allocation.
9. **Operational Efficiency:** Identify bottlenecks and inefficiencies in operations through data-driven insights, leading to process improvements and cost savings.
10. **Reporting and Visualization:** Create reports and dashboards in Excel to visualize data, making it easier for stakeholders to understand and act upon the insights derived.

## CHAPTER 4

**PROJECT IN SUPERMARKET DATA**

Predicting sales in supermarkets involves using historical sales data, customer behavior patterns, and various external factors to forecast future sales trends. This process is crucial for effective inventory management, demand planning, and overall business strategy. By analyzing past data, supermarkets can anticipate customer demand, optimize stock levels, reduce waste, and improve profitability.

Using tools like Excel and more advanced techniques such as machine learning algorithms, businesses can develop predictive models that help in making informed decisions. These models consider factors like seasonality, promotions, and market trends, enabling supermarkets to stay competitive and meet customer needs efficiently. Predictive sales analysis is a key component in driving the operational success and financial growth of supermarkets..

### Data Collection

* + Gather Historical Data: Collect historical sales data, including daily or weekly sales figures, product categories, and transaction details.
  + Collect Related Data: Obtain additional data such as customer demographics, promotional activities, seasonality, and external factors like holidays or economic indicators.

### Data Cleaning and Preprocessing

* + Remove Inconsistencies: Clean the data by removing duplicates, handling missing values, and correcting any errors or anomalies.
  + Feature Engineering: Create new features or variables from existing data that may enhance the predictive power of the model (e.g., day of the week, holiday indicators).
  + Normalization/Standardization: Normalize or standardize the data if

required, especially for algorithms sensitive to data scales.

### Exploratory Data Analysis (EDA)

* + Visualize Data Trends: Use charts and graphs to identify patterns, trends, and relationships within the data, such as seasonality or correlation between variables.
  + Identify Key Drivers: Determine which factors most significantly impact sales, such as promotions, product categories, or customer segments**.**

### Model Selection

* + Choose a Predictive Model: Select appropriate statistical or machine learning models for sales prediction. Common models include:
    - Time Series Models: ARIMA, SARIMA, or Prophet for forecasting based on historical sales trends.
    - Regression Models: Linear regression, decision trees, or random forests for predicting sales based on multiple variables.
    - Advanced Techniques: Neural networks, XGBoost, or ensemble models for more complex prediction tasks.

### Model Training

* + Split Data: Divide the data into training and testing sets to evaluate model performance.
  + Train the Model: Use the training data to fit the selected model, adjusting parameters to optimize performance.
  + Cross-Validation: Implement cross-validation techniques to ensure the model generalizes well to unseen data.

### Model Evaluation

* + Evaluate Performance: Test the model on the testing set and evaluate its accuracy using metrics like Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R² score.
  + Refine the Model: Iterate on model selection and tuning to improve accuracy, possibly incorporating more features or trying different algorithms**.**

### Sales Forecasting

* + Generate Predictions: Use the trained model to forecast future sales, producing estimates for different time periods or product categories.
  + Scenario Analysis: Perform what-if analysis to understand how changes in certain variables (e.g., price changes, promotional activities) might impact sales.

### Visualization and Reporting

* + Visualize Forecasts: Create charts, graphs, and dashboards to visually represent the sales forecasts and make them accessible to decision- makers.
  + Prepare Reports: Summarize findings, predictions, and recommendations in a report that highlights key insights and actionable steps**.**

### Implementation and Monitoring

* + Implement Forecasts: Use the predictions to guide inventory management, marketing strategies, and other operational decisions.
  + Monitor Performance: Continuously monitor the accuracy of the predictions, adjusting models as necessary to account for new trends or data

### Feedback and Improvement

* + Incorporate Feedback: Regularly update the model with new data and feedback from the actual sales outcomes to improve its predictive accuracy**.**

## CHAPTER 5

**CASE STUDY AND ANALYSIS**

I have collected the data from Supremarket to predicting the sales of 16 products

I Analysis and found the nine items which sold more in that supermarket The three month’s dataset is given below:

