

Intent of the application

The purpose of this program is to explore the Iris dataset as the first programming assignment for the Data Mining course, perform exploratory operations, display features of the dataset, and create plots for the dataset.

The plots will be of the sepal length, sepal width, and petal length, which are characteristics in the observations of the dataset.

Dataset to be used

The dataset to be used is the Iris dataset, for which more information can be found here: <https://stat.ethz.ch/R-manual/R-devel/library/datasets/html/iris.html>

Use case

The use case of this application is to demonstrate basic operations with datasets, as the first steps to take when exploring a dataset for the first time and show the contents of the iris dataset.

It will be used as an example on how to explore a dataset, create basic plots, and how to interpret these results.

There are no primary actors in this case, as the purpose of this application is only to be used as an example.

Variables

Sepal.Length: The length of the sepal, which is the outer part of the flower that encloses a developing bud.

Sepal.Width: The width of the sepal.

Petal.Length: The length of the petal of the flower, which are leaves that surround the reproductive parts of a flower.

Petal.Width: The width of the petal of the flower.

Labels

Species: Describes the species of the flower associated to the previous measurements.

Data import

In this application, there is no input needed from the user.

Proposed Libraries

datasets: Used to import the Iris dataset

ggplot2: Used to create plots from the dataset

Library source

datasets: <https://cran.r-project.org/package=dataset>

ggplot2: <https://github.com/tidyverse/ggplot2>

Analysis of results

We can see a relationship, in which most of the points on the bottom-right part of the plot tend to be darker, which with the color gradient code, we know it means that most of these points have a petal length within 1 and 3, while most of the points on the upper-left part of the plot tend to be lighter, which means they are between 4 and 6.

In other words, the higher the sepal width, the lower the sepal length, the petal length would decrease.

The lower the sepal width, the higher the sepal length, the petal length would increase.