# INTRODUCTION

Data Analytics is a cornerstone of modern business operations, enabling Organizations to derive insights from complex datasets. With the exponential growth of e-commerce and retail platforms, business face increasing challenges in understanding sales dynamics and customer prefernces.

This study address these challenges by creating an innovative data anlaysis framework tailored for sales data. By integrating business intelligence tools, such power BI, with robust data preprocessing and visualization techniques, this research provide comprehensive solution to identify sales patterns, predict future trends, and streamline decision-making process.

Data-driven sales analysis and visualization is the process of using data to better understand how a business is performing, particularly in terms of sales, and then presenting that information in visual formats like charts or graphs. This helps businesses make smarter decisions based on real insights from their data.

With the increasing volume of data generated across industries, businesses and organizations need efficient solutions to interpret and utilize data effectively. The Data Analysis Dashboard enables users to convert raw data into actionable insights through visual representation and statistical computations.

Visualization tools then translate these numerical findings into clear and concise charts and graphs, making complex data easily digestible. This combination allows businesses to make informed decisions regarding inventory management, marketing strategies, and customer engagement, ultimately leading to increased sales, improved customer satisfaction, and optimized operational efficiency.

## 1.1 AI & ML Technology

Data analysis and visualization are essential components of data-driven decision making, particularly in fields like e-commerce, retail, healthcare, and finance. By using advanced technologies, businesses can extract valuable insights from vast datasets and present them in visually compelling formats to aid in better decision making.

Here, we will explore techonologies key used for data analysis and visualization, including tools,techniques and applications .

Data analysis and visualization technologies are transforming the way businesses interact with data. Whether it's through advanced tools like Power BI and Tableau for creating interactive dashboards or Python libraries like Matplotlib for customized visualizations, these technologies allow businesses to gain deep insights, make data-driven decisions, and stay competitive.

Finally, business intelligence platforms and visualization libraries translate these analytical findings into interactive dashboards and reports, enabling stakeholders to readily comprehend and utilize the insights for informed decision-making. This integrated technological approach empowers businesses to optimize their sales strategies, enhance customer experiences, and drive overall growth.

The project integrates multiple technologies, including:

∙ TAData Processing: Python (Pandas, NumPy)

∙ Visualization: Power BI, Tableau, Matplotlib

∙ Machine Learning: Scikit-Learn, TensorFlow

∙ Database Management: SQL, NoSQL

**1.2. SYSTEM SPECIFICATION**

**1.2.1 Hardware Configuration**

## 

|  |  |
| --- | --- |
| **Feature** | **Requirements** |
| Processor(CPU) | Intel Core i5 or  AMD Ryzen 5  (Multi-core) |
| RAM(memory) | 8GB-16GB |
| Storage | 256GB-512GB SSD |

## 1.2.2 Software Configuration

## 

|  |  |
| --- | --- |
| **Features** | **Requirements** |
| IDE | Jupter Notebook,Visual Studio Code |
| Programming languages | Python,SQL |
| Tools | Power BI, Tableau |
| Browser | Google Chrome |

## 2. SYSTEM STUDY

### 2.1 Existing System

Traditional data analysis relies on spreadsheets and manual reporting, which are prone to errors, lack automation, and offer limited scalability.

A current system of data-driven sales analysis and visualization will generally leverage multiple data sources, including customer relationship management(CRM)applications, transactional databases, and external market information.With the applications of visualizations such as interactive dashboards, heatmaps, and and trend charts, salesforces and management can easily trends, monitor KPI’s and indentify actionable insights.These software programs intergrate predict future sales patterns and customers activites.

Data-driven sales analysis and visualization are fundamental to modern e commerce and retail, enabling businesses to move beyond guesswork and embrace informed decision-making. Existing systems excel at collecting vast datasets from sales transactions, website interactions, and customer demographics. This data is then processed through analytical techniques, such as trend analysis, customer segmentation, and inventory forecasting, to uncover meaningful patterns.

These insights are made readily accessible through intuitive visualizations like interactive dashboards, charts, and graphs, empowering businesses to optimize pricing, personalize marketing, improve inventory management, and ultimately enhance the customer experience. By leveraging these tools, retailers and

commerce platforms can identify emerging trends, predict future sales, and gain a competitive edge in a dynamic market.

## 2.1.1Drawbacks

∙ Time-consuming manual data processing

∙ Limited visualization capabilities

∙ Lack of real-time insights

∙ Inefficient data handling for large datasets

### 2.2 Proposed System

A proposed system for data-driven sales analysis and visualization aims to revolutionize e-commerce and retail decision-making by providing actionable insights through the effective handling and presentation of sales data. This system would involve:

#### 2.2.1 Data Aggregation and Integration

∙ It would consolidate data from diverse sources, including online sales platforms, point-of-sale systems, customer relationship management (CRM) databases, and social media.

