

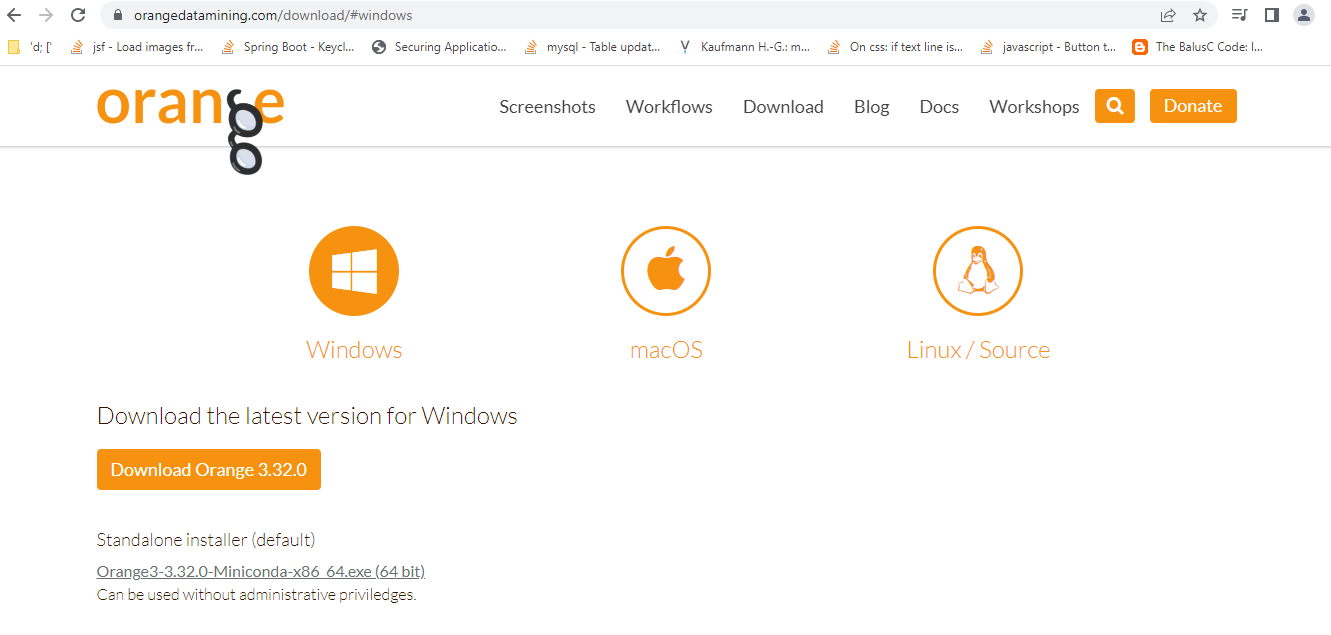
Pada modul kali ini kita akan mencoba Instalasi Tools untuk Data Mining yaitu Orange Data Mining. Data Mining salah satu teknik implementasi dari pembelajaran mesin (*machine learning*) yang saat ini banyak digunakan di berbagai bidang. Data mining menjadi salah satu primadona teknik pemodelan dan pengambilan keputusan berdasarkan data yang dapat menghasilkan wawasan (*insight*) dan perkiraan/prediksi (*forecast*) yang dapat membantu kita di kehidupan sehari-hari. Dengan aplikasi Orange Data Mining kita bisa melakukan proses data mining tanpa perlu menulis kode program sama sekali.

1. **Orange Data Mining**

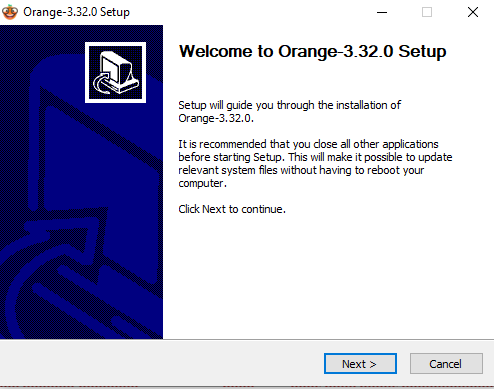
Orange merupakan aplikasi yang dikembangkan oleh University of Ljubljana, dengan target menghadirkan sebuah aplikasi untuk mengolah data secara visual tanpa perlu pengalaman membuat program sebelumnya. Pada Orange, untuk melakukan data mining kita akan menggunakan sistem widget. Setiap widget punya fungsinya masing-masing dan dapat menerima input atau mengeluarkan output. Misalnya kita ingin membaca data dari sebuah file, kita bisa menggunakan widget File untuk membaca data dari file tersebut dan menggunakan widget Data Table untuk menampilkan isi data yang sudah dibaca. Sama halnya jika kita ingin membuat sebuah model untuk melakukan data mining seperti klasifikasi atau regresi. Kita bisa menambahkan widget seperti K-nearest neighbors (KNN) dan Naive Bayes untuk membuat model dan memberikan data ke dalam model tersebut dengan cara menghubungkan sumber data dengan model dengan cara menarik garis penghubung.

1. **Instalasi Orange Data Mining**

Untuk unduh Orange Data Mining bisa melalui alamat URL berikut ini <https://orangedatamining.com/download/#windows>. Pada modul ini menggunakan sistem operasi windows, selain operasi windows silakan disesuaikan dengan sistem operasi masing-masing.

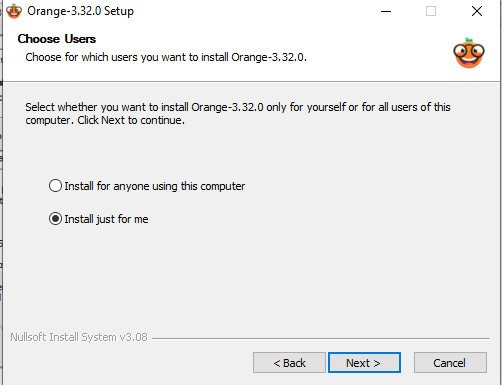


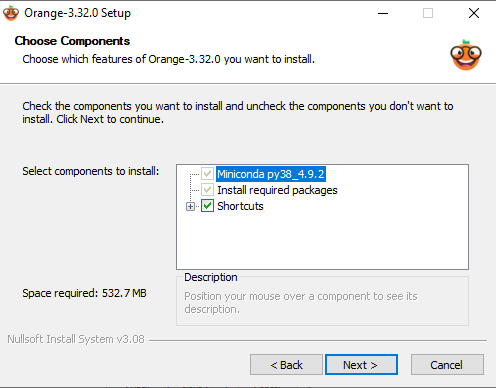
Setelah selesai unduh silakan di double klik file nya lalu ikuti tahap-tahap selanjutnya sampai aplikasi sukses terinstall.

