Visual Insights Report – Iris Dataset (EDA)

1. Distribution of Numerical Features: Histograms show that petal length and petal width have clear separation among species, whereas sepal features overlap significantly. This indicates petal features are more informative.

2. Outlier Analysis: Mild outliers are observed mainly in sepal width. No extreme anomalies were found that could severely affect model performance.

3. Correlation Analysis: Strong positive correlation exists between petal length and petal width, indicating multicollinearity. Sepal features show weaker correlations.

4. Feature Importance: Petal length and petal width are the most influential features for predicting Iris species.

5. Species Separation: Iris-setosa is clearly separable from other species using petal features, making classification easier.

6. Data Quality: Dataset is clean with no missing values, making it suitable for machine learning models without heavy preprocessing.

Summary: The EDA confirms that petal measurements dominate predictive power, and visualization effectively highlights data patterns, correlations, and feature behavior.