∙ This integration creates a unified view of sales performance, customer behavior, and market trends.

#### 2.2.2 Interactive Visualization

∙ Data visualization tools would transform complex data into easily understandable charts, graphs, and dashboards.

∙ Interactive dashboards would allow users to explore data, drill down into specific metrics, and gain a comprehensive understanding of sales performance.

∙ These visual representations would help decision makers quickly grasp key performance indicators.

#### 2.2.3 Real-Time Monitoring

#### 

∙ The system would provide real-time monitoring of sales data, enabling businesses to react quickly to changing market conditions and customer demands.

∙ Alerts and notifications would be triggered by significant changes in sales patterns, allowing for proactive intervention.

#### 2.2.4 Enhanced Decision-Making

∙ By providing clear and actionable insights, the system would empower e commerce and retail businesses to make informed decisions regarding product pricing, inventory management, marketing campaigns, and customer service.

∙ This would ultimately lead to increased sales, improved customer satisfaction, and enhanced profitability.

∙ Real-time data integration

∙ Interactive visual analytics

∙ Machine learning-powered insights

∙ Predictive analytics for future trend analysis

#### 2.2.1 Features

**2.2.1.1. Sales Performance Analysis:**

∙ Real-time sales tracking and monitoring.

∙ Sales trend analysis (e.g., daily, weekly, monthly, yearly).Sales performance by product, category, region, and channel.

∙ Key performance indicator (KPI) tracking (e.g., sales revenue, conversion rate, average order value).Cohort analysis.

**2.2.1.2 Customer Behaviour Analysis**

∙ Customer segmentation based on demographics, purchase history, and behavior.

∙ Customer lifetime value (CLTV) calculation.

∙ Purchase pattern analysis (e.g., frequency, recency, monetary value). ∙

Customer churn analysis.

∙ Basket analysis.

**2.2.1.3 Inventory and Product Analysis**

∙ Inventory level tracking and forecasting.

∙ Product performance analysis (e.g., best-selling products, slow-moving products).

∙ Product bundling and cross-selling recommendations.

∙ Stock Keeping Unit(SKU) level analysis.

**2.2.1.4** **Marketing Campaign Analysis**

∙ Campaign performance tracking (e.g., ROI, conversion rate). ∙

Customer acquisition cost (CAC) analysis.

∙ Marketing channel effectiveness analysis.

∙ A/B testing analysis.

**2.2.1.5 Predictive Analytics**

∙ Sales forecasting based on historical data and trends.

∙ Demand forecasting for inventory management.

∙ Customer churn prediction.

∙ Anomaly detection.

**2.2.1.6 Data Visualization and Reporting**

∙ Interactive dashboards with customizable views.

∙ Variety of visualization tools (e.g., charts, graphs, maps).

∙ Automated report generation and distribution.

∙ Drill-down capabilities for detailed analysis.

**2.2.1.7 Alerts and Notifications**

∙ Real-time alerts for significant changes in sales patterns.

∙ Customizable alerts for specific KPIs.

∙ Automated notifications for inventory thresholds.

**2.2.1.8 User Management and Security**

∙ Role-based access control.

∙ Data encryption and security measures.

∙ Audit trails.

**2.2.1.9 Mobile Accessibility**

∙ Mobile friendly dashboards and reports.

∙ Ability to monitor sales data on mobile devices.

∙ Dynamic Data Visualization: Interactive charts, graphs, and tables ∙

Automated Reporting: Scheduled reports and alerts

∙ Data Filtering & Customization: User-defined parameters for tailored

analysis

∙ Machine Learning Integration: Predictive modeling for decision-making ∙ Scalability & Adaptability: Suitable for different industries

#### 2.2.2 Modules

Data driven sales visualizations and analysis are of fundamental importance to inform decision-making in retail and e-commerce businesses. Using the force of data, businesses can access insights about customers, product, market, and efficiency.

commerce platforms can identify emerging trends, predict future sales, and gain a competitive edge in a dynamic market.

**2.2.2.1 segement sales analysis**

∙ Data-driven Insights: By analyzing sales data, businesses can track how products are performing over time, identify top-selling items, and recognize underperforming products.

∙ Key Metrics: Average order value (AOV), conversion rates, customer acquisition costs (CAC), and lifetime value (LTV) can be calculated and visualized in charts, heatmaps, or bar graphs.

∙ Visualization Tools: Tools like Power BI, Tableau, and Google Data Studio can create interactive dashboards that display real-time sales data for easy interpretation

**2.2.2.2 Customer segmentation**.

∙ Understanding Customer Behaviour: By analyzing customer demographics, buying patterns, and preferences, businesses can segment their customer base into distinct groups.

∙ Personalized Marketing: This segmentation enables targeted marketing campaigns, personalized product recommendations, and dynamic pricing strategies that resonate with specific customer segments.

