

Week 3

Hosain Masba Tamim

Objectives & Outcomes

- Planned:

- Work with and visualize datasets using Python libraries.

- Achieved:

- Basic knowledge about Numpy, Pandas, Matplotlib and Seaborn
- Data pre-processing using Numpy, Pandas
- Data visualization using Matplotlib and Seaborn

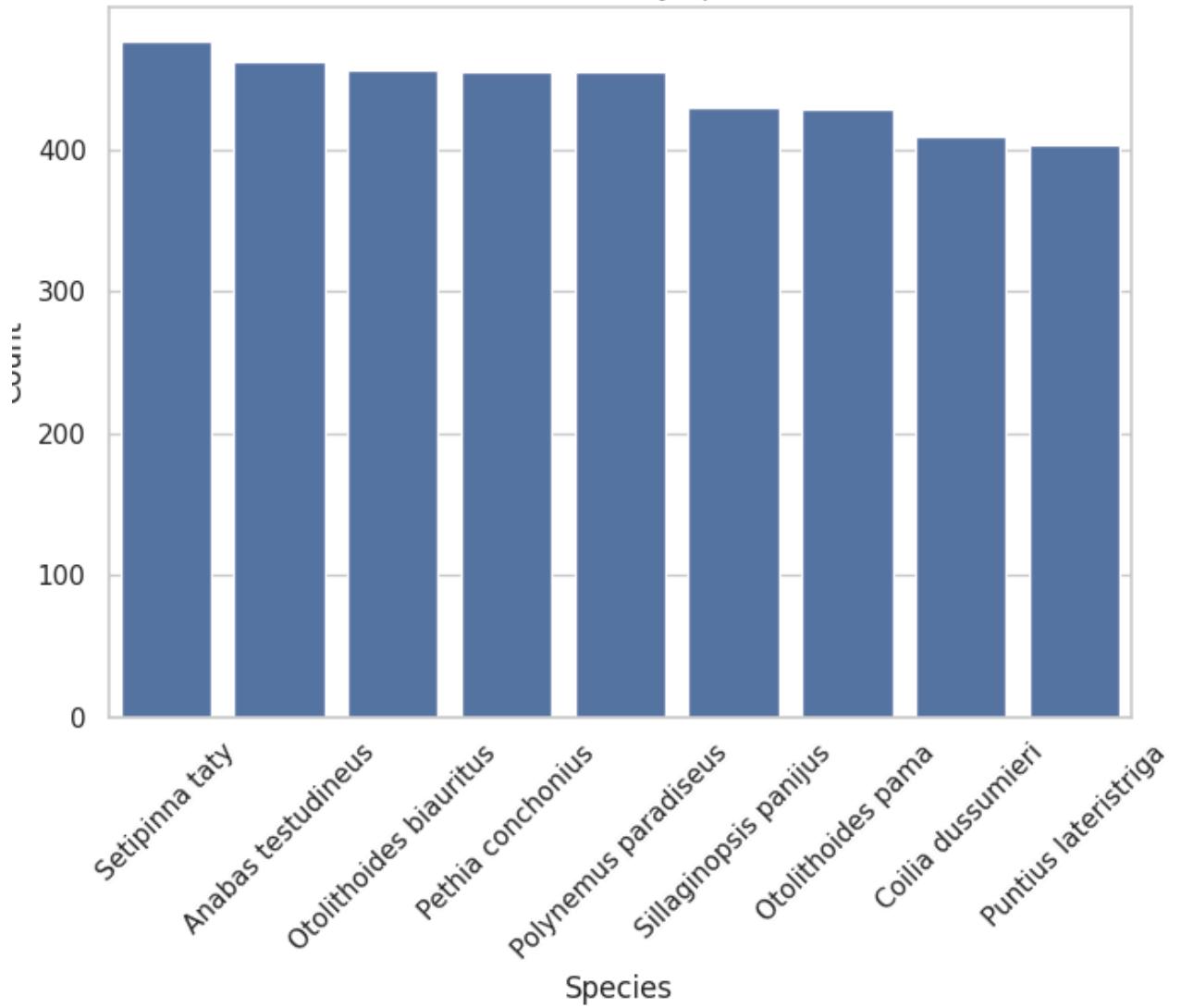
Script: [*fish.py*](#)