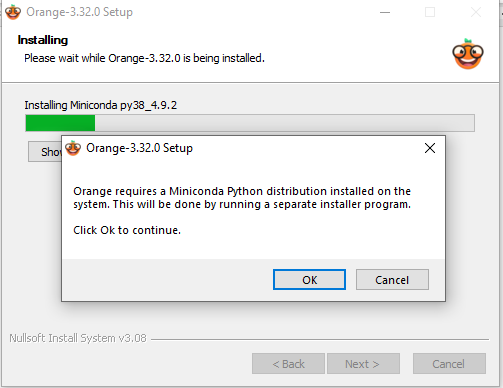


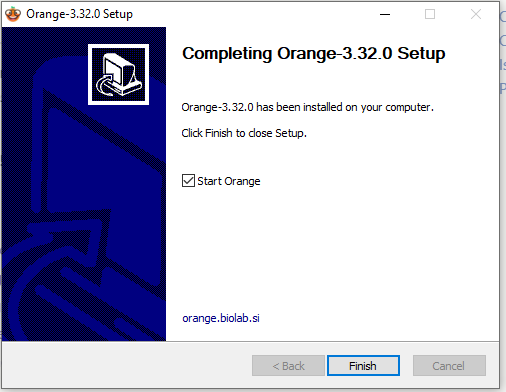
Graphical user interface, text, application, email

Description automatically generated

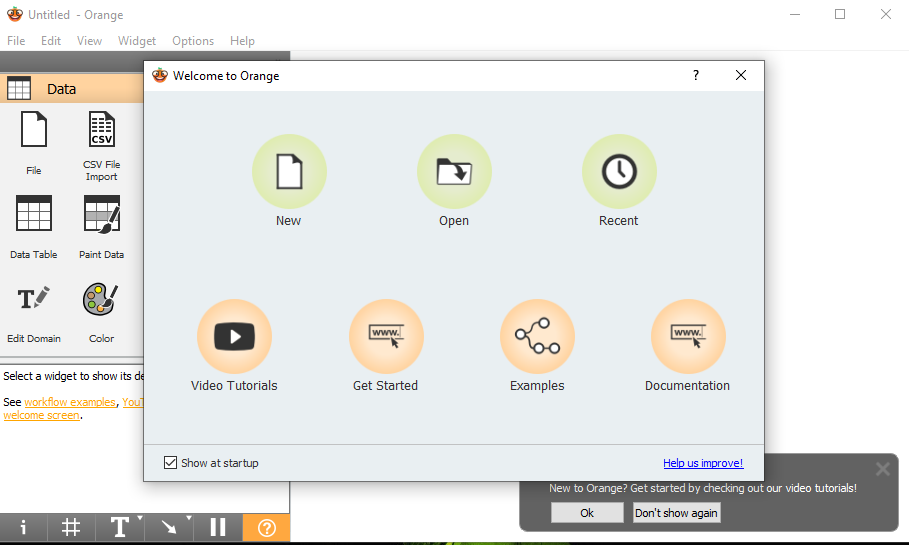








Berikut tampilan awal orange



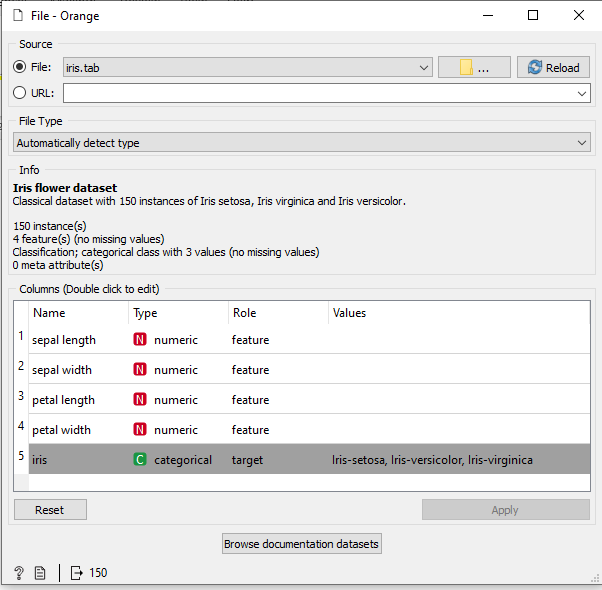
1. **Penggunaan Orange Data Mining**

* Pilih widget data kemudian klik file.

Graphical user interface, application

Description automatically generated

* Double klik file tersebut lalu pilih file dan sumber data yang diinginkan, disini kita menggunakan data set yang sudah disediakan oleh orange.