∙ Visualization: Customer segmentation can be visualized through pie charts, scatter plots, and geographic heat maps, helping businesses identify key markets or regions for growth.

##### 2.2.2.3. Product performance and Inventory Management

∙ Inventory Optimization: Data-driven analysis helps predict demand for different products based on historical sales data, seasonality, and market trends, which improves inventory management.

∙ Product Lifecycle: By tracking the lifecycle of a product, businesses can determine when to discount, promote, or phase out a product. ∙ Visualization: Inventory levels, reorder points, and stock-outs can be visualized using Gantt charts or line graphs.

##### 2.2.2.4 Predictive Analytics

∙ Forecasting Sales: Predictive analytics, based on historical sales data and external factors (e.g., weather, promotions, holidays), helps forecast future sales and demand.

**3. SYSTEM DESIGN**

### 3.1 User Interface (Form Design)

∙ Dashboard UI: A web-based or application-based dashboard with a clean and intuitive layout.

∙ Navigation Panel: Users can switch between different analytics sections (e.g., Reports, Predictions, Data Sources).

∙ Filters & Controls: Options to filter datasets, apply date ranges, and customize visualizations.

∙ Interactive Elements: Drill-down capabilities for deeper insights into specific datasets.

### 3.2 Data Input Design

∙ File Upload: Users can upload CSV, Excel, or JSON files for analysis.

∙ Database Connectivity: Real-time data ingestion from SQL/NoSQL databases and APIs.

∙ Automated Data Pipeline: Preprocessing and cleansing of raw data before analysis.

∙ Integration with IoT and Streaming Data: (For real-time analytics) Data is fetched dynamically from external sources.

## 3.3 Data Processing & Output Design

Analytical Engine:

∙ Statistical models for trend analysis.

∙ Machine Learning algorithms for predictive analytics.

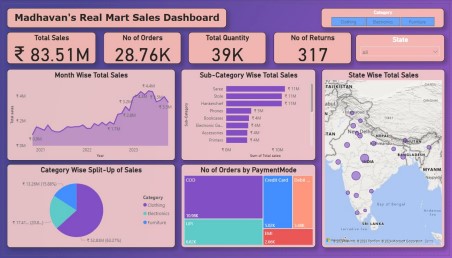
∙ Anomaly detection for data inconsistencies.

Visualization Layer: Charts, graphs, and heatmaps powered by Power BI/Tableau/Matplotlib.

∙ Real-time dynamic updates based on new data input.

∙ Export reports in PDF, Excel, and shareable links.

User Insights & Reporting: Automated report generation with key insights. ∙ Scheduled report emails and alert notifications.



## Fig:1

## 3.4 Database Design

∙ Data Warehouse Architecture

∙ Raw Data Table stores unprocessed user data.

∙ User Preferences Table stores user settings and customizations. ∙ Analytics Table stores computed insights and predicitons.

∙ User Preferences Table stores user settings and customizations.

**3.5 SYSTEM DEVELOPMENT**

**3.5.1. Description of Modules**

1. **Data Collection and Import Module:**

This module is responsible for gathering sales-related data from multiple sources such as transaction databases, e-commerce platforms, or CSV/Excel files. The imported data may include attributes like customer details, product categories, sales amounts, dates, and regions. The module ensures proper data loading using tools such as pandas for structured handling.

1. **Data Preprocessing and Cleaning Module:**

This module prepares the raw sales dataset for analysis. It includes handling missing values, removing duplicates, correcting data types, normalizing numerical features, and encoding categorical variables such as product categories or customer segments. Proper preprocessing ensures the dataset is consistent and ready for meaningful analysis.

1. **Exploratory Data Analysis (EDA) Module:**

This module enables initial exploration of the dataset to uncover patterns and trends. It includes generating descriptive statistics, correlation heatmaps, and summary reports. Visualization tools like matplotlib and seaborn are used to identify seasonality, customer preferences, and sales distribution across regions or products.

1. **Sales Forecasting and Machine Learning Module:**

This module applies statistical models and machine learning algorithms to forecast future sales. It includes splitting the dataset into training and testing sets, training models such as Linear Regression, Decision Trees, or ARIMA, and evaluating them for predictive performance. The module helps in demand prediction and inventory planning.

1. **Customer Segmentation Module:**

This module applies clustering techniques (e.g., K-Means or Hierarchical Clustering) to group customers based on purchasing behavior, frequency, and revenue contribution. It allows businesses to identify high-value customers, churn risks, and tailor marketing strategies accordingly.

1. **Visualization and Reporting Module:**

This module focuses on presenting insights in an interactive and user-friendly way. It includes generating dashboards, sales trend graphs, product performance comparisons, and heatmaps of regional sales. Visualization libraries like matplotlib, seaborn, or BI tools (e.g., Tableau, Power BI) are used to support decision-making.

