Exploratory Data Analysis Report

Comprehensive Data Analysis Report

Generated on: August 17, 2025 at 11:17 PM

# Table of Contents

1. Executive Summary  
2. Data Overview  
3. Data Quality Assessment  
4. Statistical Analysis  
5. Visual Analysis  
6. Key Insights  
7. Conclusions and Recommendations  
8. Appendix

# 1. Executive Summary

This report presents a comprehensive exploratory data analysis of the dataset. The analysis reveals key patterns, trends, and insights that can inform business decisions and strategic planning. The multi-agent system has collaboratively analyzed the data to provide actionable insights.

# 2. Data Overview

The dataset contains 10,000 records with 9 variables.

|  |  |  |
| --- | --- | --- |
| Variable | Data Type | Description |
| customer\_id | int64 | Unique identifier for each customer |
| age | int64 | Customer age in years |
| gender | object | Customer gender (Male/Female) |
| income | float64 | Annual income in currency units |
| purchase\_amount | float64 | Amount spent on purchase |
| product\_category | object | Category of product purchased |
| customer\_satisfaction\_score | float64 | Customer satisfaction rating (1-5 scale) |
| region | object | Geographic region of customer |
| purchase\_date | datetime64[ns] | Date of purchase |

# 3. Data Quality Assessment

Missing Values Analysis:

|  |  |  |
| --- | --- | --- |
| Variable | Missing Count | Missing Percentage |
| income | 500 | 5.00% |

# 4. Statistical Analysis

Numerical Variables Summary:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Statistic | customer\_id | age | income | purchase\_amount | customer\_satisfaction\_score |  |  |  |
| count | 10000.00 | 10000.00 | 9500.00 | 10000.00 | 10000.00 |  |  |  |
| mean | 5000.50 | 39.98 | 50675.52 | 99.83 | 3.92 |  |  |  |
| std | 2886.90 | 14.00 | 18895.66 | 99.29 | 0.87 |  |  |  |
| min | 1.00 | 18.00 | 20000.00 | 0.00 | 1.00 |  |  |  |
| 25% | 2500.75 | 29.00 | 36354.07 | 28.37 | 3.33 |  |  |  |
| 50% | 5000.50 | 39.00 | 50040.31 | 69.55 | 4.01 |  |  |  |
| 75% | 7500.25 | 50.00 | 63828.87 | 137.92 | 4.70 |  |  |  |
| max | 10000.00 | 80.00 | 139581.69 | 846.54 | 5.00 |  |  |  |

Categorical Variables Analysis:

gender:

|  |  |
| --- | --- |
| Category | Count |
| Female | 5082 |
| Male | 4918 |

product\_category:

|  |  |
| --- | --- |
| Category | Count |
| Sports | 2049 |
| Electronics | 2015 |
| Books | 2005 |
| Clothing | 1983 |
| Home | 1948 |

region:

|  |  |
| --- | --- |
| Category | Count |
| North | 2537 |
| West | 2513 |
| South | 2498 |
| East | 2452 |

# 5. Visual Analysis

The following visualizations provide insights into the data patterns and relationships.

Key Visualizations Generated:

* • Numerical Distributions
* • Gender Distribution
* • Product Category Distribution
* • Region Distribution
* • Purchase Date Distribution
* • Correlation Heatmap
* • Purchase Date Timeseries

# 6. Key Insights

* • Dataset contains 10,000 records with 9 variables
* • Missing data detected in 1 variables
* • Numerical variables analyzed: customer\_id, age, income, purchase\_amount, customer\_satisfaction\_score
* • Average customer income: $50,675.52
* • Average purchase amount: $99.83
* • Categorical variables analyzed: gender, product\_category, region
* • Gender distribution: {'Female': 5082, 'Male': 4918}
* • Most popular product category: Sports

# 7. Conclusions and Recommendations

Based on the comprehensive analysis, the following conclusions and recommendations are provided:

* • The dataset provides comprehensive information for business analysis
* • Data quality assessment completed with 500 missing values identified
* • Customer income and purchase behavior analysis reveals spending patterns
* • Customer satisfaction metrics provide insights into service quality
* • Recommendations include further segmentation analysis and predictive modeling
* • Regular data updates and quality monitoring are recommended

# 8. Appendix

Technical Details and Additional Information