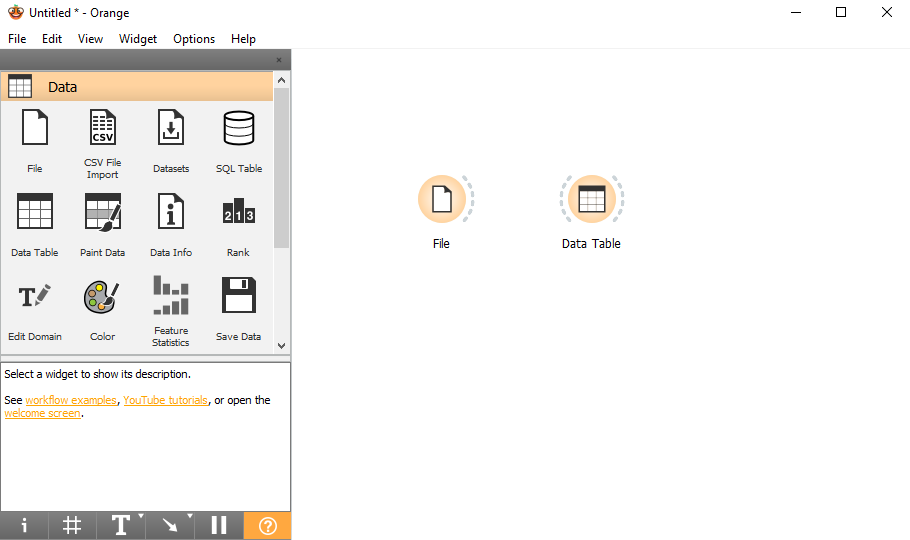


Graphical user interface, text

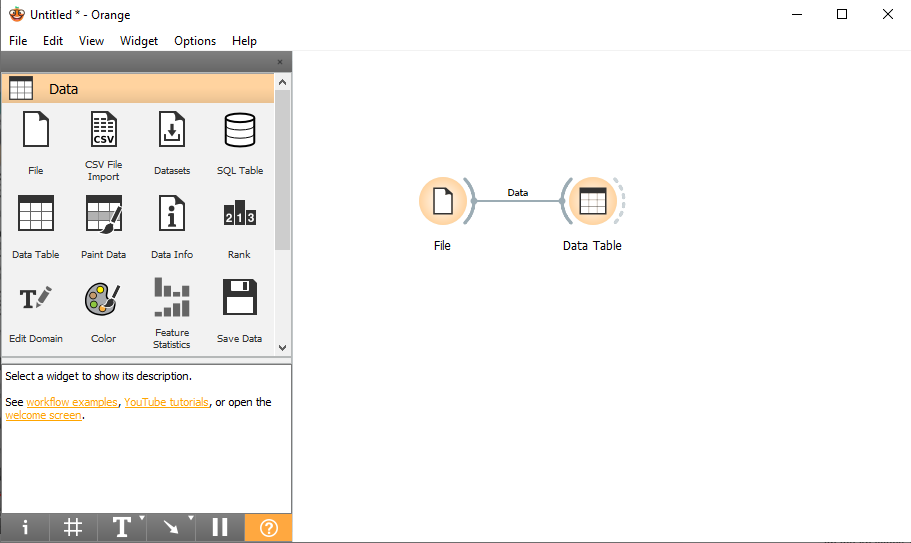
Description automatically generated

Pilih iris.tab kemudian setelah dipilih klik tanda silang (close) dialog file tersebut

* Klik widget Data Table



* Lalu tarik garis dari kanan (output) widget file ke kiri (input) widget Data Table



* Kemudian klik Data Table, akan tampil data berupa tabel

Table

Description automatically generated with medium confidence

* Tambahkan widget pivot table, lalu hubungkan garis dari kanan (output) widget file ke kiri (input) widget pivot table dan double klik widget pivot table.

Graphical user interface, application

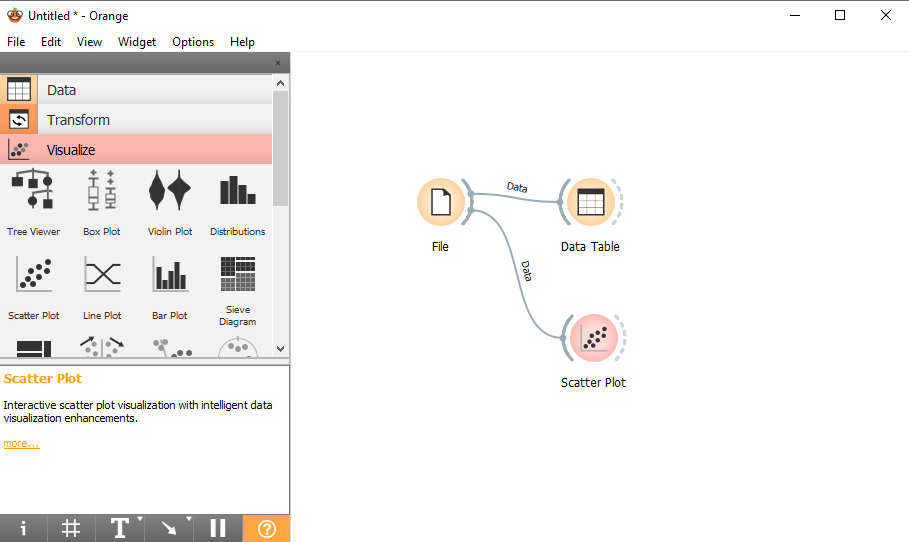
Description automatically generated

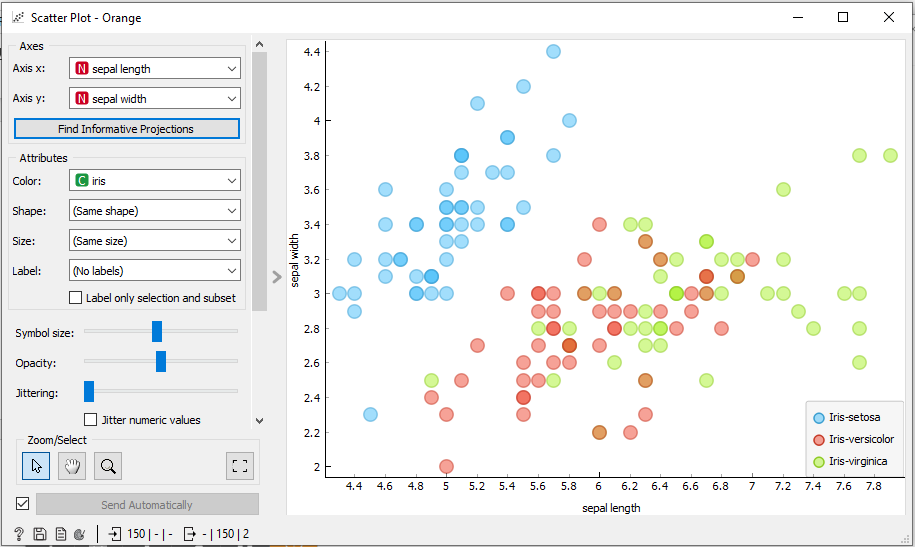
A screenshot of a computer

Description automatically generated

Silakan ceklist count, sum, mean, median, min, max sehingga tampil seperti gambar diatas.