1. **User Interaction and Query Module:**

This module provides a simple interface for managers and analysts to query specific sales insights. For example, a user may input a product ID or time range to retrieve sales performance or predictions. It ensures non-technical stakeholders can access insights without coding.

1. **System Integration and Workflow Control Module:**

This module integrates all other modules and manages the workflow of the sales analysis system. It controls the sequence of tasks (data collection → preprocessing → modeling → visualization), ensures data is passed correctly across modules, and maintains system efficiency for continuous updates and analysis.

**4. SOFTWARE TESTING AND IMPLEMENTATION**

**SOFTWARE TESTING AND RESULTS**

The testing phase evaluates the accuracy and effectiveness of the implemented data-driven sales analysis models in uncovering insights, predicting trends, and assisting decision-making. The evaluation process focuses on quantitative performance measures, model-specific results, and visualization-driven insights.

**1. Performance Metrics**

The following standard metrics are used to evaluate the performance of forecasting and predictive models:

* 1. **Accuracy Score:**

Measures how well the predictive model (e.g., regression or classification) matches actual sales outcomes.

* 1. **Mean Absolute Error (MAE):**

Represents the average magnitude of errors in sales predictions, providing a measure of prediction reliability.

* 1. **Root Mean Squared Error (RMSE):**

Provides a penalized view of large deviations between predicted and actual values, ensuring the robustness of the forecasting model.

* 1. **R² Score (Coefficient of Determination):**

Indicates how well the independent variables explain the variation in sales outcomes.

**1.5 Visualization of Forecast Results:**

Graphical comparisons (line plots, bar charts) of actual vs. predicted sales provide intuitive insight into model performance.

**1. Model-Specific Results**

**1.1** **Sales Forecasting Model Results:**

Regression models (e.g., Linear Regression) produced quantitative

forecasts for future sales based on historical data.

Accuracy metrics (R², MAE, RMSE) were calculated and compared across

**1.2 Customer Segmentation Results:**

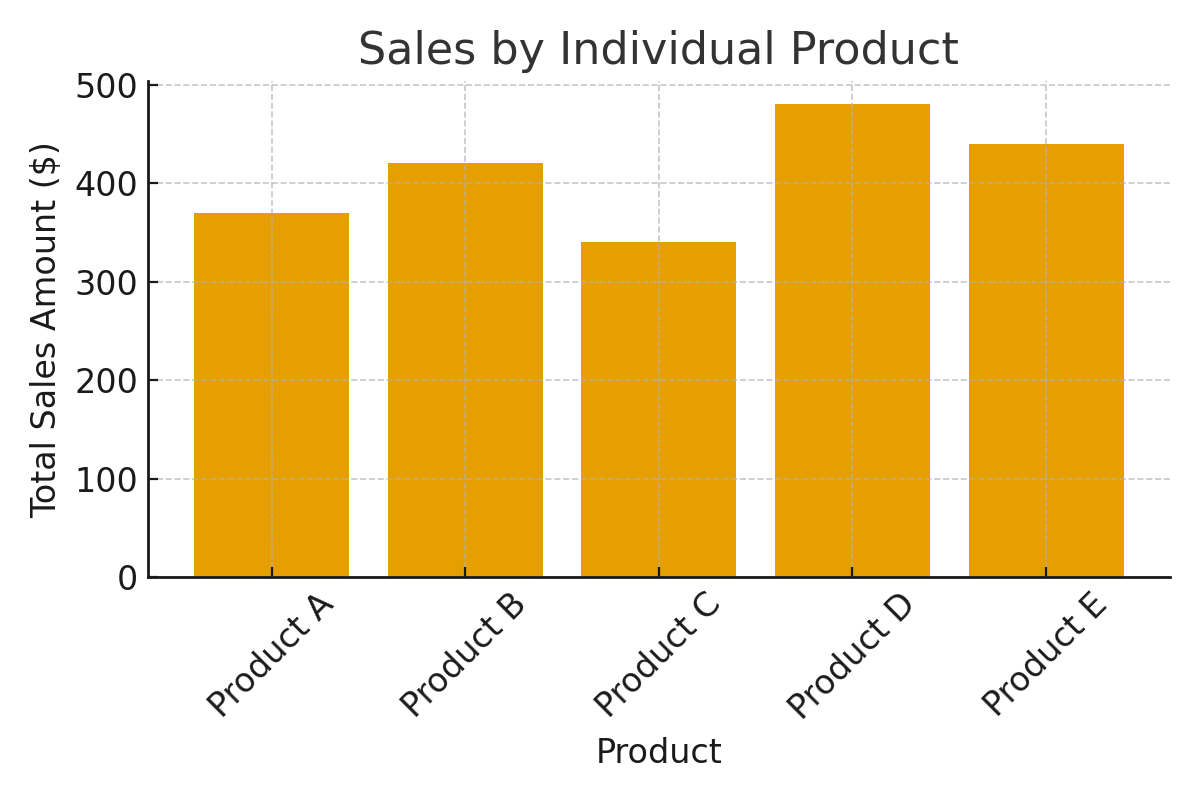
* + - Clustering models (e.g., K-Means) effectively grouped customers into categories such as high-value, frequent buyers, and churn risks.

