



**Samen aan de Slag 2025**

21 January 2025, Utrecht

# From Data to Insight: Data Analysis and Visualization with Orange

**Dr. Serkan Girgin**

[s.girgin@utwente.nl](mailto:s.girgin@utwente.nl)

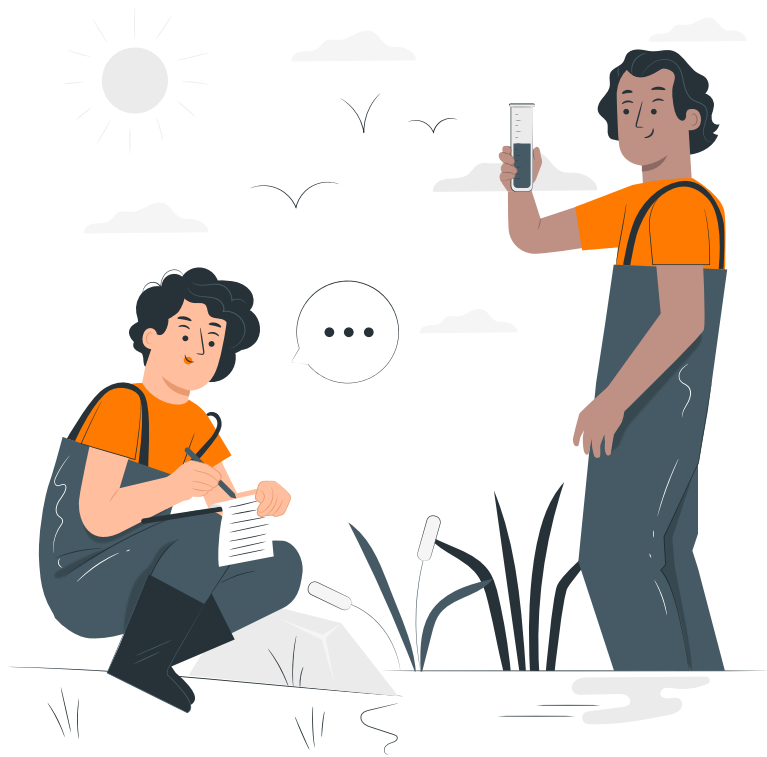
<https://linkedin.com/in/serkan-girgin/>

Head, Center of Expertise in Big Geodata Science (CRIB)  
Associate Professor, Department of Geo-information Processing