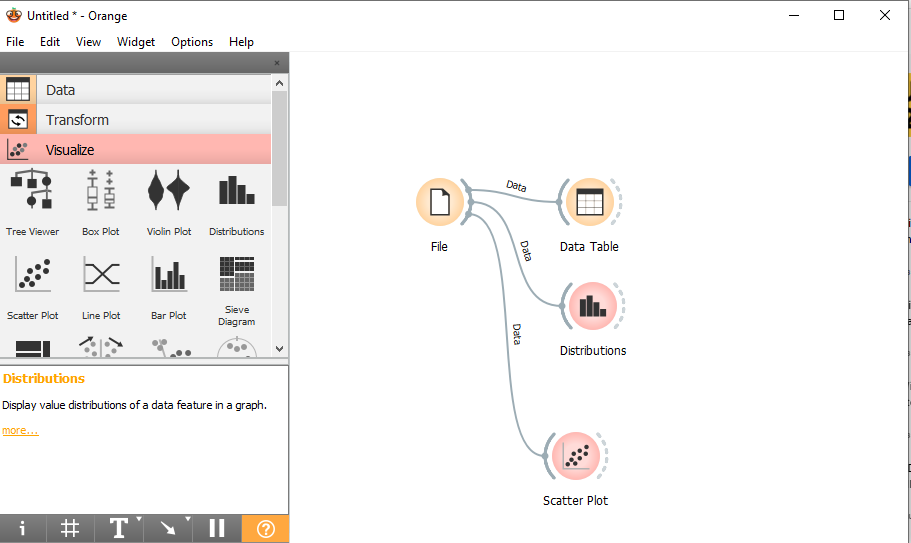
* Tambahkan widget Visualize Scatter Plot, lalu hubungkan garis dari kanan (output) widget file ke kiri (input) widget Scatter Plot dan double klik widget Scatter Plot.

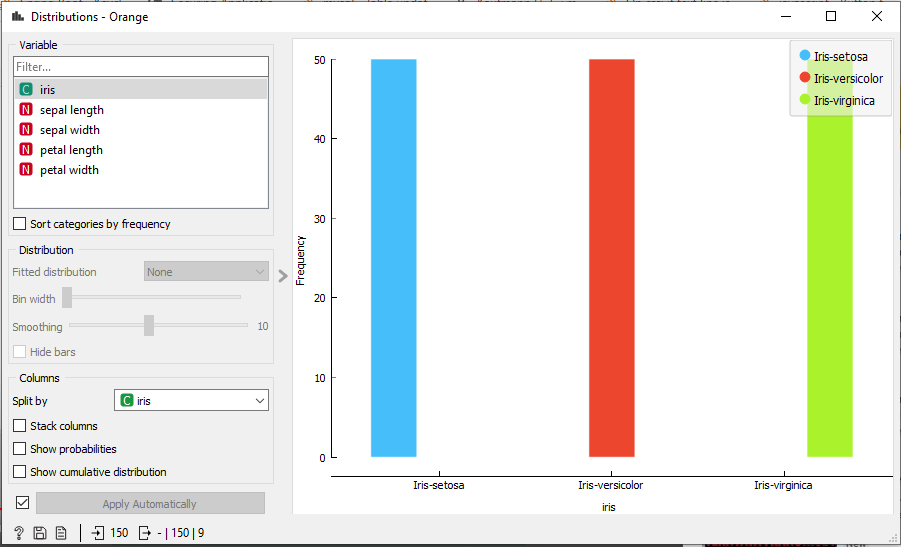




Akan terlihat distribusi data yang ditampilkan dalam diagram.

* Tambahkan widget Visualize Distribution, lalu hubungkan garis dari kanan (output) widget file ke kiri (input) widget Distribution dan double klik widget Distribution.

