

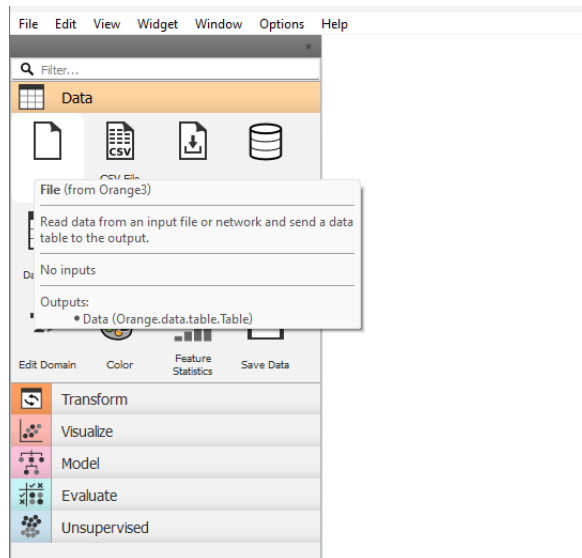
## EKSPLORASI KLASIFIKASI DENGAN ORANGE

Anggota Kelompok :

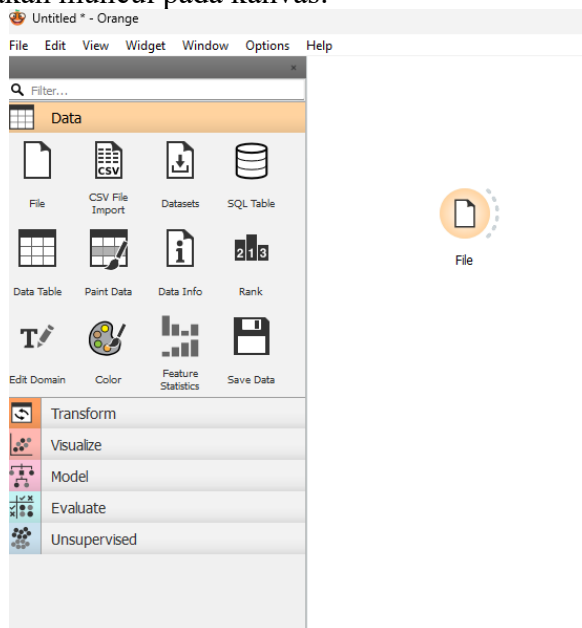
- Marta Yulianti / 215314154
- Victoria Alysha F S / 215314158
- Resiana Kinanti Jati / 215314159
- Loadtriani Oktavia S / 215314172

Langkah-langkah membaca dan menampilkan dataset dalam Orange. Dataset boleh dari sampel dalam Orange (misal Iris, titanic, weather) atau data dari sumber lain.

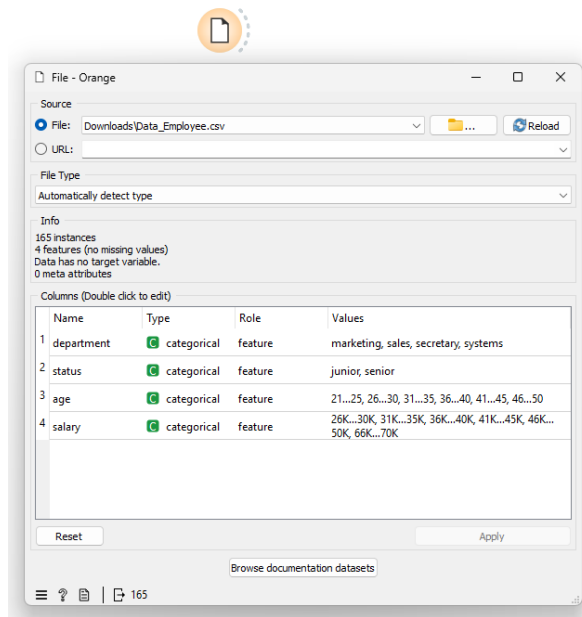
1. Pada Orange, Pilih Data yang berada pada sebelah kiri dan pada Data pilih file dengan cara klik pada file.



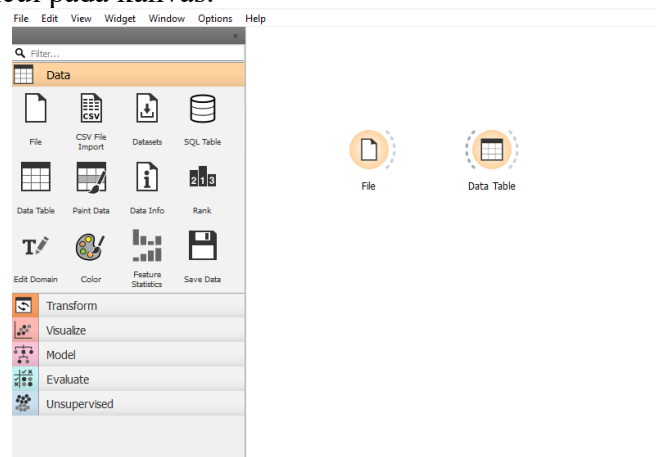
2. Setelah diklik, file akan muncul pada kanvas.