**Faculty of Geo-information Science and Earth Observation (ITC)**

**UNIVERSITY  
OF TWENTE.**

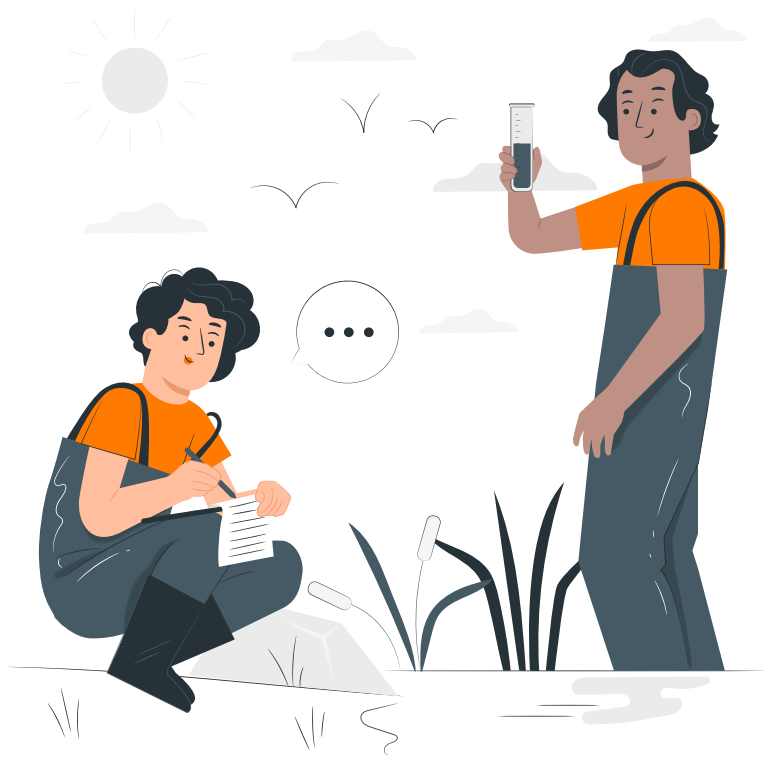




Citizen Science



Data Science



Citizen Science

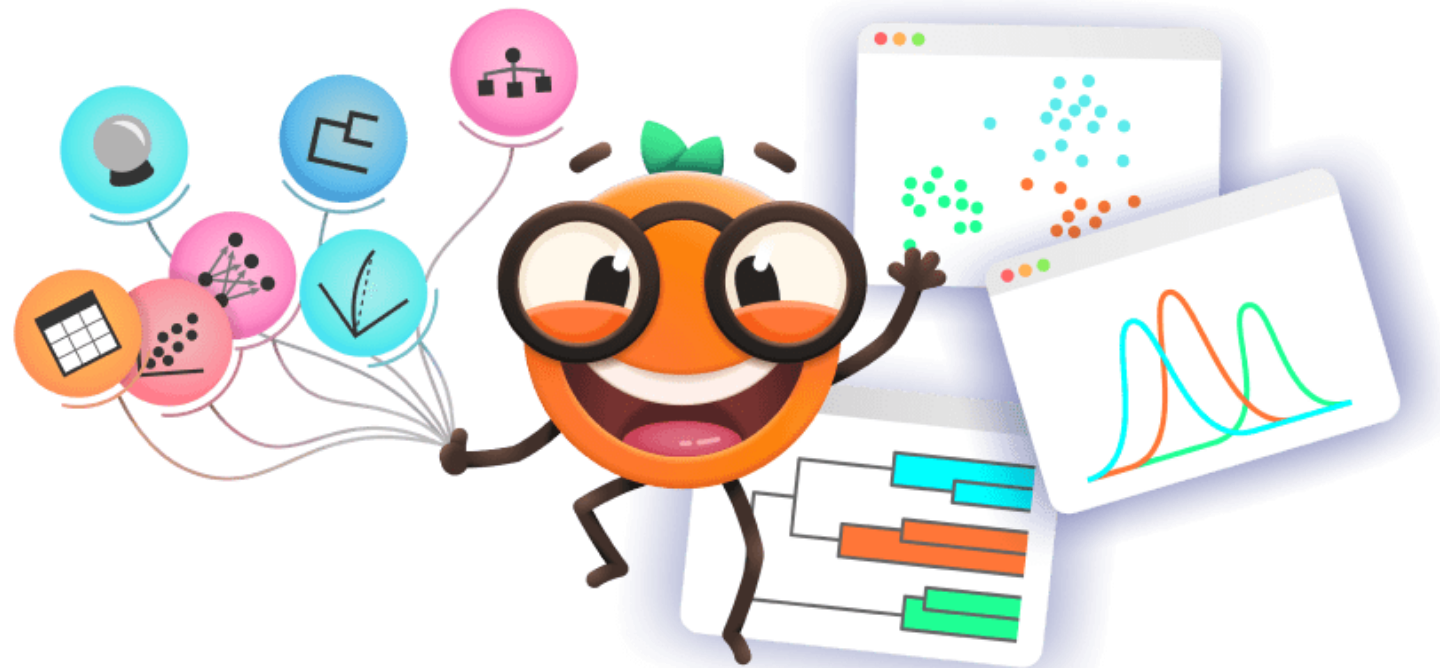


orange  
DATA MINING

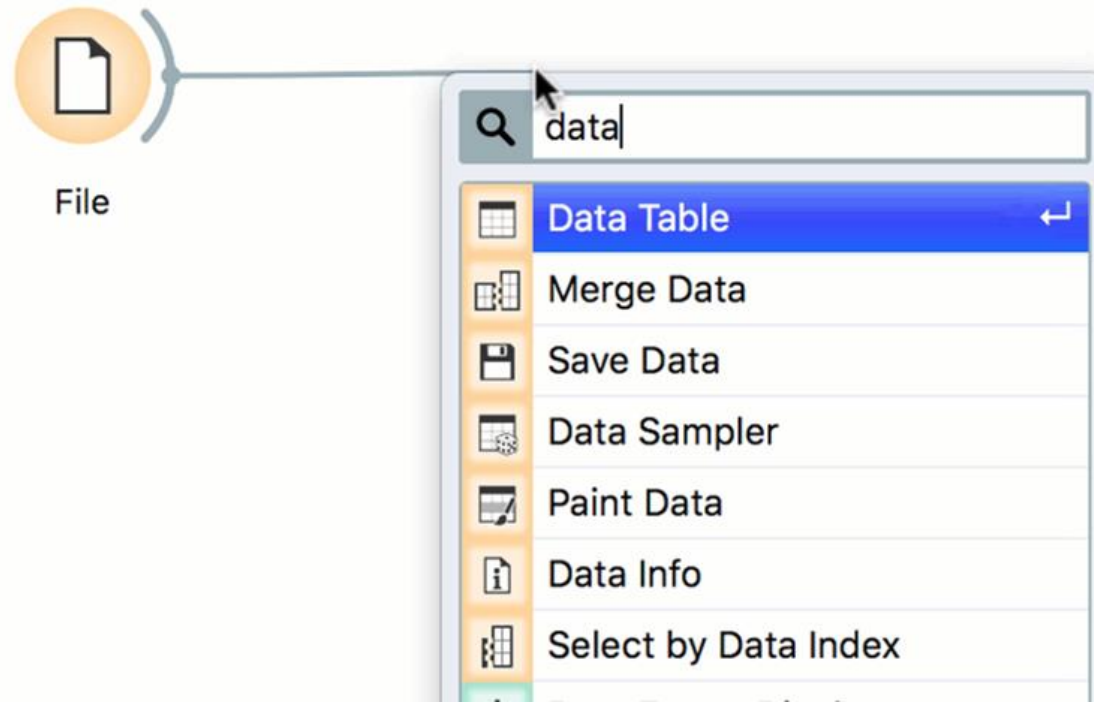


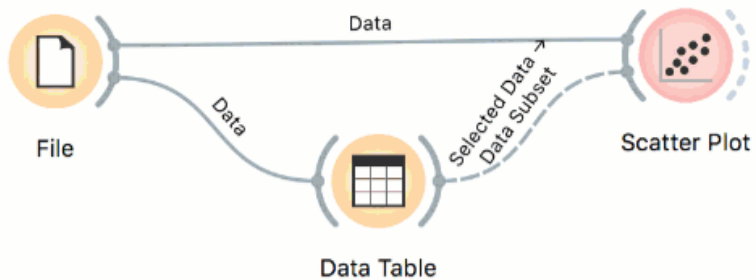
Data Science

Orange Data Mining is an open-source software that offers an interactive, workflow-based interface allowing users to build data analysis processes by combining widgets

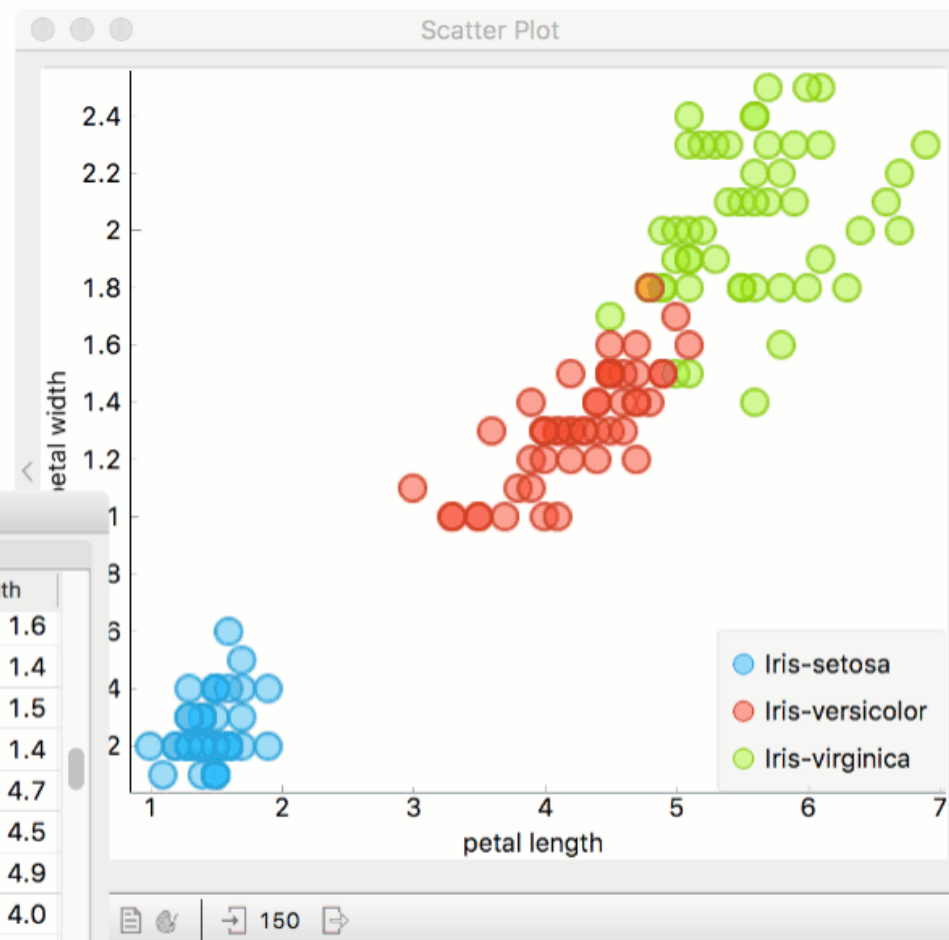


Workflows provide a **simple** approach to data analysis and visualization, **eliminating the need for programming** experience