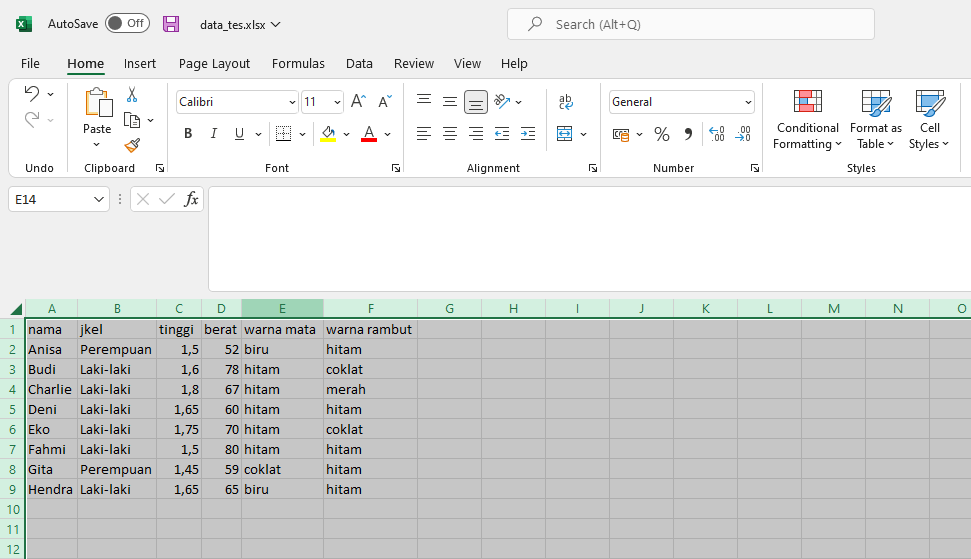




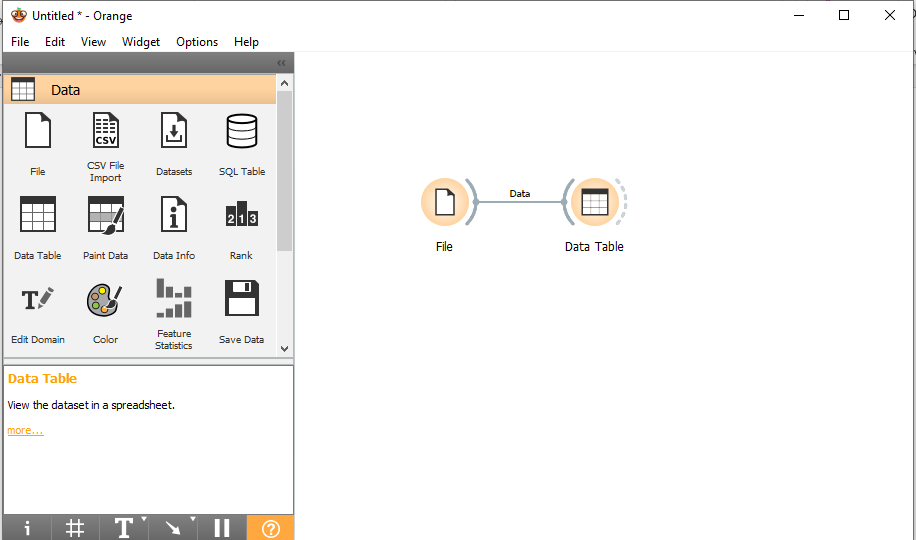
* Tambahkan widget Visualize Distribution, lalu hubungkan garis dari kanan (output) widget file ke kiri (input) widget Distribution dan double klik widget Distribution.

1. **Loading Data Orange Data Mining**

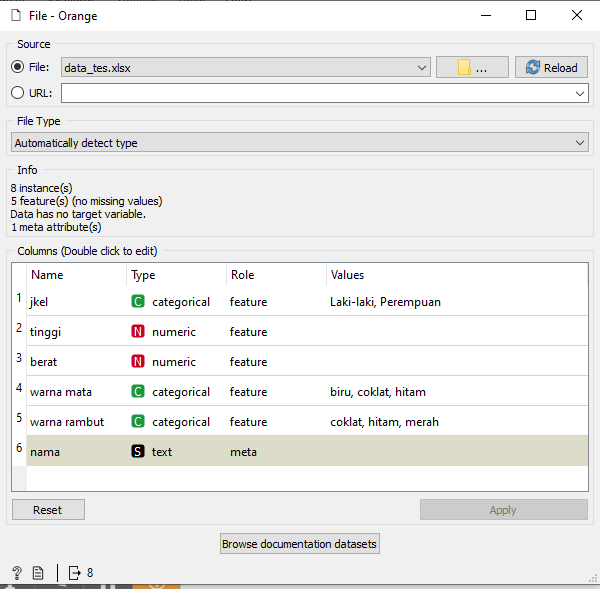
* Buat file excel seperti dibawah ini



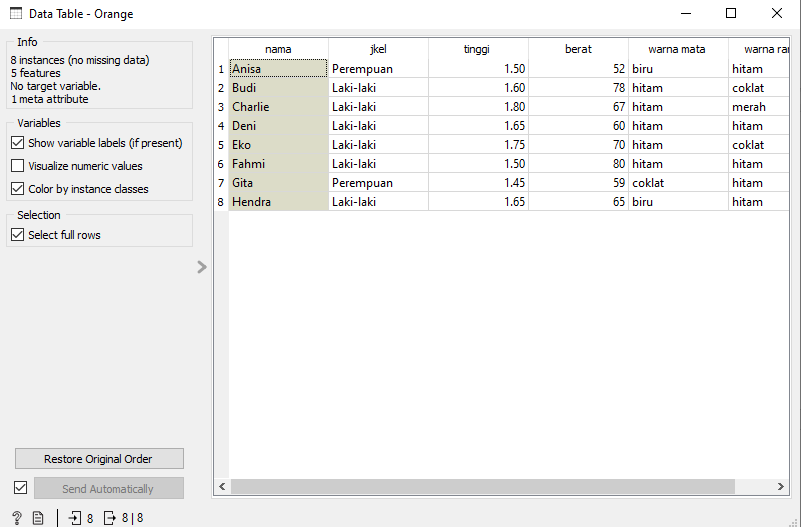
* Lalu buat kan pada orange widget file dan data table.



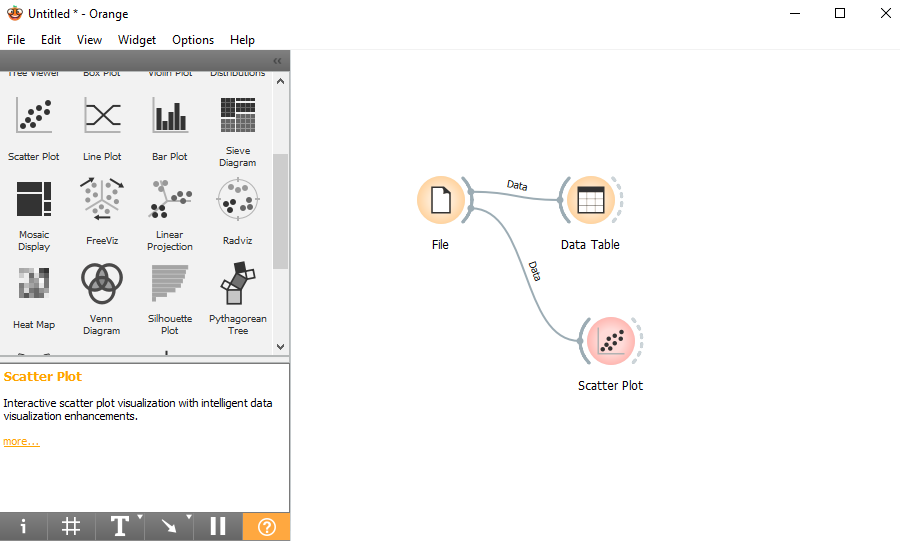
* Double klik pada widget file, kemudian pilih file excel yang tadi sudah dibuat

