# Exploratory Data Analysis (EDA) Report

## 1. Introduction

This report presents the findings from the Exploratory Data Analysis (EDA) conducted on the eCommerce Transactions dataset. The dataset includes customer, product, and transaction details. The aim of this analysis is to uncover insights into customer behavior, product performance, and sales trends.

## 2. Data Overview

### 2.1 Customers Dataset

- **Columns**: `CustomerID`, `CustomerName`, `Region`, `SignupDate`
- **Description**: Contains customer demographics.

### 2.2 Products Dataset

- **Columns**: `ProductID`, `ProductName`, `Category`, `Price`
- **Description**: Includes product information.

### 2.3 Transactions Dataset

- **Columns**: `TransactionID`, `CustomerID`, `ProductID`, `TransactionDate`, `Quantity`, `TotalValue`
- **Description**: Holds details about each transaction.

## 3. Data Cleaning

The following Python code was used for data cleaning, which included handling missing values and merging datasets.

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

# Load datasets
customers = pd.read_csv('Customers.csv')
products = pd.read_csv('Products.csv')
```

```
transactions = pd.read_csv('Transactions.csv')

# Convert dates to datetime
customers['SignupDate'] = pd.to_datetime(customers['SignupDate'])
transactions['TransactionDate'] = pd.to_datetime(transactions['TransactionDate'])

# Check for missing values
print("Missing values in Customers:\n", customers.isnull().sum())
print("Missing values in Products:\n", products.isnull().sum())
print("Missing values in Transactions:\n", transactions.isnull().sum())

# Merge datasets
merged_data = transactions.merge(customers, on='CustomerID').merge(products,
on='ProductID')
```

# 4. Key Findings

## 4.1 Total Sales by Product Category

The following bar chart presents the total sales by product category.

```
# Total sales by product category
sales_by_category =
merged_data.groupby('Category')['TotalValue'].sum().sort_values(ascending=False)
sales_by_category.plot(kind='bar', title='Total Sales by Product Category')
plt.xlabel('Product Category')
plt.ylabel('Total Sales Value')
plt.show()
```

# 5. Visualizations

Include charts and graphs such as sales trends, product performance, and customer behavior analysis. You can use Matplotlib and Seaborn for visual representation.