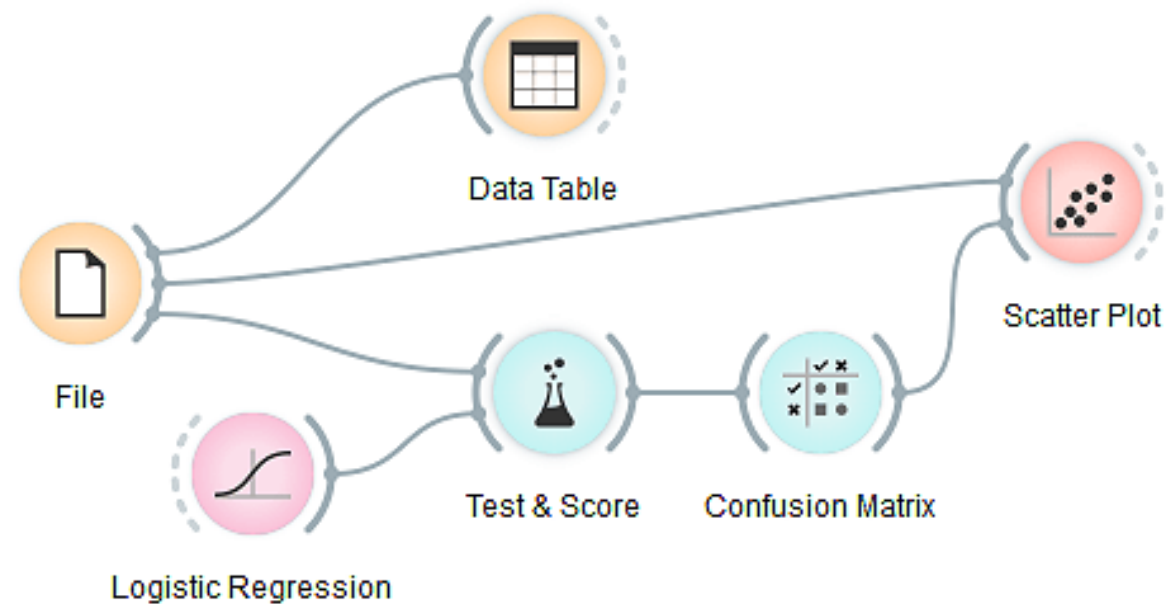


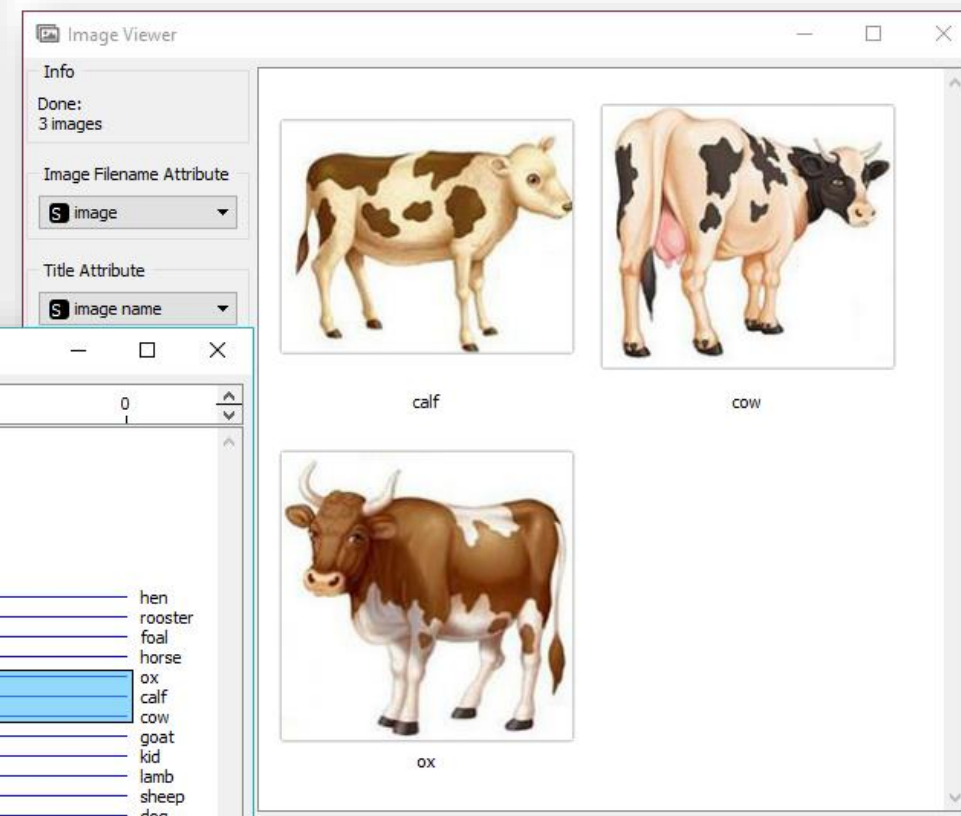
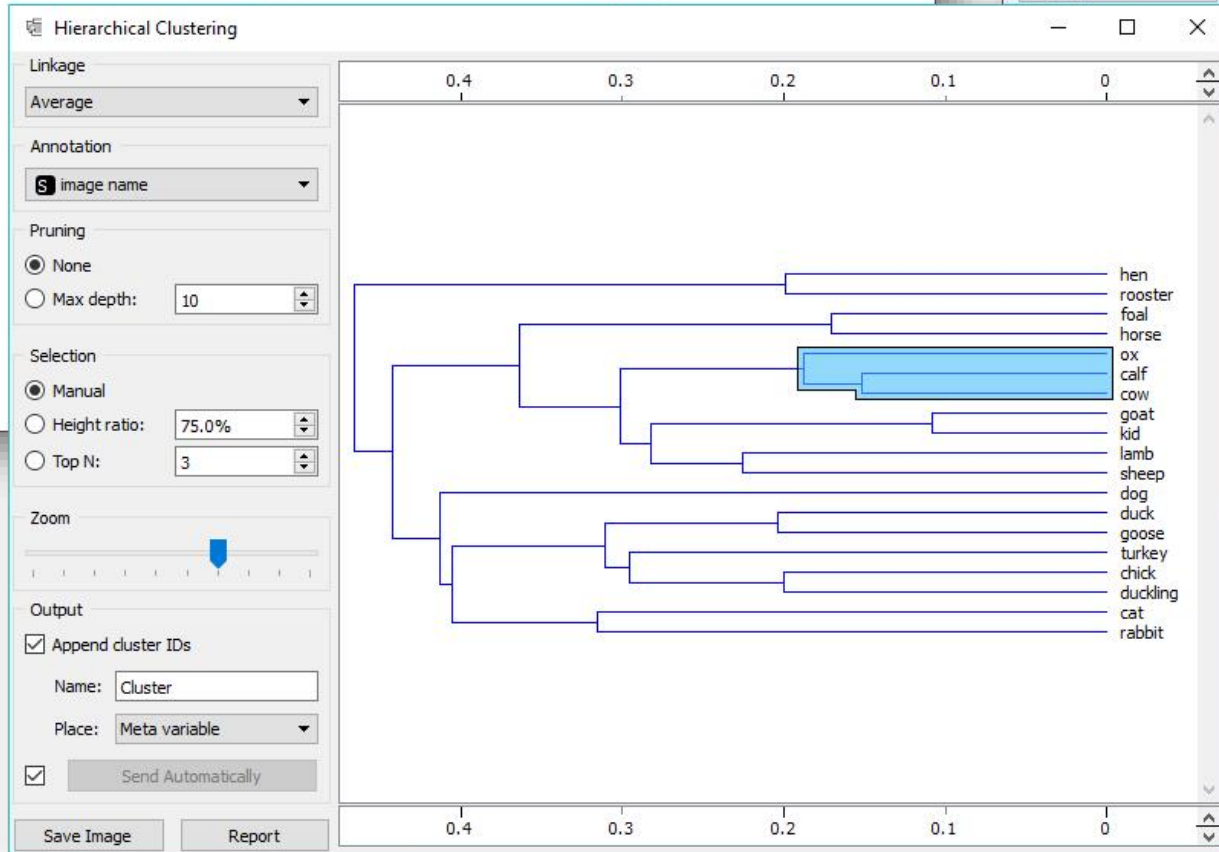