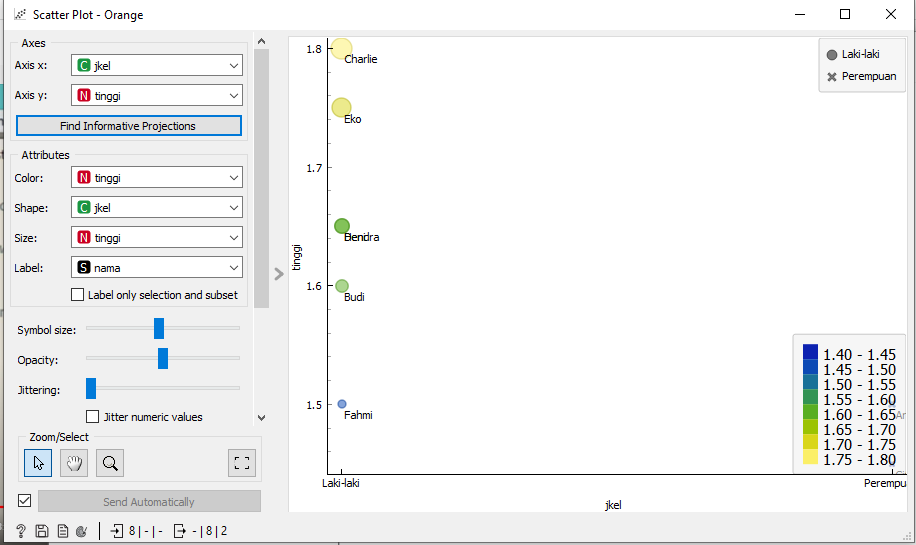


* Kemudian close lalu klik data table maka akan terlihat data yang ada di excel.



* Tambahkan widget scatter plot lalu bisa dilihat data ditampilkan dalam bentuk diagram.





1. **Making Predictions**

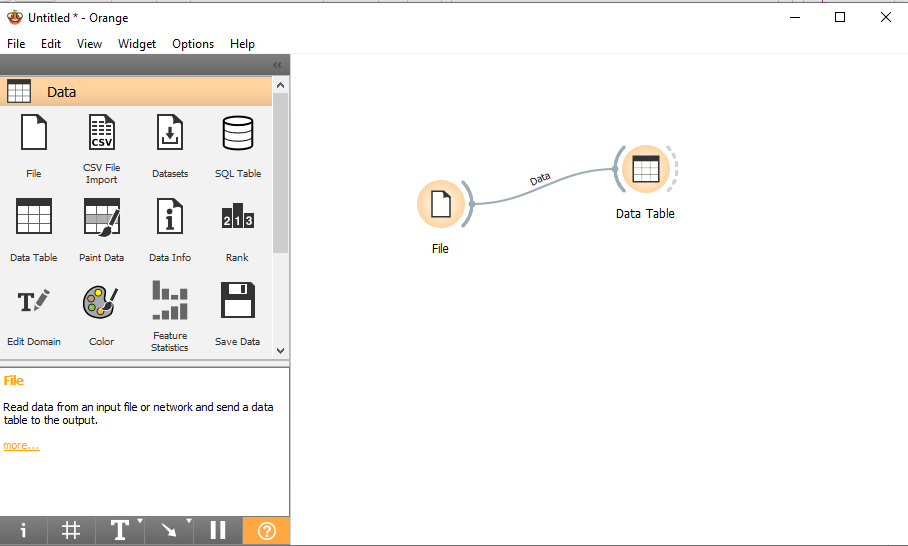
* Download data set berikut ini untuk melakukan prediksi dengan *classification tree* dan *logistic regression*

<https://raw.githubusercontent.com/ajdapretnar/datasets/master/data/fruits-and-vegetables-train.tab>

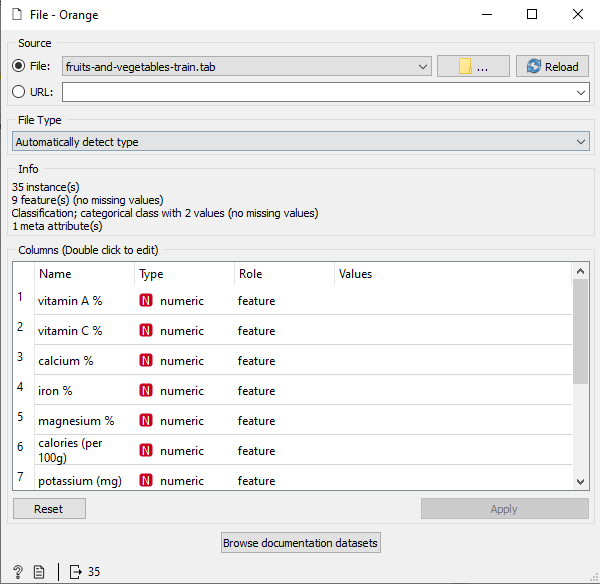
<https://raw.githubusercontent.com/ajdapretnar/datasets/master/data/fruits-and-vegetables-test.tab>

simpan dengan ekstensi .tab

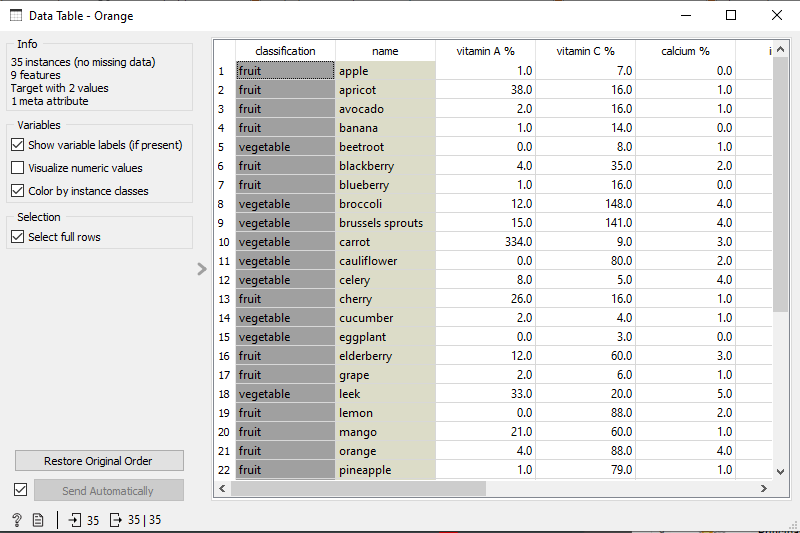
* Buat file new baru untuk membuat widget seperti berikut