* + - Visualization through scatter plots and cluster heatmaps demonstrated clear customer behavior patterns.

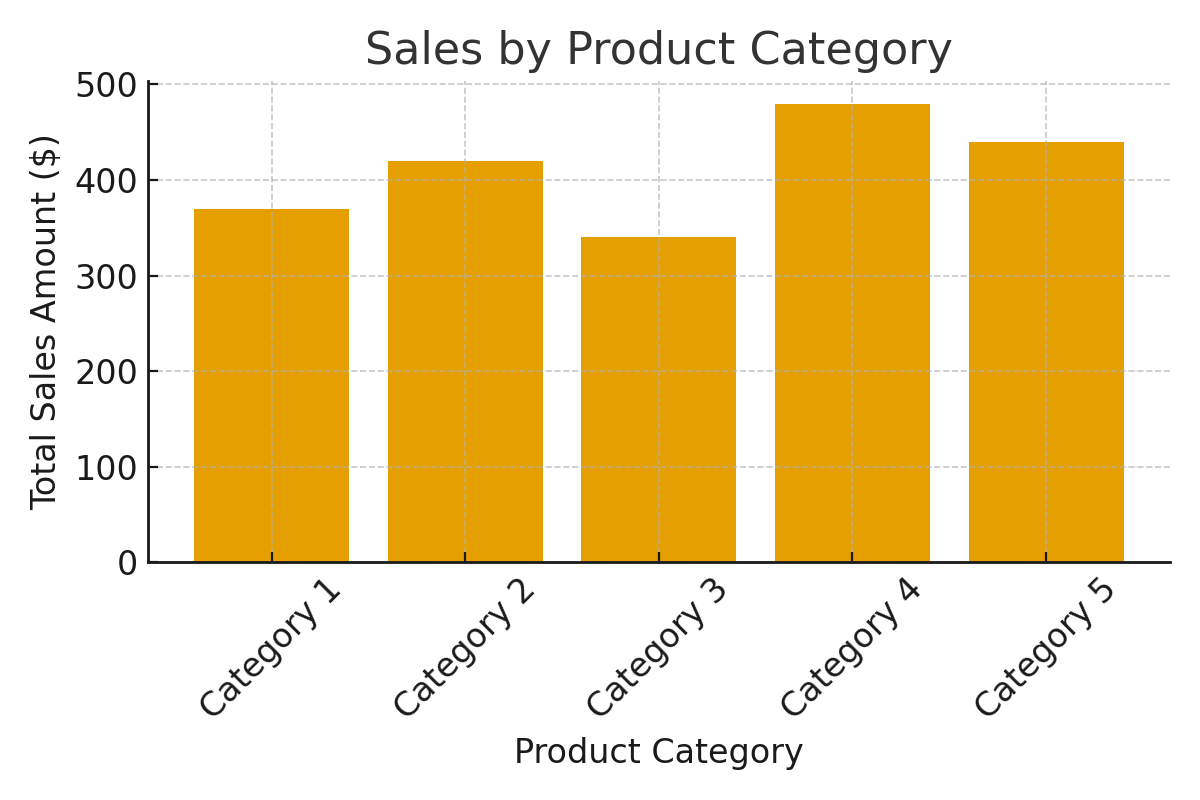
**1.3 Visualization Dashboard:**

* + - Interactive dashboards were generated using matplotlib, seaborn, or BI tools to display trends in sales performance, customer demographics, and regional growth.

* + - Graphical summaries (heatmaps, bar plots, pie charts) highlighted top-performing products, seasonal variations, and revenue drivers.



**Fig: 2**



**Fig: 3**

Simulation Graph: Existing Work VS Proposed Work

## ****Algorithm and Mathematical Model****

The proposed system uses **Linear Regression**and**K-Means Clustering** as the core analytical models. These algorithms are applied to predict sales trends and group customer data into meaningful clusters for better business decision-making.