Data Table				
	iris	sepal length	sepal width	petal length
47	Iris-setosa	5.1	3.8	1.6
48	Iris-setosa	4.6	3.2	1.4
49	Iris-setosa	5.3	3.7	1.5
50	Iris-setosa	5.0	3.3	1.4
51	Iris-versicolor	7.0	3.2	4.7
52	Iris-versicolor	6.4	3.2	4.5
53	Iris-versicolor	6.9	3.1	4.9
54	Iris-versicolor	5.5	2.3	4.0
55	Iris-versicolor	6.5	2.8	4.6
56	Iris-versicolor	5.7	2.8	4.5
57	Iris-versicolor	6.3	3.3	4.7
58	Iris-versicolor	4.9	2.4	3.3
59	Iris-versicolor	6.6	2.9	4.6



These widgets can perform tasks such as data **preprocessing**,  
**visualization**, **statistical analysis**, and **machine learning**



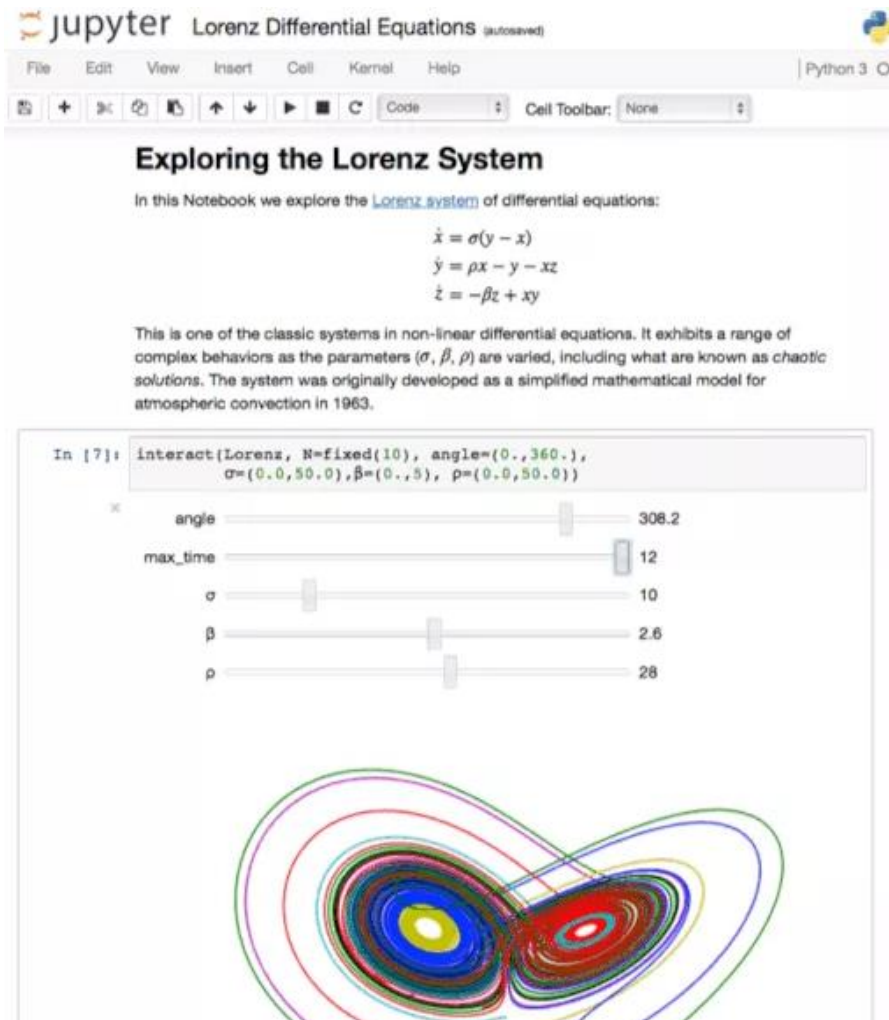














These features make Orange Data Mining a suitable tool for **citizen science**, as it makes data analysis accessible and engaging for **the public\***




Orange workflows also offer deeper **insights into the data analysis process** compared to other methods, such as interactive notebooks