### ANALYSED DATA

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Product Name** | **Cost Price (INR)** | **Selling Price (INR)** | **Quantity Sold** | **Total Sales (INR)** |
| Apples | 68.69 | 78.08 | 1318 | 102907.93 |
| Bananas | 143.59 | 166.69 | 1146 | 191026.22 |
| Bread | 115.16 | 138.76 | 520 | 72155.85 |
| Butter | 97.83 | 116.06 | 1340 | 155519.35 |
| Cheese | 40.28 | 46.66 | 666 | 31073.53 |
| Chicken | 140.72 | 162.38 | 991 | 160909.63 |
| Eggs | 85.21 | 97.99 | 1199 | 117547.01 |
| Fish | 142.67 | 168.32 | 1030 | 173375.17 |
| Milk | 57.45 | 65.93 | 1505 | 99234.97 |
| Oranges | 73.22 | 83.56 | 1482 | 123888.64 |
| Potatoes | 18.45 | 20.57 | 2136 | 43929.28 |
| Rice | 43.51 | 50.16 | 1128 | 56683.01 |
| Sugar | 28.67 | 33.04 | 1433 | 47358.36 |
| Tomatoes | 25.43 | 29.11 | 1247 | 36296.46 |
| Wheat | 40.93 | 46.85 | 1937 | 90772.98 |
| Yogurt | 48.25 | 57.03 | 1038 | 59228.58 |

**TABLE :4.1 Analysed Data**

## PIVOT TABLES

|  |  |  |
| --- | --- | --- |
| **Row Labels** | **Sum of Quantity**  **Sold** | **Sum of Total Sales**  **(INR)** |
| Apples | 1318 | 102907.93 |
| Bananas | 1146 | 191026.22 |
| Bread | 520 | 72155.85 |
| Butter | 1340 | 155519.35 |
| Cheese | 666 | 31073.53 |
| Chicken | 991 | 160909.63 |
| Eggs | 1199 | 117547.01 |
| Fish | 1030 | 173375.17 |
| Milk | 1505 | 99234.97 |
| Oranges | 1482 | 123888.64 |
| Potatoes | 2136 | 43929.28 |
| Rice | 1128 | 56683.01 |
| Sugar | 1433 | 47358.36 |
| Tomatoes | 1247 | 36296.46 |
| Wheat | 1937 | 90772.98 |
| Yogurt | 1038 | 59228.58 |
| **Grand Total** | **20116** | **1561906.97** |

### TABLE 4.2 : PIVOT TABLE

**TO FIND THE PROFIT OF PRODUCTS**

PROFIT = SELLING PRICE -COST PRICE

Explanation:

* Selling Price (SP): The price at which the product is sold to customers.
* Cost Price (CP): The price at which the product was purchased or the cost incurred to produce the product.

## Profit Calculation:

Profit for each item can be calculated as:

Profit for Each Product:

* + Apples: INR 12,376.02
  + Bananas: INR 26,472.60
  + Bread: INR 12,272.00
  + Butter: INR 24,428.20
  + Cheese: INR 4,249.08
  + Chicken: INR 21,465.06
  + Eggs: INR 15,323.22
  + Fish: INR 26,419.50
  + Milk: INR 12,762.40
  + Oranges: INR 15,323.88
  + Potatoes: INR 4,528.32
  + Rice: INR 7,501.20
  + Sugar: INR 6,262.21
  + Tomatoes: INR 4,588.96
  + Wheat: INR 11,467.04
  + Yogurt: INR 9,113.64

### Key Insights:

* Most Profitable Product: Bananas with a profit of INR 26,472.60.
* Least Profitable Product: Cheese with a profit of INR 4,249.08.

Bananas are providing the highest profit, while Cheese yields the lowest profit, indicating that more focus could be placed on products like Bananas to maximize profitability. Conversely, Cheese may need attention, whether in terms of pricing, marketing, or reducing costs

