📊 Exploratory Data Analysis (EDA) Report – Iris Dataset

This document compares the original Mini Guide steps with the actual code implementation for the Iris dataset, and highlights extra enhancements for deeper insights.

## 📋 Comparison Table

|  |  |  |  |
| --- | --- | --- | --- |
| **Step** | **Mini Guide Description** | **Implemented Code** | **Extra Feature (If Any)** |
| 1 | Generate summary statistics (mean, median, std, etc.) | Yes | Included .info(), .isnull(), .sum(), and shape |
| 2 | Create histograms and boxplots for numeric features | Yes | Used plt.subplot layout for clean display |
| 3 | Use pairplot/correlation matrix for feature relationships | Yes | Added heatmap with correlation values |
| 4 | Identify patterns, trends, or anomalies in the data | Yes | Included violin plots, count plots, and 3D scatter |
| 5 | Make basic feature-level inferences from visuals | Yes | Multiple print() inferences embedded after EDA |
| 6 | — | Yes | Used plotly for interactive 2D/3D scatter visualizations |
| 7 | — | Yes | Checked dataset balance (countplot per species) |
| 8 | — | Yes | Correlation between petal features analyzed numerically |

## Summary:

The implemented EDA not only followed all the suggested steps from the mini guide but also introduced meaningful enhancements like interactive Plotly graphs, class balance checking, violin plot analysis, and deep inferences. These additions make the dataset thoroughly explored and ready for modeling and feature selection.