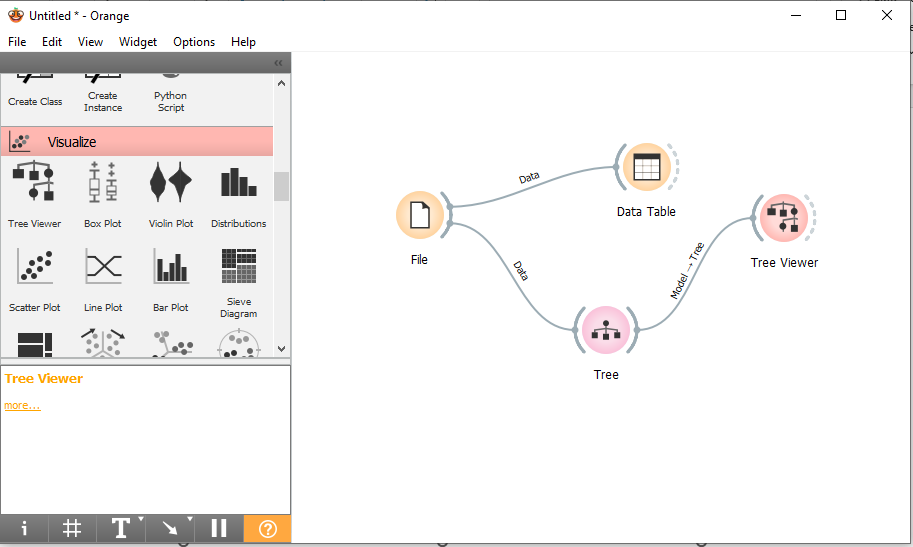
* Sama seperti sebelumnya ambil data set pada widget file yaitu fruits-and-vegetables-train.tab



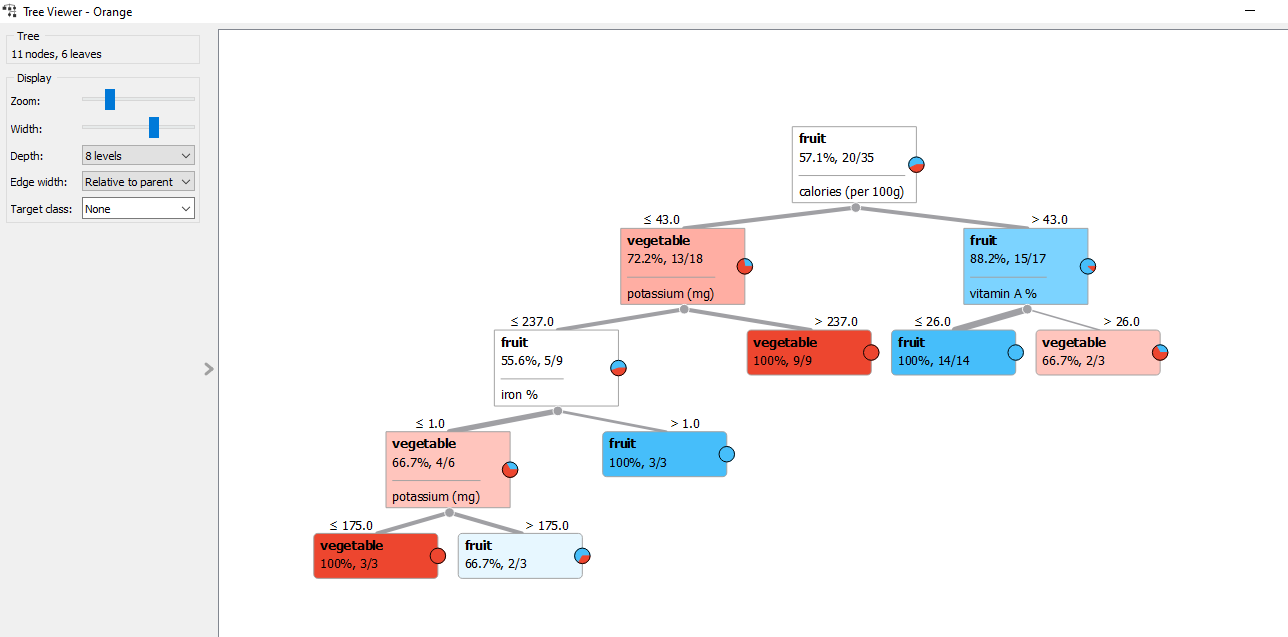
* Double klik widget Data Table maka isi data akan tampil seperti ini.



* Tambahkan widget tree dan tree viewer sehingga menjadi seperti ini



* Double klik tree viewer nanti akan terlihat klasifikasi data sesuai data set



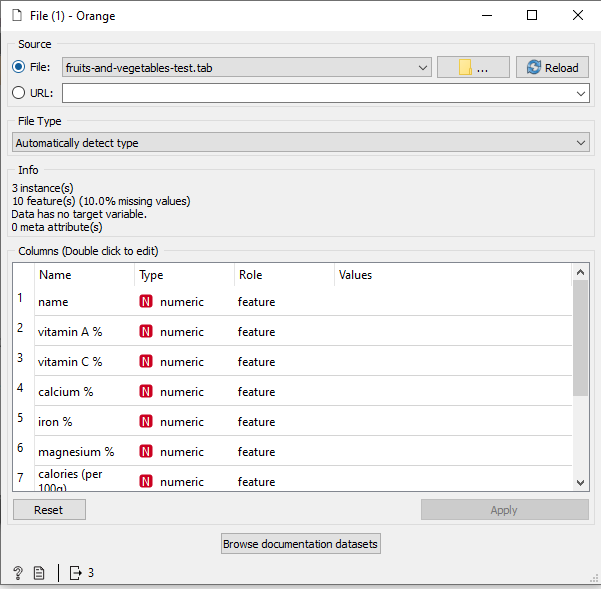
Pada gambar diatas terlihat klasifikasi fruit dan vegetables sesuai dengan nilai masing-masing parameter. Pertama adalah kalori, jika kalori lebih kecil sama dengan 43 maka masuk kategori vegetables lalu jika lebih besar 43 masuk kategori fruit. Nilai 43 adalah nilai yang tertera pada dataset untuk kalori pada fruit dan vegetables. Kemudian selanjutnya adalah klasifikasi sesuai parameter yang lain.