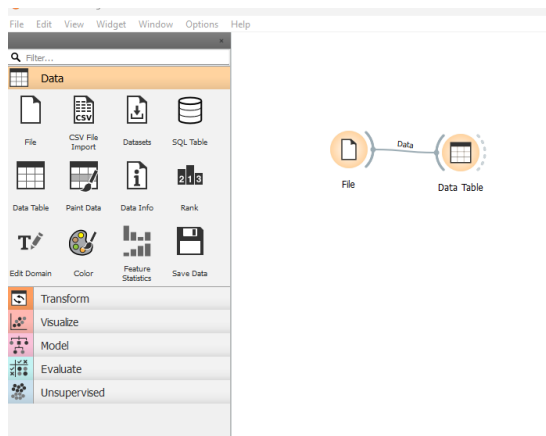
3. Kemudian klik 2 kali pada widget/icon yang berada pada kanvas untuk memilih data apa yang ingin digunakan. Disini menggunakan dataset Data\_Employee.csv



4. Jika ingin melihat Data table, pada Data pilih dengan cara klik Data table. Setelah diklik, akan muncul pada kanvas.



5. Untuk dapat melihat isi file, Seret garis dari sisi kanan file kemudian hubungkan ke sebelah kiri data file.

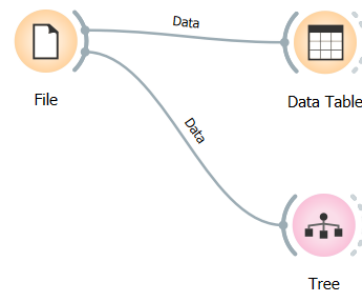
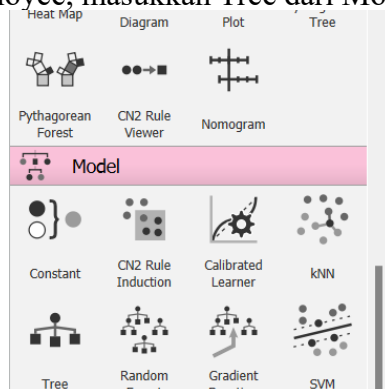


6. Kemudian klik 2 kali untuk melihat isi dari Data Table dari dataset yang digunakan

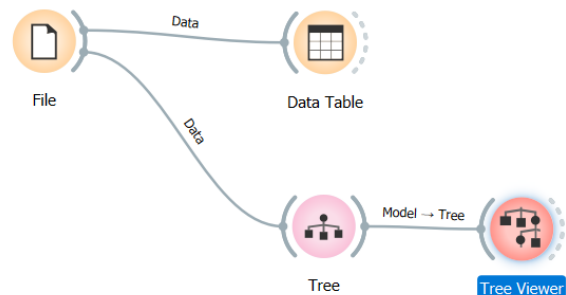
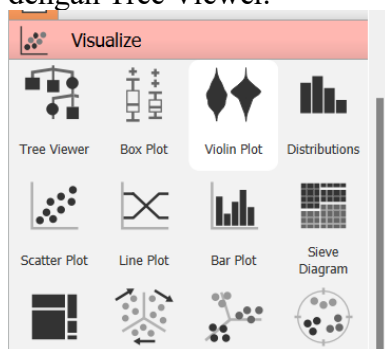
	department	status	age	salary
1	sales	senior	31...35	46K...50K
2	sales	senior	31...35	46K...50K
3	sales	senior	31...35	46K...50K
4	sales	senior	31...35	46K...50K
5	sales	senior	31...35	46K...50K
6	sales	senior	31...35	46K...50K
7	sales	senior	31...35	46K...50K
8	sales	senior	31...35	46K...50K
9	sales	senior	31...35	46K...50K
10	sales	senior	31...35	46K...50K
11	sales	senior	31...35	46K...50K
12	sales	senior	31...35	46K...50K
13	sales	senior	31...35	46K...50K
14	sales	senior	31...35	46K...50K
15	sales	senior	31...35	46K...50K
16	sales	senior	31...35	46K...50K
17	sales	senior	31...35	46K...50K
18	sales	senior	31...35	46K...50K
19	sales	senior	31...35	46K...50K
20	sales	senior	31...35	46K...50K
21	sales	senior	31...35	46K...50K