FileEditViewInsertCellKernelWidgetsHelp



Code

 Voila

```
In [13]: # importing libraries

from __future__ import print_function
from ipywidgets import interact, interactive, fixed, interact_manual
from IPython.core.display import display, HTML

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import plotly.express as px
import folium
import plotly.graph_objects as go
import seaborn as sns
import ipywidgets as widgets

In [14]: # loading data right from the source:
death_df = pd.read_csv('https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_githubusercontent/csse_gh_data/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_deaths_global.csv')
confirmed_df = pd.read_csv('https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_githubusercontent/csse_gh_data/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_confirmed_global.csv')
recovered_df = pd.read_csv('https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_githubusercontent/csse_gh_data/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_recovered_global.csv')
country_df = pd.read_csv('https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_githubusercontent/csse_gh_data/csse_covid_19_data/csse_covid_19_historical/csse_covid_19_historical.csv')

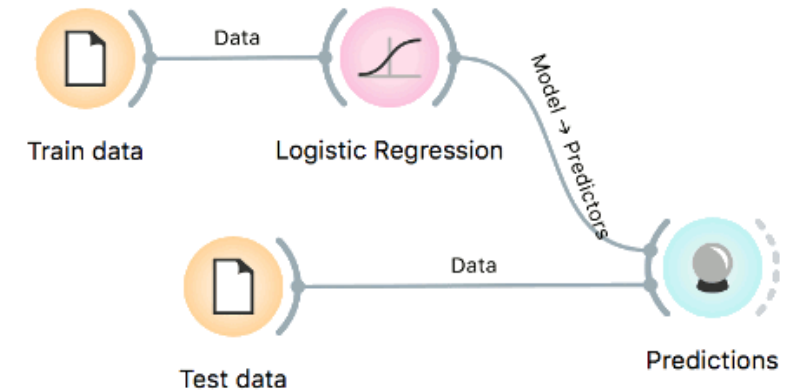
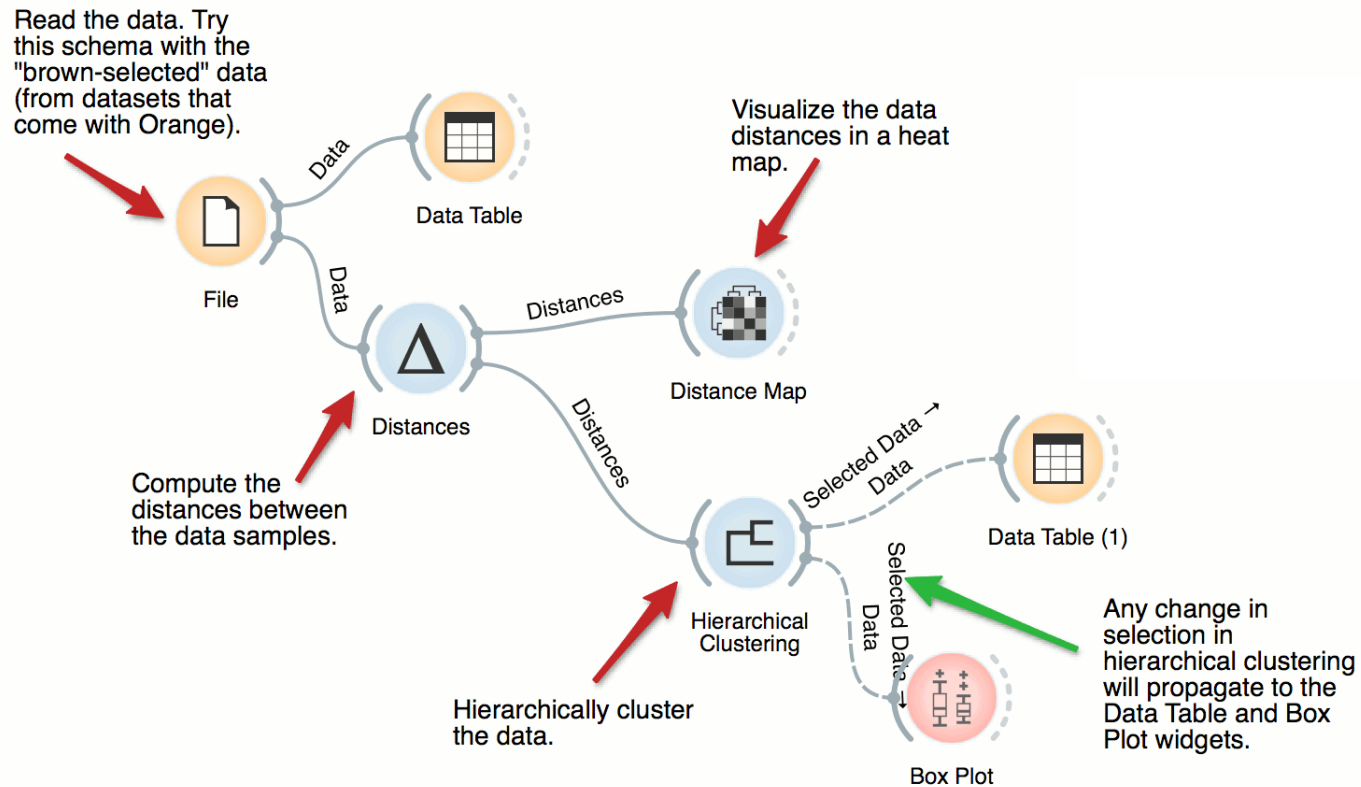
In [15]: confirmed_df.head()

In [16]: recovered_df.head()

In [17]: death_df.head()

In [18]: country_df.head()
```

Orange workflows also offer deeper insights into the data analysis process compared to other methods, such as interactive notebooks



Detailed documentation is available, including details about the widgets

## Orange Visual Programming

### Navigation

[Loading your Data](#)  
[Building Workflows](#)  
[Exporting Models](#)  
[Exporting Visualizations](#)  
[Learners as Scorers](#)  
[Report](#)

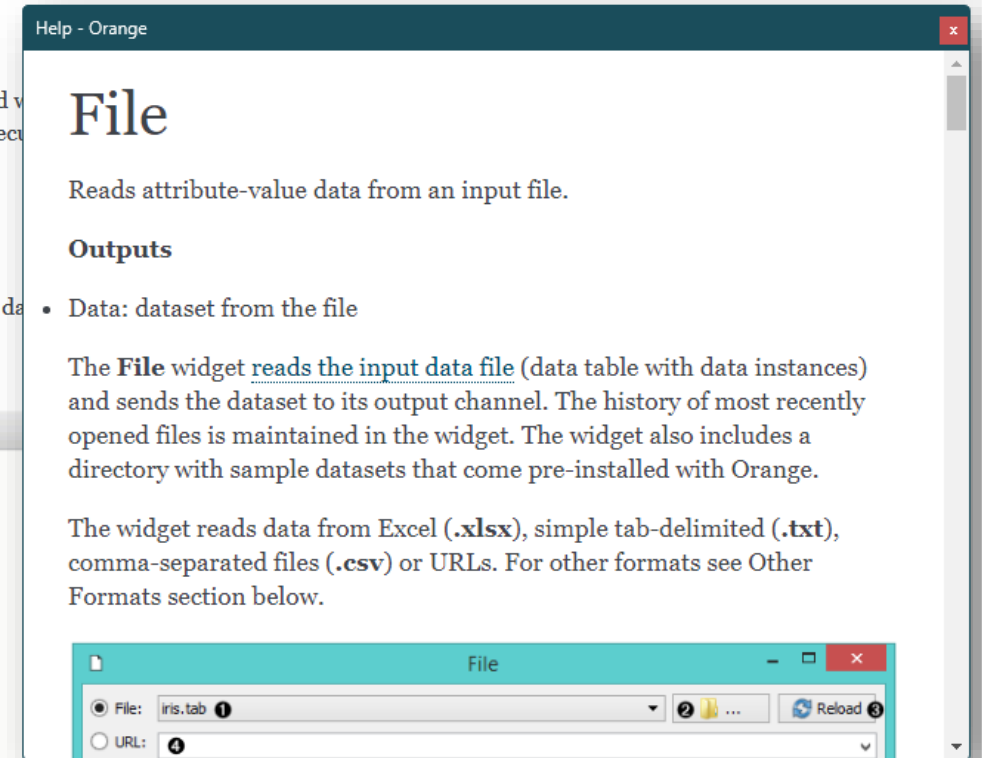
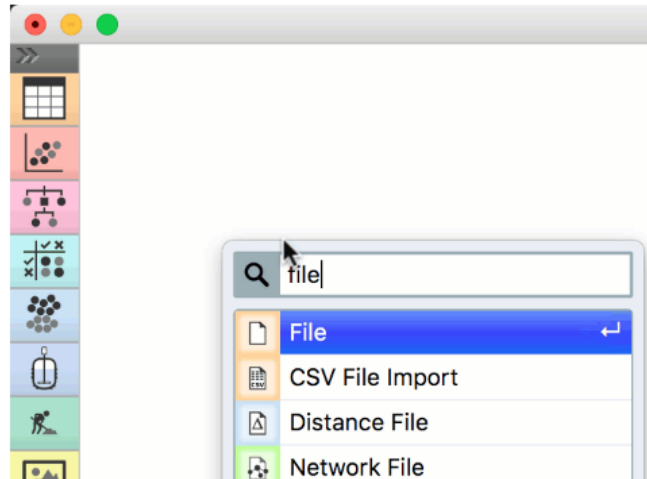
[File](#)  
[CSV File Import](#)  
[Datasets](#)  
[SQL Table](#)  
[Save Data](#)  
[Data Info](#)  
[Aggregate Columns](#)  
[Data Table](#)  
[Select Columns](#)  
[Select Rows](#)  
[Data Sampler](#)  
[Transpose](#)  
[Discretize](#)  
[Continuize](#)

