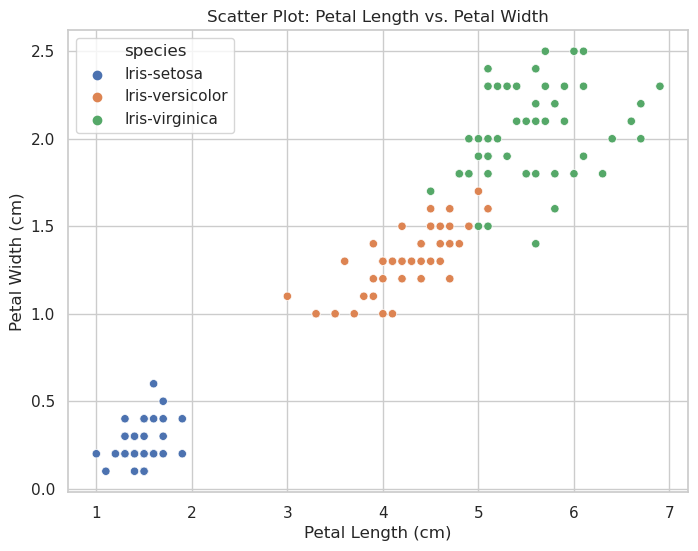
**Question 1:**

I utilized the Iris dataset in this question, which was first introduced by Ronald A. Fisher in 1936 ("Iris", 2023). It consists of four separate measurements (length, width of sepals, and petal) for three iris flower species (Iris setosa, Iris versicolor, and Iris virginica). This dataset, which contains 50 samples from each of the three species, is an excellent option for showcasing the strengths and limitations of various visualization methods.

Scatter Plots:

The scatter plot represents the values of two numerical variables using dots. Each dot on the horizontal and vertical axes represents a value for a single data point. Scatter plots are used to investigate the relationships between variables.

Example Scatter Plot (Petal Length vs. Petal Width):

The scatter plot shown above depicts the petal length and width of the species in the iris dataset. Each dot represents individual iris flower from the dataset. The horizontal horizontal location of each point denotes the length (in centimeters) and the vertical position indicates the width (in meters) of that flower petal. The graphic shows a generally strong positive link between a flower petal’s length and width.

**Effectiveness:**

Scatter plots are effective when observing and demonstrating correlations between two numerical variables. In this, a scatter plot's dots not only represent the values of individual data points but reveal patterns when the data is viewed as a whole.

Scatter plots are also commonly used to identify correlational relationships. In these circumstances, we want to know what a good forecast for the vertical value would be if we were given a specific horizontal value. The variable on the horizontal axis is frequently referred to as an independent variable, while the variable on the vertical axis is called the dependent variable. Variable relationships can be described as: positive or negative, strong or weak, linear or nonlinear.

Advantages:

* Allows for the identification of trends, clusters, or outliers in the data.
* Can be used for visualizing pairwise relationships between two continuous variables.

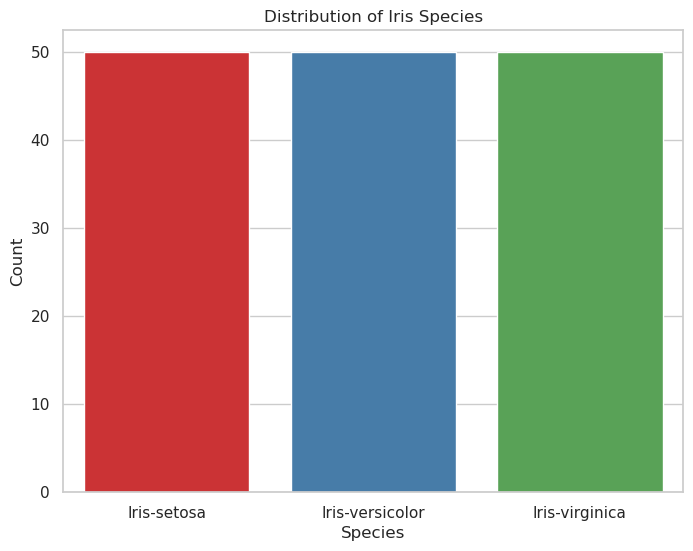
Limitations:

* When we have a large number of data points to plot, scatter plot may suffer the problem of over-plotting.
* Not easy to interpret scatter plots

Bar Chart:

A bar chart displays numerical values for categorical feature levels as bars. Levels are represented on one axis of the chart, and values are plotted on the other. Each categorical value is assigned one bar, and the length of each bar corresponds to the value of the bar. For easier value comparison, bars are plotted on the same baseline.

Example Bar Chart:



The bar chart shown above depicts distribution of Iris species. The horizontal axis is plotted with the Iris species, and the height of each bar correlates to the number total number of each flower. This chart shows that the dataset has a total of 50 flowers for each specie.

**Effectiveness:**

Bar charts are effective in comparing discrete data across different categories. They are also ideal for visualizing summary statistics or distributions.

Advantages:

* Bar charts are useful for comparing discrete data across categories.
* Bar charts are simple to understand and well-known.

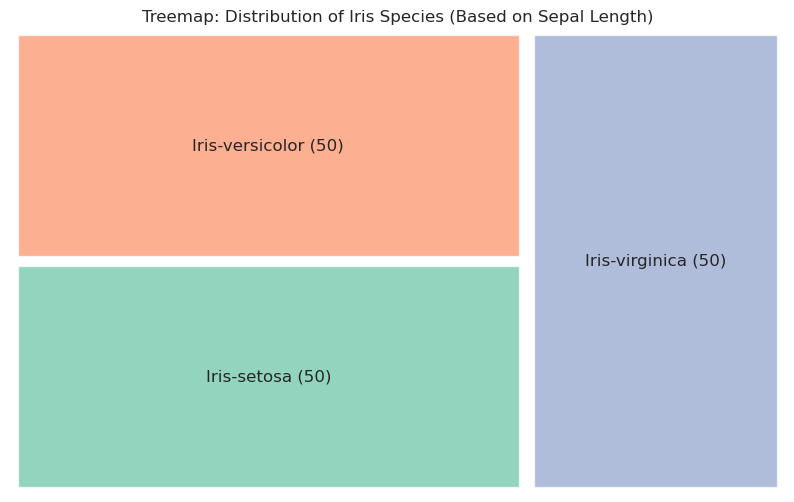
Limitations:

* Bar charts are less useful for displaying correlations between many continuous variables.
* Bar charts are not suited for representing hierarchical data.

**Tree-map:**

Treemaps are hierarchical data representations. They are composed of a sequence of nested rectangles with sizes proportionate to the associated data value. A huge rectangle symbolizes a branch of a data tree, and it is broken into smaller rectangles that reflect the size of each node within that branch.

Example Treemap:

This tree-map shows the distribution of Iris species based on the sepal length. The red box shows the number of the Iris-versicolor, the green box shows the setosa and the blue box shows the virginica specie.

**Effectiveness:**

Treemaps are useful for representing several variables inside a hierarchical structure and visualizing hierarchical data. They also offer a concise, hierarchical picture of data.

Advantages:

* The ability to recognize patterns and establish relationships between two categories or elements in a hierarchical data structure.
* Space utilization while showing tens of thousands of data points, with drill-down capability as needed.
* Displaying several items accurately at the same time, including "part to whole" ratios. This facilitates data visualization.
* Size and color keys are used to visualize various properties. Color-coding categories and subcategories to match the parent categories is possible.

Limitations:

* Tree-maps are not suitable for activities requiring exact comparisons.
* Tree-maps are not for a balanced tree. The fundamental objective of a treemap becomes extremely difficult in these instances.
* Cannot be used if the data is not hierarchical.
* Tree-maps can be visually overwhelm to users because they are frequently used to show extremely massive data sets.

**References**

Iris. (2023, July 3). In Wikipedia. https://en.wikipedia.org/wiki/Iris\_flower\_data\_set