# Documentation for Sales and Customer Analysis Project

**Overview**

This project involved performing an exploratory data analysis (EDA) on a dataset containing sales and customer information for an e-commerce platform. The dataset consisted of 10,000 entries and included features such as order dates, product names, categories, regions, customer demographics, and revenue details. The primary goal was to uncover trends, analyze top-performing products and categories, and provide actionable recommendations to improve sales performance.

---

**Step-by-Step Process**

**Step 1: Data Loading and Libraries**

The analysis began by importing necessary Python libraries, including Pandas, Matplotlib, and Seaborn, to facilitate data manipulation and visualization. The dataset (ecommerce_professional_dataset.csv) was loaded, and an initial inspection confirmed its structure and completeness.

---

**Step 2: Data Cleaning**

The dataset was already clean, requiring no additional preprocessing. Key observations:

- **No missing values**
- **No duplicate entries**
- **No outliers detected**

---

**Step 3: Exploratory Data Analysis (EDA)**

The following analyses and insights were derived from the data:

**1. Monthly Revenue Analysis**

- Average monthly revenue was consistent from January 2023 to January 2024.
- A significant revenue drop was observed in February 2024, where sales were **50% lower than the average of other months**.

**2. Top Trending Product Analysis**

- Revenue generation across products was almost uniform with minor variations.
- The top revenue-generating products included:
    1. Stopwatch
    2. Phone
    3. Mouse
    4. Headphones
    5. Tablet
    6. Laptop
    7. Keyboard
- However, the differences in revenue between these products were marginal.

**3. Best Category with High Revenue**

- Categories showed distinct performance:
    - **Electronics**: $12,622,119 (~50% of total revenue)
    - **Accessories**: $9,447,779 (~38% of total revenue)
    - **Wearables**: $3,323,718 (~12.5% of total revenue)
- Electronics contributed the largest share of revenue, followed by Accessories and Wearables.

**4. Best Region with High Revenue**

- Revenue was evenly distributed across all four regions, indicating balanced sales performance geographically.

**5. Customer Age Distribution**

- The average customer age was found to be **40 years**.

---

**Recommendations for the Business**

Based on the analysis, the following recommendations are proposed:

1. **Investigate Revenue Drop in February 2024**
    - Examine factors contributing to the significant decline in sales for February 2024.
    - Consider promotional campaigns or targeted marketing to boost sales during slower months.
2. **Focus on Electronics and Accessories**
    - Invest in advertising and inventory for electronics, as they generate the most revenue.
    - Explore opportunities to increase the revenue margin for Accessories and Wearables.
3. **Enhance Product Differentiation**
    - Since revenue across products is almost uniform, consider strategies to differentiate top-performing products to create new opportunities for growth.
4. **Understand Customer Demographics**
    - With an average customer age of 40 years, tailor marketing campaigns to suit the preferences of this demographic.
5. **Maintain Regional Sales Consistency**
    - Since revenue is well-balanced across regions, continue with region-specific marketing efforts to sustain performance.

---

**Deliverables**

- **Exploratory Data Analysis (EDA)**: Completed in the Jupyter Notebook with detailed code, explanations, and visualizations.
- **Full Documentation**: This document will be provided as a separate PDF file.

---