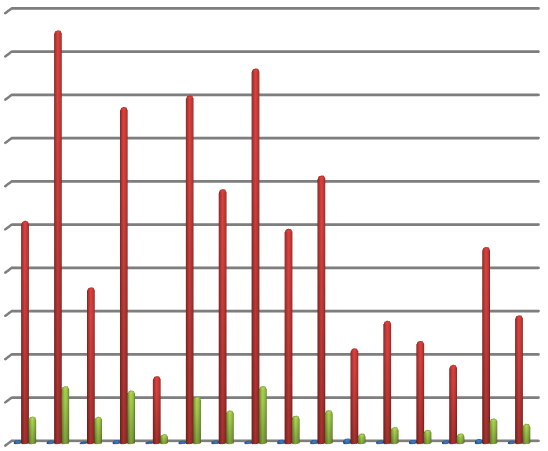
### PROFIT FOR ITEM

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Product Name** | **Cost Price**  **(INR)** | **Selling Price (INR)** | **Quantity Sold** | **Total Sales (INR)** | **Profit (INR)** |
| Apples | 68.69 | 78.08 | 1318 | 102907.93 | 12374.23 |
| Bananas | 143.59 | 166.69 | 1146 | 191026.22 | 26468.81 |
| Bread | 115.16 | 138.76 | 520 | 72155.85 | 12273.05 |
| Butter | 97.83 | 116.06 | 1340 | 155519.35 | 24433.05 |
| Cheese | 40.28 | 46.66 | 666 | 31073.53 | 4245.43 |
| Chicken | 140.72 | 162.38 | 991 | 160909.63 | 21548.55 |
| Eggs | 85.21 | 97.99 | 1199 | 117547.01 | 15185.43 |
| Fish | 142.67 | 168.32 | 1030 | 173375.17 | 26448.45 |
| Milk | 57.45 | 65.93 | 1505 | 99234.97 | 12739.45 |
| Oranges | 73.22 | 83.56 | 1482 | 123888.64 | 15323.32 |
| Potatoes | 18.45 | 20.57 | 2136 | 43929.28 | 4540.32 |
| Rice | 43.51 | 50.16 | 1128 | 56683.01 | 7504.35 |
| Sugar | 28.67 | 33.04 | 1433 | 47358.36 | 6264.49 |
| Tomatoes | 25.43 | 29.11 | 1247 | 36296.46 | 4584.44 |
| Wheat | 40.93 | 46.85 | 1937 | 90772.98 | 11484.19 |
| Yogurt | 48.25 | 57.03 | 1038 | 59228.58 | 9058.96 |

**TABLE 4.3: PROFIT FOR ITEMS**

To determine which items are giving a profit to the salesman based on the data provided, we can compare the "cost price" with the "selling price" If the "selling price" for an item is higher than the "cost price" then the item is likely generating a profit.By showing in graph visualization

## GRAPHS



200000

180000

160000

140000

120000

100000

80000

60000

40000

20000

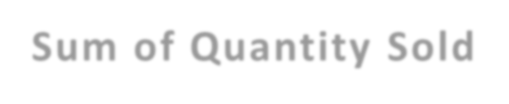
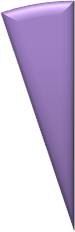
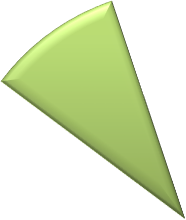
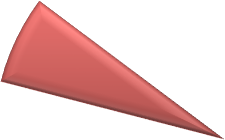
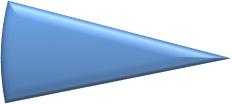
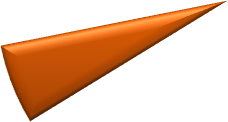
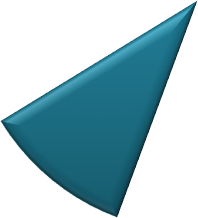
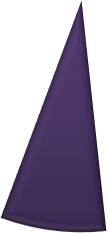
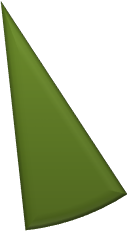
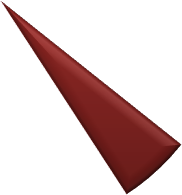
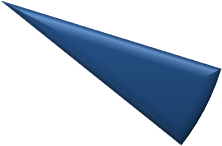
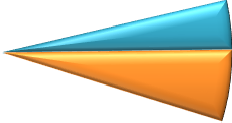
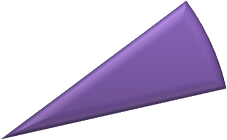
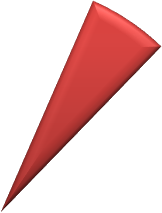
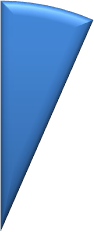
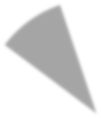
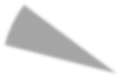
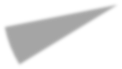
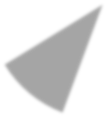
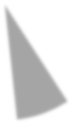
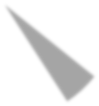
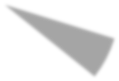
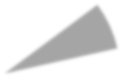
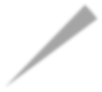
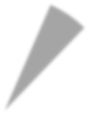
0

