

Lab 1: Introduction to Orange

The first step is to get and install the software on your machine. Alternatively, you can log in to TU Dublin Lab and click the orange icon.

[Orange Data Mining - Download](https://orangedatamining.com/download/#macos)

The first thing you need to do is to go to the first video of the online tutorials. If you do not have headphones, do not worry; use the subtitles.

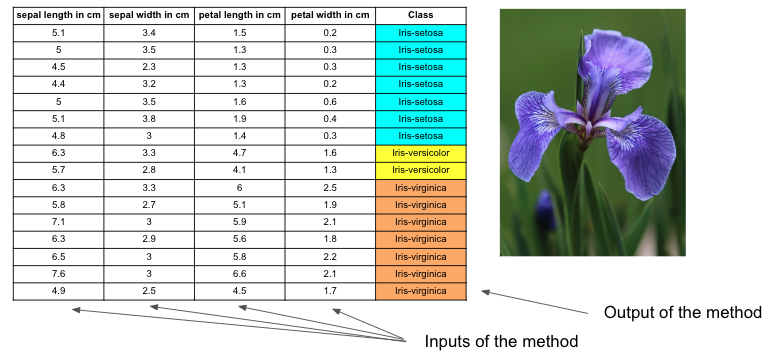
[Getting Started with Orange 01: Welcome to Orange](https://youtu.be/HXjnDIgGDuI)

**Part I:** [Getting Started with Orange 01: Welcome to Orange](https://youtu.be/HXjnDIgGDuI)

The first video is just a quick overview of how Orange works. In this video, you would do the following things:

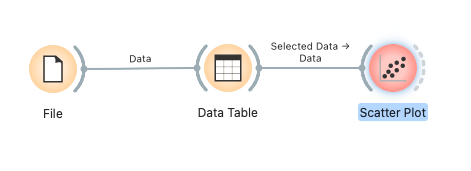
* General overview of how Orange works.
* Use the File widget to load the Iris dataset. The Iris dataset is one of the classic datasets used in ML. <https://www.kaggle.com/datasets/uciml/iris>
* Use the Scatter plot to visualise the table.
* Save your file and load it again.
* Explore the different options and try to understand the logic behind them.
* Now, try to use the “Data Table” widget to select instances of only two classes.

**Part II:** [Getting Started with Orange 02: Data Workflows](https://youtu.be/lb-x36xqJ-E)

This second video is very similar to the first one, but it is best if you open a new file. This way, you can keep your exercise tidy in different files. As in the first exercise, we are going to use a file and a data table. The exercise will allow you to inspect the different attributes of the dataset.

**Part III:** [Getting Started with Orange 03: Widgets and Channels](https://youtu.be/2xS6QjnG714)

This exercise is about visualising subsegments of the



**Part IV:** [Getting Started with Orange 04: Loading Your Data](https://youtu.be/MHcGdQeYCMg)

This is about the dataset format that can be used to integrate databases with Orange.

**Part V:** [Getting Started with Orange 05: Hierarchical Clustering](https://youtu.be/dJ5z2SRwzgs)

In this exercise, we will examine hierarchical clustering, which groups objects according to similarity. First, we will use the Euclidian distance to determine how close one object is to the other. Then, we will plot a dendrogram and see if the elements of a similar class are in the same group.