## Building Workflows

The core principle of Orange is visual programming, which means each analytical step is contained within a widget. Widgets are placed on the canvas and connected into an analytical workflow, which is executed from left to right. Orange never passes data backwards.

### Simple workflow

Let us start with a simple workflow. We will load the data with the File widget, say the famous *Iris* dataset. Right-click on the canvas. A menu will appear. Start typing "File", then press Enter to confirm the selection. [File](#) widget will be placed on the canvas.



<https://orange3.readthedocs.io/projects/orange-visual-programming/>

# Example workflows are provided to facilitate rapid adoption

Text Mining

Bioinformatics

Fairness

Survival Analysis

Classification

Clustering

Hierarchical Clustering

Cox Regression

Scatter Plot

Visualization

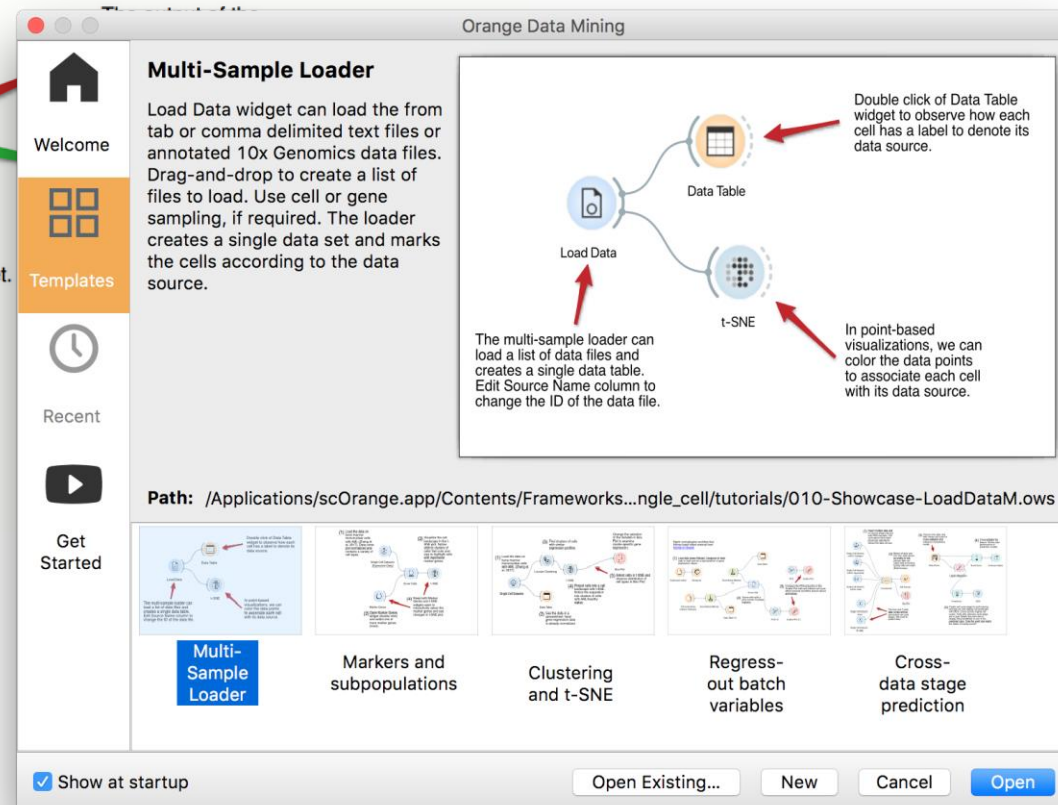
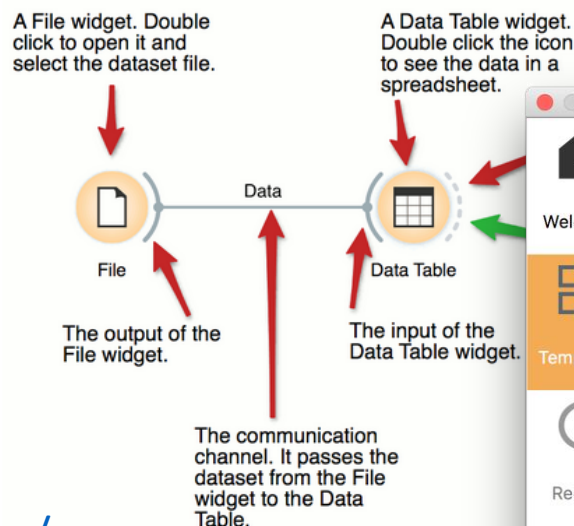
Data Table, Data Loading

## File and Data Table

The basic data mining units in Orange are called widgets. In this workflow, the File widget reads the data. File widget communicates this data to Data Table widget that shows the data in a spreadsheet. The output of File is connected to the input of Data Table.