### ****Linear Regression****

Linear Regression is used to establish a relationship between dependent and independent variables. In this system, the **dependent variable (Y)** represents the predicted sales amount, while **independent variables (X₁, X₂, …, Xₙ)** represent factors such as product category, quantity, or time period.  
The mathematical equation for the model is expressed as:

Y = β₀ + β₁X₁ + β₂X₂ + ⋯ + βₙXₙ + ε

Where:

* = Predicted Sales
* = Intercept of the line
* = Coefficients of predictors
* = Input features (e.g., time, category, quantity)
* = Error term

The goal is to minimize the **Mean Squared Error (MSE)** to achieve the best-fit line that predicts sales values accurately. The formula for MSE is:

MSE = \frac{1}{n} \sum\_{i=1}^{n} (Y\_i - Ŷ\_i)^2

Where is the actual sales and is the predicted sales.

### ****(ii) K-Means Clustering****

K-Means Clustering is used to group similar customers or products based on purchasing behavior and sales performance. It works by minimizing the **Within-Cluster Sum of Squares (WCSS)**, represented mathematically as:

WCSS = \sum\_{i=1}^{k} \sum\_{x \in C\_i} \|x - μ\_i\|^2

Where:

* = Number of clusters
* = Cluster
* = Mean of cluster
* = Data point

This helps identify customer segments or product groups with similar characteristics, supporting targeted marketing and sales optimization.

**IMPLEMENTATION PHASE**

The implementation phase of the Data-Driven Sales Analysis and Visualization System was executed primarily in Python, supported by libraries such as Pandas,num py ,matplotlib,seaborn and scikit-learn.

Environment Setup:

Development was carried out in environments like Jupyter Notebook and Google Colab, ensuring access to required datasets (CSV/Excel files).

**1. Data Loading and Preprocessing:**

**○** Historical sales data was imported from structured sources (e.g.,SALES \_DATASET.CSV).

○ Data cleaning steps included handling missing values, removing duplicates, and normalizing numerical features.

○ Categorical features such as product categories and customer segments were encoded into numerical formats for analysis.

**2. Model Development and Training:**

○ Predictive models for sales forecasting (Linear Regression, Decision Trees) were trained using historical data.

○ Customer segmentation was performed using clustering algorithms.

○ Train-test splitting was implemented to ensure unbiased performance evaluation.

**3. Prediction and Forecasting:**

○ The trained models predicted future sales volumes and revenue.

○ Insights were generated to assist in inventory planning, marketing strategy, and

Demand forecasting.

**4. Visualization and User Interaction:**

○ A visualization layer was implemented using plots, charts, and dashboards.

○ Users can query sales performance by product, region, or time period to retrieve insights in real time.

**5. System Integration:**

○ Modules for data collection, preprocessing, analysis, forecasting, and visualization were integrated into a single workflow.

○ The system ensures smooth execution from raw data input to decision-support

Making.

**CONCULSION**

The Data Analysis Dashboard is an innovative solution for streamlining data analysis and visualization. By integrating modern data science tools and machine learning, the system empowers users with powerful analytics capabilities. The dashboard improves decision-making efficiency and provides a scalable solution adaptable across various industries.

By transforming raw data into actionable insights, businesses can optimize operations, personalize customer experiences, and achieve significant competitive advantages. This shift towards data-centric strategies empowers organizations to navigate market complexities with greater agility and precision, ultimately driving sustainable growth and profitability in the dynamic landscape of modern commerce.