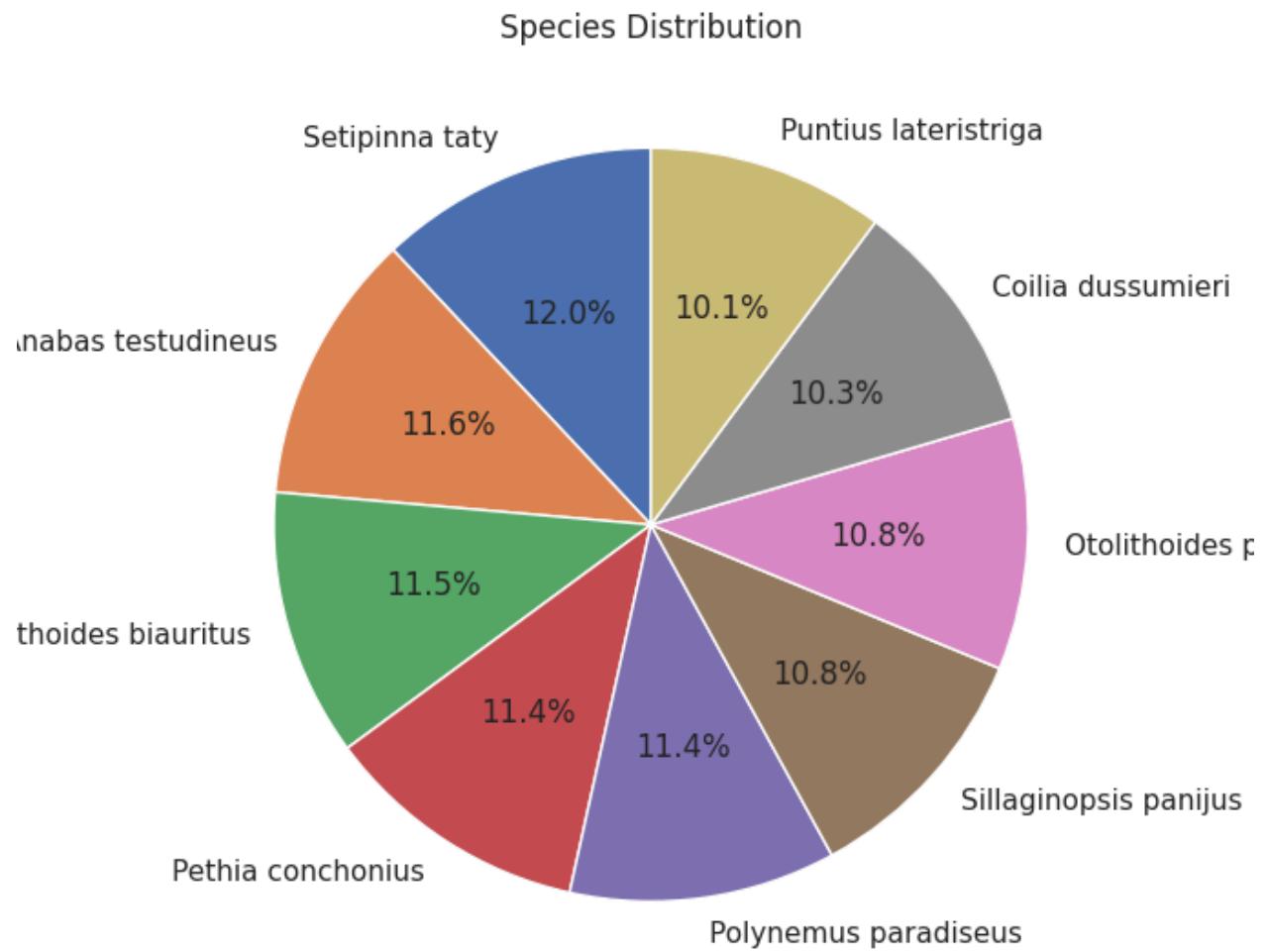
Code & Explanations

- **What it does :** This Python script demonstrates data preprocessing, and visualization for a fish dataset using Pandas, NumPy, Matplotlib, and Seaborn.
- **How it works :** The fish.py script loads a fish dataset (fish_data.csv) . It cleans it by removing duplicate data and invalid values and converts columns to numeric types. It then performs basic exploratory analysis, including printing the first rows, dataset info, descriptive statistics, and species counts. Lastly, it creates several visualizations bar and pie charts for species distribution, scatter plots for length vs weight and w/l ratio vs weight, a correlation heatmap, and a boxplot of weight by species using Seaborn and Matplotlib to help understand patterns and relationships in the data.

- **Screenshots :**

Count of Fish by Species





Length vs Weight by Species