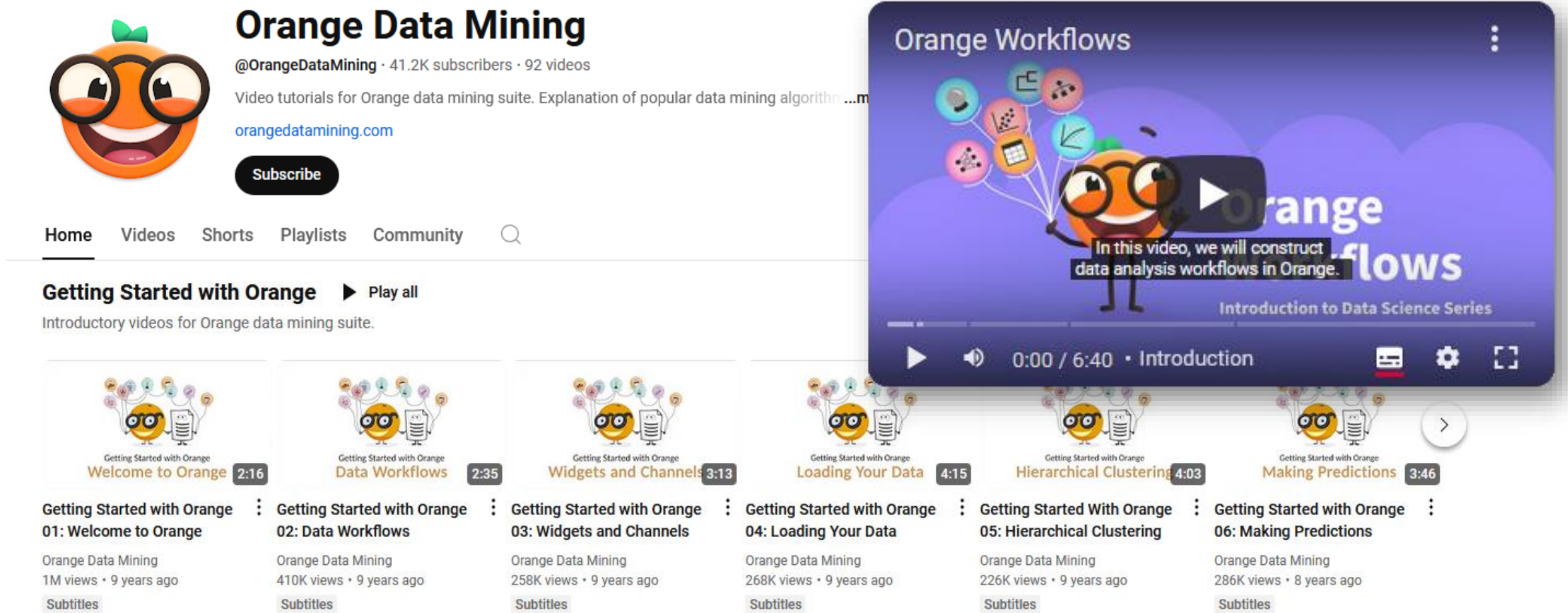
Download

<https://orangedatamining.com/examples/>





Video tutorials are provided to help users learn basic and advanced use cases



**Orange Data Mining**  
@OrangeDataMining · 41.2K subscribers · 92 videos  
Video tutorials for Orange data mining suite. Explanation of popular data mining algorithms ...  
[orangedatamining.com](http://orangedatamining.com)  
Subscribe

Home Videos Shorts Playlists Community

**Getting Started with Orange** ▶ Play all  
Introductory videos for Orange data mining suite.

**Orange Workflows**  
In this video, we will construct data analysis workflows in Orange.  
Introduction to Data Science Series  
0:00 / 6:40 · Introduction

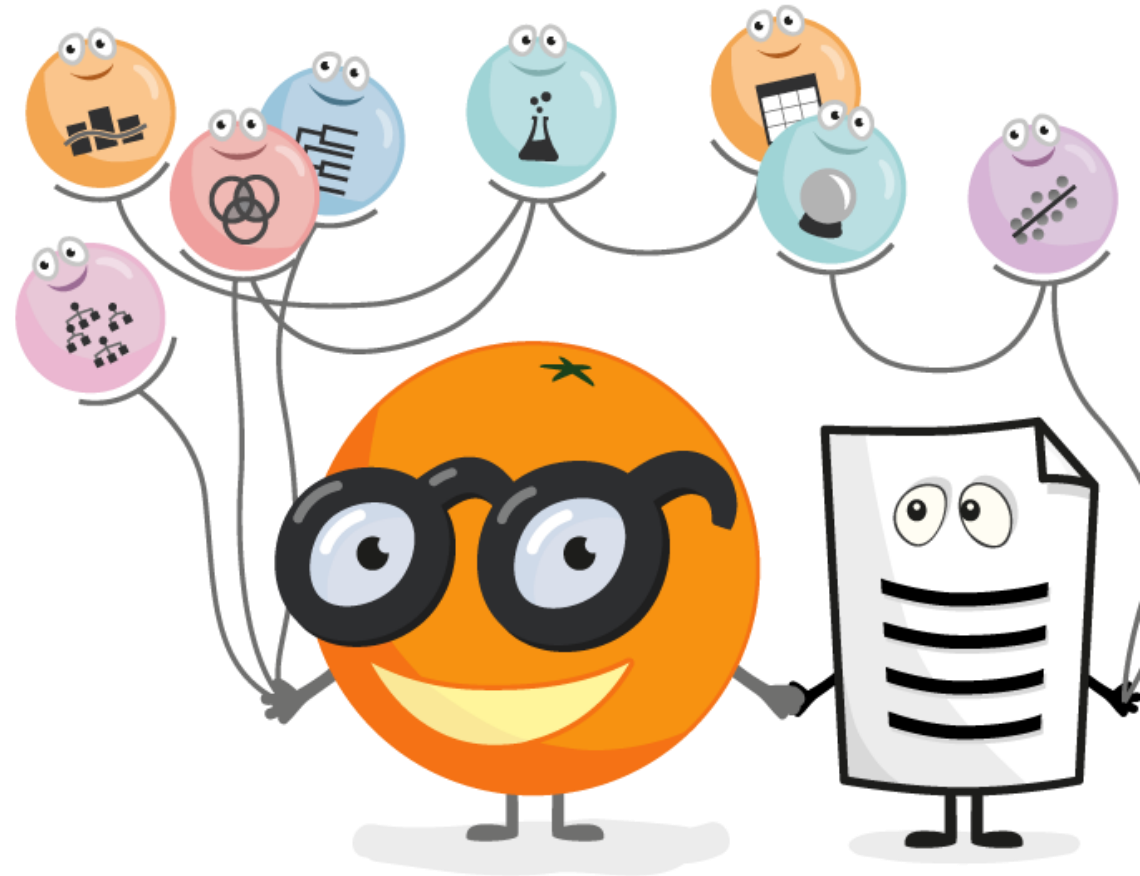
**Getting Started with Orange**

- 01: Welcome to Orange** 2:16  
Orange Data Mining  
1M views · 9 years ago  
Subtitles
- 02: Data Workflows** 2:35  
Orange Data Mining  
410K views · 9 years ago  
Subtitles
- 03: Widgets and Channels** 3:13  
Orange Data Mining  
258K views · 9 years ago  
Subtitles
- 04: Loading Your Data** 4:15  
Orange Data Mining  
268K views · 9 years ago  
Subtitles
- 05: Hierarchical Clustering** 4:03  
Orange Data Mining  
226K views · 9 years ago  
Subtitles
- 06: Making Predictions** 3:46  
Orange Data Mining  
286K views · 8 years ago  
Subtitles