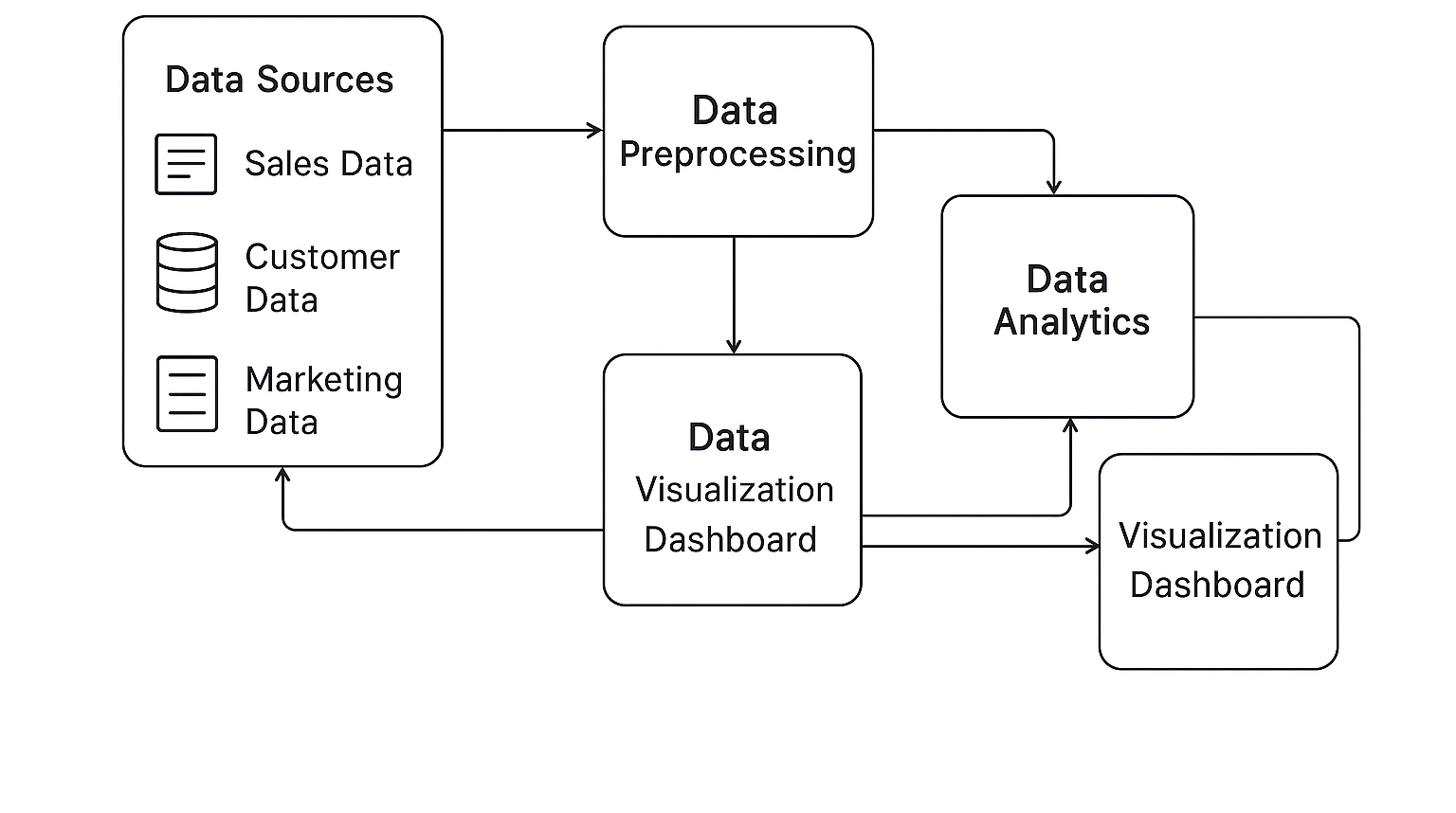
**BIBLIOGRAPHY**

1. "Data Science Handbook: A Guide for Beginners" - Available at: https://www.datascience.com/resources
2. "Machine Learning for Data Analysis" - Available at: https://www.ml4data.com/freebook
3. "Introduction to Power BI for Data Visualization" - Available at: https://powerbi.microsoft.com/en-us/learning/
4. "SQL for Data Analysis" - Available at: https://www.sqlzoo.net/
5. Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS quarterly*, 36(4), 1165-1188
6. Davenport, T. H., & Harris, J. G. (2007). *Competing on analytics: The new science of winning*. Harvard Business School Press..
7. Few, S. (2012). *Show me the numbers: Designing tables and graphs to enlighten*. Analytics Press.

**APPENDICES**

### Data Flow Diagram

### 

****

### B. Table Structure

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Order ID** | **Sale ID** | **Date** | **Hours** | **Order** |
| 2854884835480 | 8725409595544 | 11-11-2020 | 01:39:33 AM | TS1011 |
| 2864488448152 | 8752046604440 | 16-11-2020 | 04:43:01 PM | TS1012 |
| 2864491954328 | 8752055845016 | 16-11-2020 | 04:45:58 PM | TS1013 |
| 2864491954328 | 8752055877784 | 16-11-2020 | 04:45:58 PM | TS1013 |
| 2864497066136 | 8752070656152 | 16-11-2020 | 04:49:47 PM | TS1014 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Fulfilment** | **ship-service-level** | **Transaction** | **Sale type** | **Sales channel** |
|  |  | type |  |  |
| Merchant | Standard | product | Order | Online Store |
| Merchant | Standard | product | Order | Point of Sale |
| Amazon | Expedited | product | Order | Point of Sale |
| Merchant | Standard | product | Order | Point of Sale |
| Amazon | Expedited | product | Order | Point of Sale |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Net quantity** | **Gross sales** | **Discounts** | **Net sales** | **Shipping** |
| 1 | 200 | 0 | 200 | 0 |
| 2 | 9000 | 0 | 9000 | 0 |
| 1 | 1800 | 0 | 1800 | 0 |
| 1 | 1500 | 0 | 1500 | 0 |
| 1 | 3000 | 0 | 3000 | 0 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Taxes** | **Total sales** | **PaymentMode** | **Category** | **Sub-Category** | **State** | **Returns** |
| 0 | 200 | COD | Electronics | Electronic  Games | Uttar  Pradesh | 0 |
| 0 | 9000 | EMI | Furniture | Chairs | Delhi | 0 |
| 0 | 1800 | EMI | Furniture | Bookcases | Uttar  Pradesh | 0 |
| 0 | 1500 | Credit Card | Electronics | Printers | Maharashtra | 0 |
| 0 | 3000 | Credit Card | Electronics | Phones | Madhya Pradesh | 0 |

