ANUDIP FOUNDATION

A PROJECT REPORT ON

E – COMMERCE SALES ANALYSIS

Student ID – AF0410412

Student Name – NISARGA C M

INTRODUCTION

E-commerce has transformed global retail, offering convenience and accessibility. With the surge in online transactions, businesses generate vast amounts of data, which, when analysed, provide valuable insights into consumer behaviour, sales trends, and product performance. Effective data analysis helps optimize inventory, enhance marketing strategies, and improve customer engagement. This study leverages Python's powerful libraries—pandas, numpy, and matplotlib—to analyse e-commerce data, uncover trends, and support data-driven decision-making. By exploring key metrics such as revenue, customer segmentation, and product demand, this analysis aims to help businesses maximize profitability, improve operational efficiency, and deliver a personalized shopping experience for their customers.

OBJECTIVE

The primary objectives of this e-commerce analysis include:

- Understanding customer purchasing patterns and behaviours.
- Identifying top-selling products and underperforming items.
- Analysing revenue trends over time.
- Evaluating customer segmentation for targeted marketing.

- Detecting anomalies or fraud-related activities in transactions.
- Providing data-driven recommendations to improve business efficiency.

METHODOLOGY

The analysis follows a structured methodology that includes:

a. Data Collection

- The dataset used consists of historical sales transactions, including order details, customer information, and product attributes.
- The data is loaded and pre-processed using Python libraries such as **pandas** and **numpy**.

b. Data Pre-processing

- Handling missing values through imputation or removal.
- Cleaning inconsistent data formats (e.g., date formats, string normalization).
- Removing duplicate records to ensure data integrity.

c. Exploratory Data Analysis (EDA)

- Generating descriptive statistics to understand data distribution.
- Visualizing trends using **matplotlib** and **seaborn**.
- Identifying relationships between variables through correlation analysis.

d. Key Performance Metrics Calculation

- Total sales revenue, average order value, and customer lifetime value (CLV).
- Monthly and seasonal trends in sales performance.
- Customer segmentation using **RFM (Recency, Frequency, Monetary) analysis**.

e. Data Visualization and Insights

- Time-series analysis of revenue growth.
- Bar charts showcase best-selling products and customer demographics.

KEY FINDINGS

Based on the analysis, the following insights were discovered:

- **Top-Selling Products:** Certain product categories consistently outperform others, with seasonal spikes in sales.
- **Customer Segmentation:** A small percentage of loyal customers contribute to a significant portion of the revenue.
- **Sales Trends:** Revenue exhibits strong fluctuations around major sales events and holidays.
- **Underperforming Products:** Several products have low conversion rates and need strategic pricing or promotions.
- **Geographical Insights:** Certain regions show higher engagement in online shopping than others.
- **Anomalies in Transactions:** A few instances of unusually high-value orders could indicate fraudulent transactions.

SOFTWARE AND HARDWARE REQUIREMENTS

1. SOFTWARE

- **Python** (for data processing and visualization)
- **Jupyter Notebook** (for interactive analysis)
- **Libraries Used:**

  - pandas – for data manipulation
  - matplotlib and seaborn – for visualization

2. HARDWARE

- Minimum **4GB RAM**, recommended **8GB+ RAM** for smooth computation.
- Processor: **Intel i5 or equivalent**, recommended **i7+** for large datasets.
- Storage: **At least 20GB free space** for dataset storage and processing.

IMPLICATIONS

The findings of this analysis have several business implications:

- **Revenue Optimization:** By identifying high-demand products, businesses can focus on stocking and marketing them effectively.
- **Targeted Marketing:** Understanding customer behaviour allows personalized marketing strategies to enhance engagement.
- **Inventory Management:** Reducing stock of slow-moving items prevents excess inventory costs.
- **Fraud Detection:** Monitoring unusual transaction patterns can help mitigate risks associated with fraudulent activities.
- **Strategic Planning:** Insights into seasonal trends help businesses plan marketing campaigns efficiently.

RESULTS

The analysis provided actionable insights into:

- The best-performing product categories.
- Key revenue-generating customers.
- Seasonal and regional sales trends.
- Areas for potential improvement in marketing and inventory management.

CONCLUSION

This Python-based e-commerce analysis successfully identified critical patterns in customer behaviour, sales performance, and product trends. The insights derived from this study can assist businesses in making data-driven decisions to optimize revenue, improve customer satisfaction, and enhance operational efficiency. Future work can focus on advanced predictive modelling and recommendation systems to further personalize customer experience.