<https://orangedatamining.com/getting-started/>

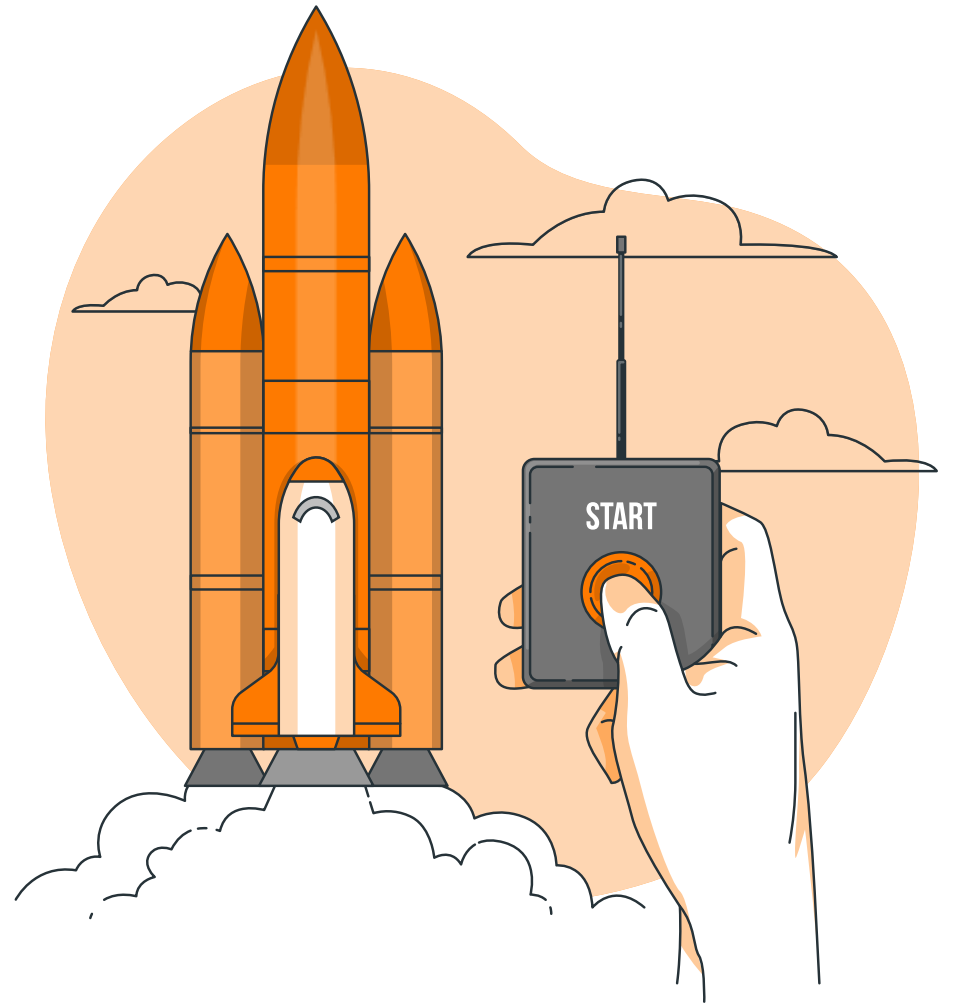
This workshop aims to **introduce** Orange and **explore its capabilities**

- Installation and getting started with Orange
- Reading and displaying data
- Creating data
- Clustering
- Making predictions
- Model evaluation and scoring (self study)
- Image analytics
- Text analytics
- Mapping



# Let's start!

<https://orangedatamining.com/>



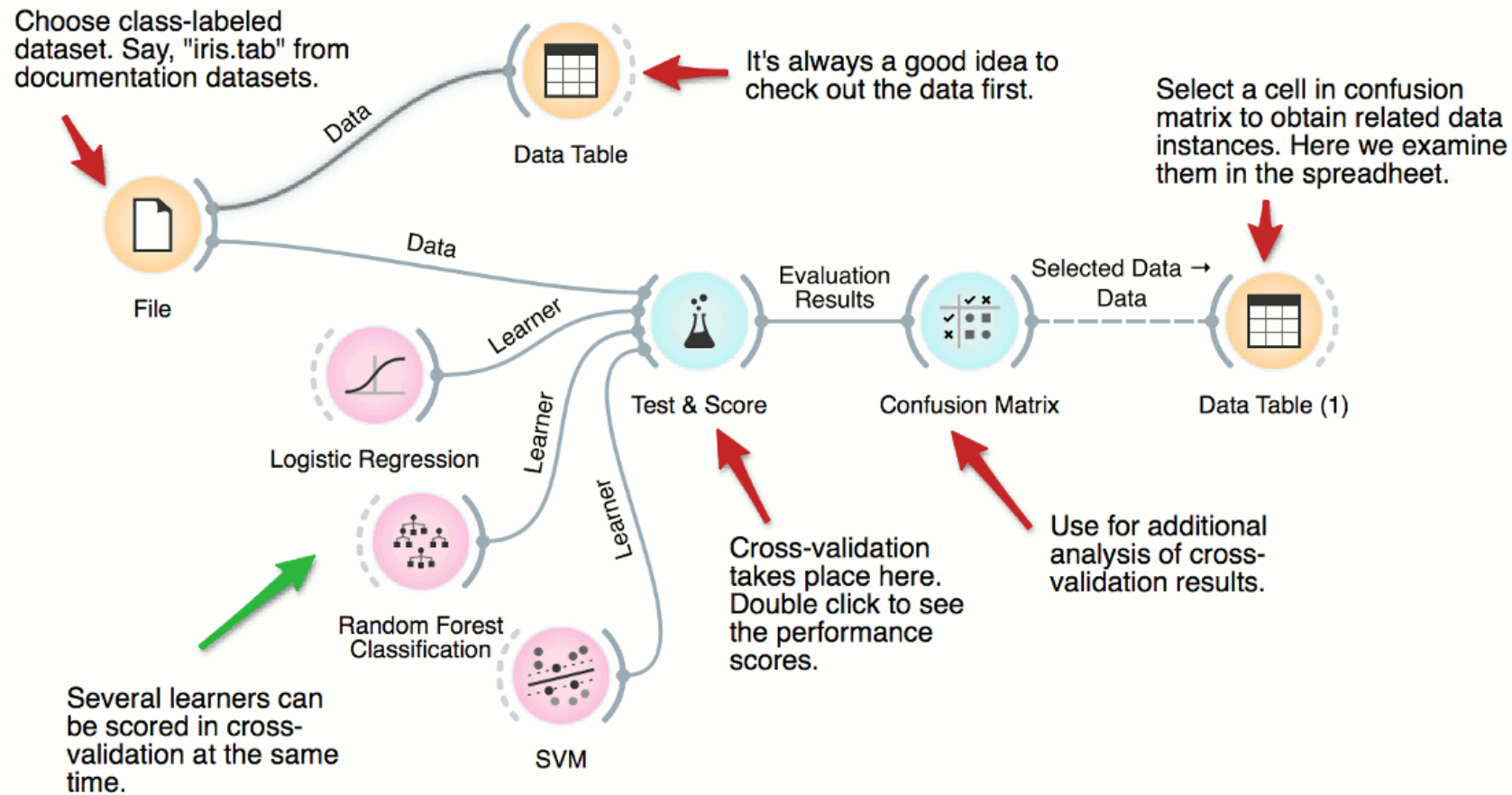


# Domestic Animals Dataset



<https://tinyurl.com/orange-animals>

# Self study: How good are data mining methods on your classification dataset?



# Let's discuss!



<https://wooclap.com/SAMEN25>