### C. Sample code

import pandas as pd import matplotlib.pyplot as plt import seaborn as sns

# Sample data (this would be your actual sales data) data = {

'Date': pd.date\_range(start='2023-01-01', periods=100, freq='D'), 'Product':

['Product A', 'Product B', 'Product C', 'Product D', 'Product E'] \* 20,

'Category': ['Category 1', 'Category 2', 'Category 3', 'Category 4', 'Category 5'] \* 20,

'Sales\_Amount': [120, 200, 150, 180, 210, 250, 220, 190, 300, 230] \* 10, 'Quantity': [5, 7, 6, 8, 4, 10, 9, 5, 12, 6] \* 10

}

# Creating DataFrame df = pd.DataFrame(data) # Data Preprocessing df['Date'] = pd.to\_datetime(df['Date'])

# Grouping by date for daily sales trend analysis daily\_sales = df.groupby('Date')['Sales\_Amount'].sum().reset\_index() # Grouping by product category for product sales comparison category\_sales = df.groupby('Category')['Sales\_Amount'].sum().reset\_index() # Grouping by product for individual product sales comparison product\_sales = df.groupby('Product')['Sales\_Amount'].sum().reset\_index() # Visualization 1: Daily Sales Trend plt.figure(figsize=(10, 6)) sns.lineplot(x='Date', y='Sales\_Amount', data=daily\_sales, marker='o') plt.title('Daily Sales Trend') plt.xlabel('Date') plt.ylabel('Total Sales Amount ($)') plt.xticks(rotation=45) plt.grid(True) plt.tight\_layout() plt.show()

# Visualization 2: Sales by Product Category

plt.figure(figsize=(10, 6))

sns.barplot(x='Category', y='Sales\_Amount', data=category\_sales,

palette='viridis') plt.title('Sales by Product Category') plt.xlabel('Product Category') plt.ylabel('Total Sales Amount ($)')

plt.xticks(rotation=45) plt.tight\_layout() plt.show()

# Visualization 3: Sales by Individual Product plt.figure(figsize=(10, 6)) sns.barplot(x='Product', y='Sales\_Amount', data=product\_sales, palette='Set2') plt.title('Sales by Individual Product') plt.xlabel('Product') plt.ylabel('Total Sales Amount ($)') plt.xticks(rotation=45) plt.tight\_layout() plt.show()

# Analysis for Decision Making # Top 5 highest-grossing products top\_products = product\_sales.nlargest(5, 'Sales\_Amount') print("Top 5 Products by Sales Amount:") print(top\_products)

# Total sales comparison between different categories total\_category\_sales = category\_sales['Sales\_Amount'].sum()

category\_sales['Percentage\_of\_Total'] = (category\_sales['Sales\_Amount'] / total\_category\_sales) \* 100 print("\nSales Percentage by Category:") print(category\_sales)

**D.SAMPLE INPUT**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Date** | **Product** | **Category** | **Sales\_ Amount** | **Quantity** |
| 2023-01-01 | Product A | Category 1 | 120 | 5 |
| 2023-01-02 | Product B | Category 2 | 200 | 7 |
| 2023-01-04 | Product C | Category 3 | 150 | 6 |
| 2023-01-05 | Product D | Category 4 | 180 | 8 |
| 2023-01-06 | Product E | Category 5 | 210 | 4 |
| 2023-01-07 | Product A | Category 1 | 250 | 10 |
| 2023-01-08 | Product B | Category 2 | 220 | 9 |
| 2023-01-09 | Product C | Category 3 | 190 | 5 |
| 2023-01-10 | Product D | Category 4 | 300 | 12 |
| 2023-01-11 | Product E | Category 5 | 230 | 6 |

**E. SAMPLE OUTPUT**

**1. TOTAL SALES BY PRODUCT**

|  |  |
| --- | --- |
| **Product** | **Total\_ Sales** |
| Product A | 370 |
| Product B | 420 |
| Product C | 340 |
| Product D | 480 |
| Product E | 440 |

**2. AVERAGE SALES PER CATEGORY**

|  |  |
| --- | --- |
| **Category** | **Avg \_ sales** |
| Category 1 | 185 |
| Category 2 | 210 |
| Category 3 | 170 |
| Category 4 | 240 |
| Category 5 | 220 |

**3. TOTAL QUANTITY SALES PER PRODUCT**

|  |  |
| --- | --- |
| **Product** | **Total\_ Quantity** |
| Product A | 15 |
| Product B | 16 |
| Product C | 11 |
| Product D | 20 |
| Product E | 10 |