

1. Data Structure

- The dataset consists of 150 rows and 5 columns.
- Features: Sepal Length, Sepal Width, Petal Length, Petal Width.
- Target: Species (Setosa, Versicolor, Virginica).
- **Data Quality:** No missing values were found.

2. Visual Insights & Observations

- **Distribution:**
 - Sepal Length and Sepal Width follow a roughly Normal (Gaussian) distribution.
 - Petal Length and Petal Width show a bimodal distribution (two peaks), suggesting clear separation between different species.
- **Class Balance:**
 - The dataset is perfectly balanced with 50 samples for each of the three species.
- **Outliers:**

- Box plots reveal a few outliers in Sepal Width for the "Virginica" species, but they are not extreme.
- **Correlation:**
- **High Positive Correlation:** Petal Length and Petal Width are highly correlated (0.96). If Petal Length increases, Width also increases.
- **Negative Correlation:** Sepal Width has a slight negative correlation with Petal Length.

3. Conclusion The dataset is clean and balanced. The "Petal" features contain the most information for distinguishing between species, making them the most important features for any future Machine Learning model.