Descriptive Statistics Project – Practical Report

1. Introduction

This project focuses on analyzing a household dataset containing variables such as Age, Household Income, Education Level, Family Size, and Area (Urban/Rural). The objective was to apply descriptive statistical techniques and visualization methods to understand the distribution, spread, and relationships within the data.

2. Dataset Description

- Age – Numerical (Continuous)
- Household Income – Numerical (Continuous)
- Family Size – Numerical (Discrete)
- Education Level – Categorical
- Area (Urban/Rural) – Categorical

3. Statistical Analysis Performed

The following statistical measures were calculated: • Mean, Median, and Mode for Age and Income • Range, Variance, and Standard Deviation • Interquartile Range (IQR) to measure income spread • Skewness and Kurtosis to understand distribution shape Interpretation: Income data showed moderate spread with slight positive skewness, indicating some higher-income outliers.

4. Data Visualizations

The following visualizations were created: • Histogram with Gaussian Curve – To examine income distribution. • KDE Plot – To observe density and distribution smoothness. • Boxplot – To identify spread and potential outliers. • Scatter Plot with Regression Line – To analyze relationship between Age and Income. • Boxplot comparison of Family Size across Education Levels. • Income comparison across Urban and Rural areas. These visualizations helped in understanding trends, spread, and correlations in the dataset.

5. Key Findings

- Income distribution approximately follows a normal distribution with slight skewness.
- There is a mild positive relationship between Age and Household Income.
- Higher education levels tend to show relatively higher income distribution.
- Urban households generally show slightly higher income compared to rural households.
- Family size variation differs across education categories.

6. Conclusion

This practical project successfully applied descriptive statistical methods and visualization techniques to analyze household data. The study provided clear insights into income distribution, demographic relationships, and data variability. The project demonstrates how statistical tools can be used to transform raw data into meaningful information for decision-making and analysis.