Sum of Quantity Sold

Sum of Total Sales (INR)

Sum of Profit (INR)

**4.4:BAR GRAPH**



**Sum of Quantity Sold**

5%

7%

Apples

10%

6%

Bananas

3%

Bread

6%

7%

Butter

7%

3%

5%

6%

6%

Cheese

Chicken Eggs Fish

11%

5%

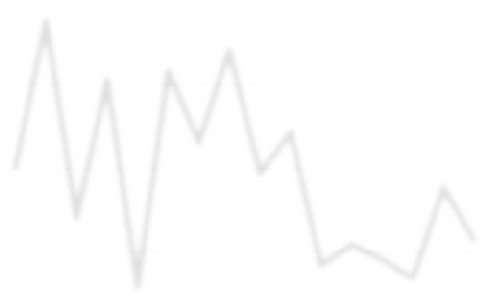
Milk

7%

7%

Oranges

## 4.5:PIE GRAPH



250000

200000

150000

Sum of Quantity Sold

100000

Sum of Total Sales (INR)

Sum of Profit (INR)

50000

0

**4.6:LINE GRAPH**

Apples

Bananas Bread Butter Cheese Chicken

Eggs Fish Milk Oranges

Potatoes

Rice Sugar Tomatoes

Wheat Yogurt

## CHAPTER 6 CONCLUSION

This dataset provides a detailed overview of the sales performance for various products, focusing on their cost price, selling price, quantity sold, total sales, and profit. It is evident that all products generated a profit, as the selling prices exceed the cost prices in each case. Among the products, Fish and Bananas yielded the highest profits, each contributing significantly to the total revenue, indicating their strong demand and efficient pricing strategy. On the other hand, despite having lower cost prices, products like Potatoes and Sugar also contributed positively to overall profitability, albeit with lower profit margins. This analysis underscores the importance of balancing cost and pricing strategies to maximize profitability across different product categories., Butter, with a total sales figure of INR 155,519.35 and a profit of INR 24,433.05, and Bananas, with total sales of INR 191,026.22 and a profit of INR 26,468.81, are among the highest profit-generating products. On the other hand, Potatoes, despite being sold in large quantities (2,136 units), generated a relatively lower profit of INR 4,540.32 due to their low cost and selling prices.. Products with higher selling prices like Bananas and Fish tend to yield more significant profits, while essential staples like Potatoes and Tomatoes, despite their lower prices, contribute to steady sales with moderate profits. This analysis underscores the need for a balanced product mix, ensuring that high-margin items are complemented by volume-driven staples to maximize overall profitability.The salesman should focus on maintaining or even increasing the stock and sales of Butter, Bananas, and Fish, as these items contribute the most to overall profits. Additionally, while Potatoes and Cheese are less profitable, they still add value to the sales portfolio and can be strategically promoted to increase their contribution to total profits.

**INTRENSHIP CERTIFICATE**





## REFERENCES:

* Alexander, M. (2016). Excel 2016 Power Programming with VBA. John Wiley & Sons.
* Berk, K. N., & Carey, P. (2010). Data Analysis with Microsoft Excel: Updated for Office 2007. Cengage Learning.
* Bill Jelen, & Tracy Syrstad. (2019). Excel 2019 VBA and Macros. Microsoft Press.
* Bullen, S., Bovey, R., & Green, J. (2010). Professional Excel Development: The Definitive Guide to Developing Applications Using Microsoft Excel, VBA, and .NET. Addison-Wesley.
* Frye, C. (2018). Excel 2019 All-in-One For Dummies. Wiley.
* Géron, A. (2019). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media. (Note: Though this book primarily focuses on machine learning, it covers data preparation and analysis techniques that are often used in Excel.)
* Gross, D. (2020). Excel 2020: A Complete Guide to Microsoft Excel, Including Tips, Tricks, and Shortcuts! Independently Published.
* Khan, N. (2016). Excel Data Analysis For Dummies. Wiley.
* McFedries, P. (2019). Excel Data Analysis: Your visual blueprint for creating and analyzing data, charts, and PivotTables. Wiley.
* Winston, W. (2016). Microsoft Excel 2016 Data Analysis and Business Modeling. Microsoft Press