

# Orange EDA - Data Bunga Iris

Instruktur: Muhammad Hilman, Ph.D

Penyusun: Fariz Darari, Ph.D

## Prasyarat

Pastikan aplikasi **Orange** sudah terpasang dengan baik di komputer anda.

## 1 Pendahuluan

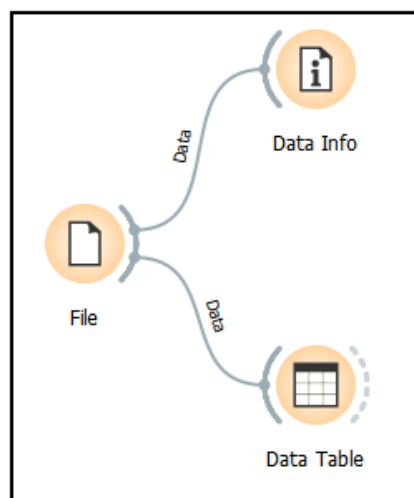
Bunga iris merupakan salah satu jenis bunga dari famili Iridaceae yang terdiri dari 300 spesies. Bunga ini memiliki warna yang beragam dan mencolok, kecuali merah terang. Keindahannya tersebar di seluruh belahan dunia dengan warna dominan yang berbeda di setiap wilayahnya.

Warnanya yang beragam menjadikan bunga iris dipercaya berkaitan dengan mitologi Yunani Kuno. Menurut Interflora, Iris merupakan dewi berwujud hujan yang diutus untuk menyampaikan pesan antara surga dan bumi dengan pelangi sebagai perantara.

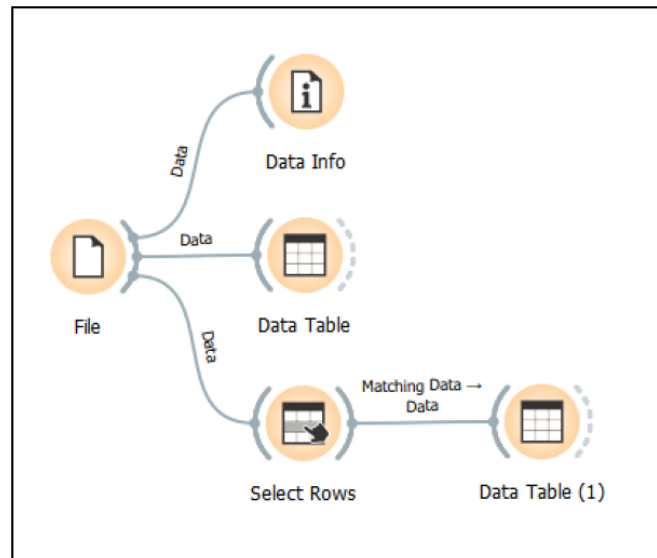
## 2 Eksplorasi Data

Pada EDA kali ini, kita akan mengeksplorasi data bunga iris yang tersedia secara *default* pada aplikasi **Orange**. Silahkan ikuti petunjuk dan jawab pertanyaan di bawah ini:

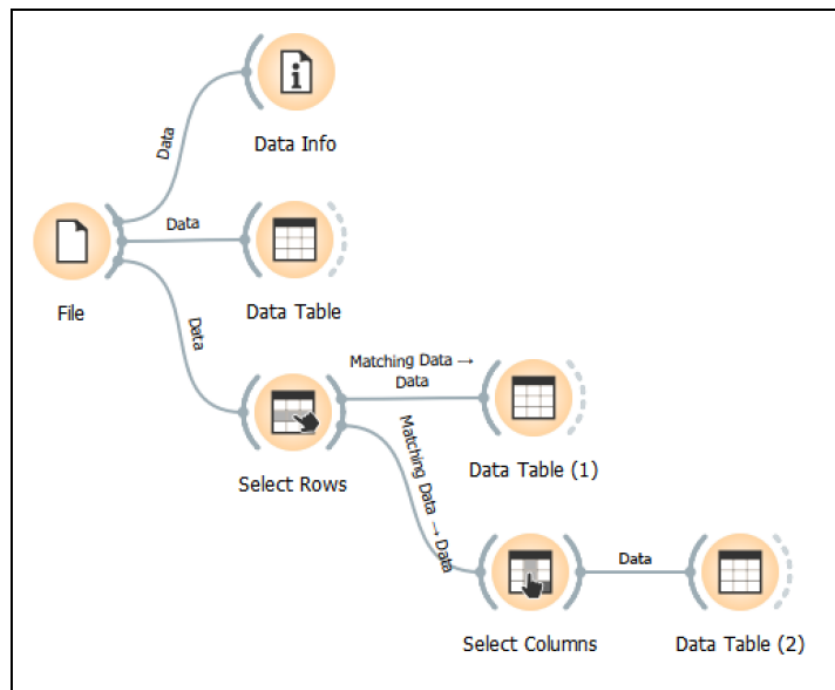
1. Buka aplikasi **Orange** dan pilih **New**.
2. Tambahkan *widget File* pada canvas. Klik *widget* tersebut dan pilih *source file iris.tab*
3. Perhatikan bagian **Info** pada *widget File*. Apa yang dapat anda observasi? Ada berapa jumlah baris pada data bunga iris? Ada berapa atribut / *feature*? Apa saja *feature* yang dapat anda temukan?
4. Tambahkan *widget Data Info* dan *Data Table* kemudian hubungkan keduanya dengan *widget File*. Observasi isi dari masing-masing *widget* tersebut



5. Tambahkan *widget Select Rows* kemudian hubungkan dengan *widget File*. Pada *widget Select Rows* pilih **Iris-versicolor** sebagai nilai dari kondisi *class iris*. Tambahkan *widget Data Table* yang baru dan hubungkan dengan *widget Select Rows*. Apa yang dapat anda observasi?



6. Tambahkan *widget Select Columns* dan hubungkan dengan *widget Select Rows*. Pada *widget Select Columns*, pilih untuk mengabaikan (*ignore*) *feature petal length* dan *petal width* dengan cara *menge-drag* kedua *feature* tersebut ke sebelah kiri. Selanjutnya, tambahkan dan hubungkan *widget Select Columns* tersebut dengan *widget Data Table* yang baru. Apa yang dapat anda observasi?



7. Tambahkan *widget Save Data* dan hubungkan dengan *widget Data Table (2)* yang sudah anda buat. *Widget Save Data* dapat digunakan untuk meng-*export* data, misalnya ke dalam format **csv**. Simpan dengan nama **Iris-versicolor.csv**. Coba buka berkas tersebut dengan aplikasi *spreadsheet* misalnya Ms. Excel.
8. Workflow yang sudah dibuat dapat disimpan dengan perintah **Save As** pada aplikasi **Orange**. Beri nama berkas ini dengan **data-processing-basics.ows**.

### 3 Visualisasi Data

Pada bagian ini, kita akan melakukan visualisasi terhadap data bunga iris yang sudah dieksplorasi pada bagian sebelumnya.

1. Buka berkas **data-processing-basics.ows** menggunakan **Orange**. Kemudian **Save As** dengan nama **eda.ows**.
2. Untuk melakukan visualisasi, kita akan mulai dengan mencoba *widget* **Feature Statistics**. *Widget* ini menyediakan cara cepat untuk melakukan inspeksi dan menemukan fitur-fitur menarik pada data. Silahkan tambahkan *widget* ini kemudian hubungkan ke *widget* **Data Table**, **Data Table (1)**, maupun **Data Table (2)**. Apa yang dapat anda observasi?
3. Selanjutnya, mari kita pindah ke section **Visualize** dan tambahkan *widget* **Box Plot** dan hubungkan dengan data seperti yang sudah dilakukan sebelumnya. Apa yang dapat anda observasi?
4. *Widget* yang tidak kalah menarik adalah **Violin Plot**. *Widget* ini mirip seperti **Box Plot** dengan menampilkan perbedaan distribusi dari data yang kita plot.
5. Untuk visualisasi yang lebih detail, anda dapat mencoba *widget* **Distributions** yang menampilkan distribusi *value* untuk suatu *feature*/atribut.
6. Pada bagian terakhir dari aktivitas ini, anda diminta untuk mengeksplorasi data dengan *widget* **Scatter Plot**. *Widget* ini akan menampilkan visualisasi relasi antara dua *feature* pada data.