* Selanjutnya tambahkan dibagian bawah seperti bagan berikut, buat widget file, data table dan predictions untuk melakukan *logistic regression*.

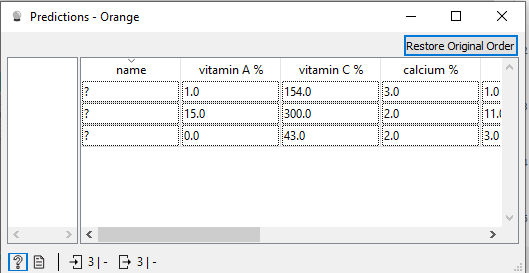
A picture containing diagram

Description automatically generated

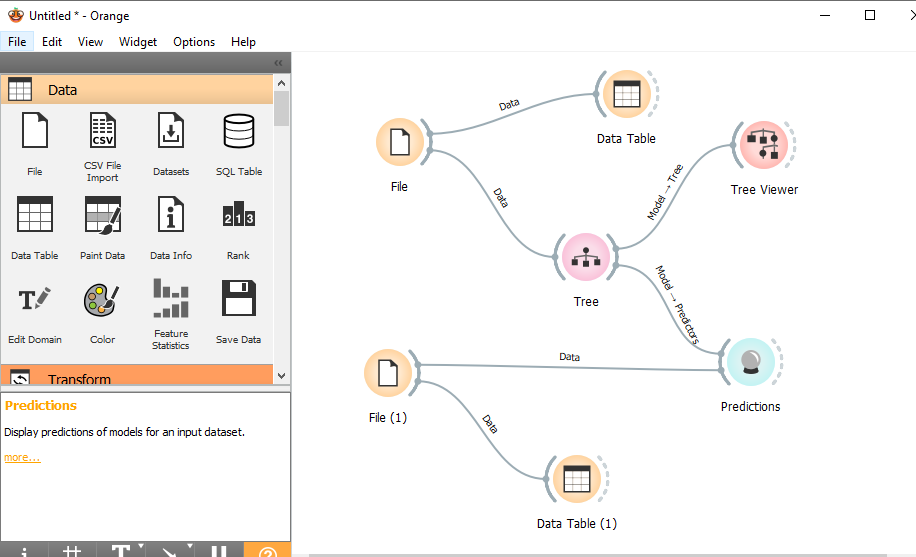
* Masukan data set pada file (1) fruits-and-vegetables-test.tab



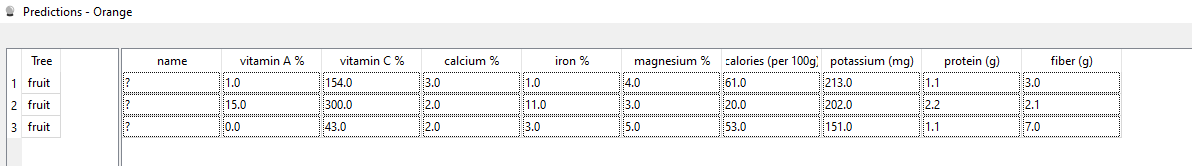
* Double klik predictions



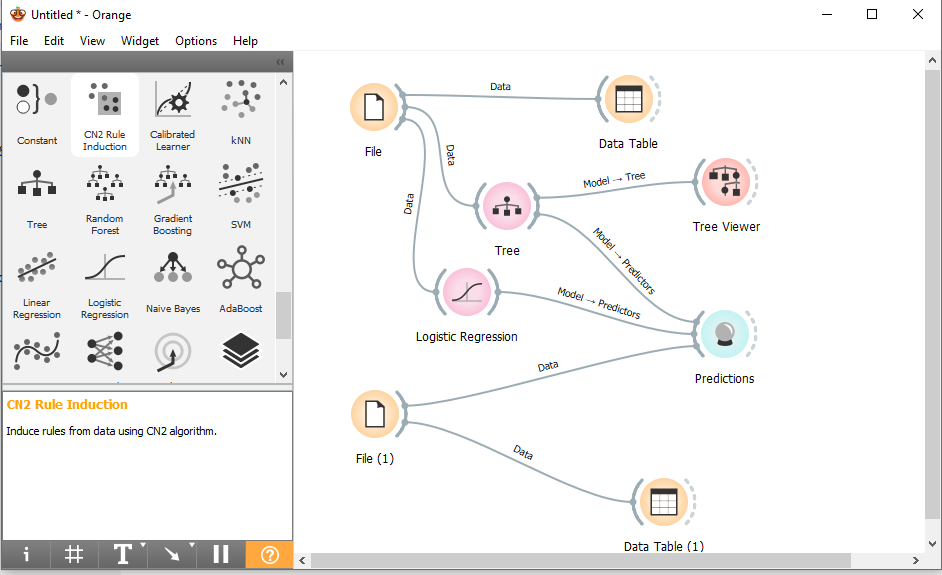
* Sambungkan output tree ke input dari predictions seperti gambar berikut



* Lakukan double klik kembali pada predictions



* Tambahkan logistic regression sesuai gambar lalu lihat hasilnya pada predictions



A picture containing table

Description automatically generated

Terlihat prediksi untuk data mengunakan tree dan juga logistic regression. Dari sini bisa diambil kesimpulan dengan tools orange ini user bisa melakukan beberapa proses terkait data mining selama data set tersedia.

*----- Sekian Terima Kasih -----*

* <https://orangedatamining.com/>