### Langkah-langkah menghasilkan decision tree dari suatu dataset.

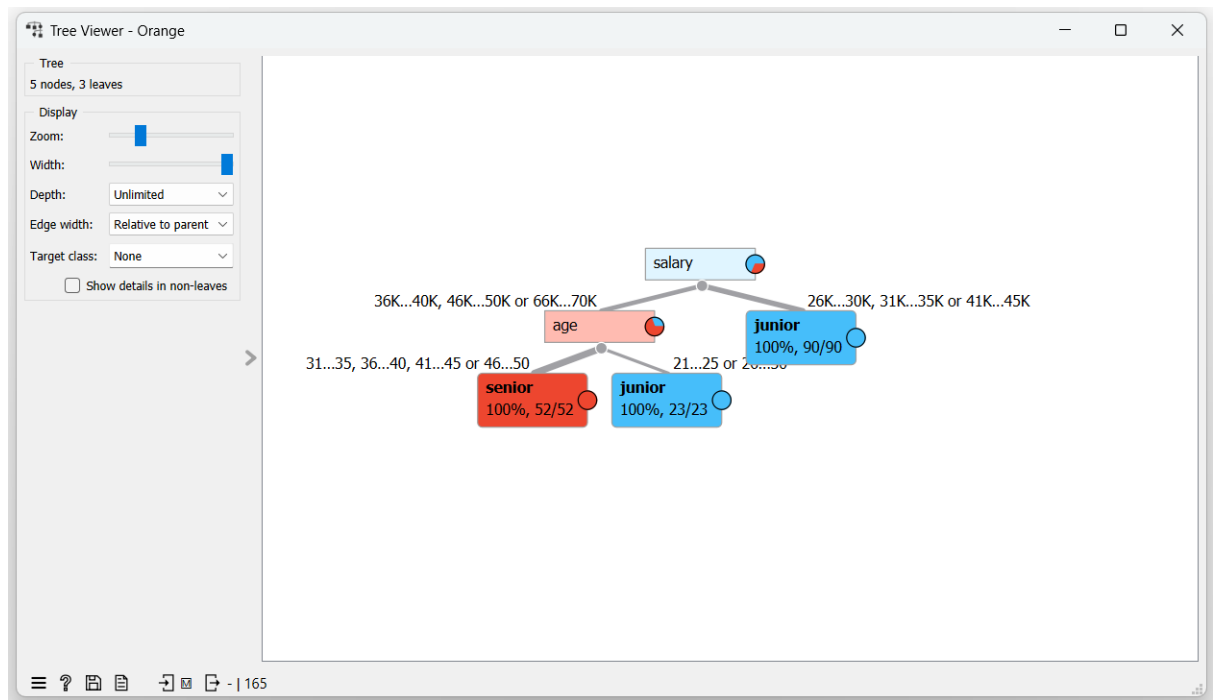
1. Pilih data yang akan dibaca dan ditampilkan menggunakan cara pada langkah-langkah sebelumnya. Setelah file dipilih dan dapat dibaca, pada kasus ini memakai data Employee, masukkan Tree dari Model, untuk membentuk model.



2. Setelah model terbentuk, selanjutnya akan divisualisasikan dalam bentuk decision tree dengan cara memasukkan Tree Viewer dari Visualize ke kanvas, lalu sambungkan Tree dengan Tree Viewer.



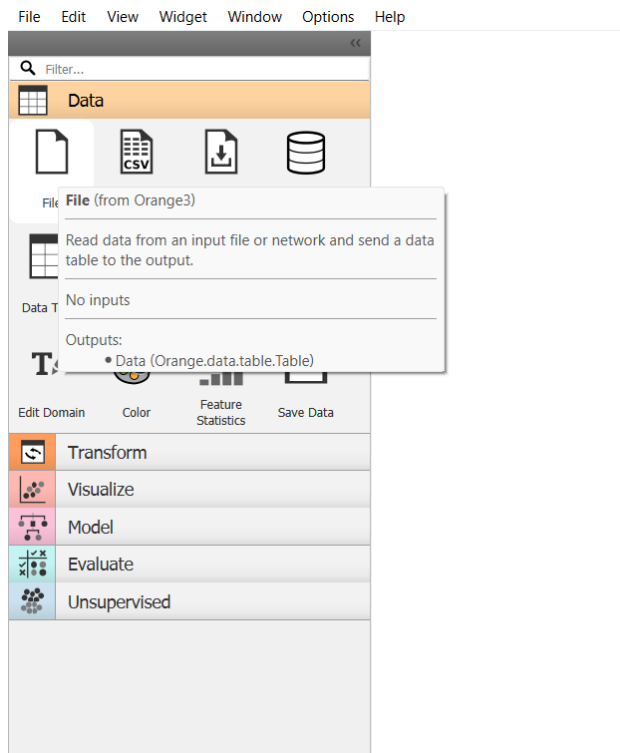
3. Klik dua kali pada Tree Viewer, nanti akan ada window yang berisi decision tree berdasarkan data Employee.



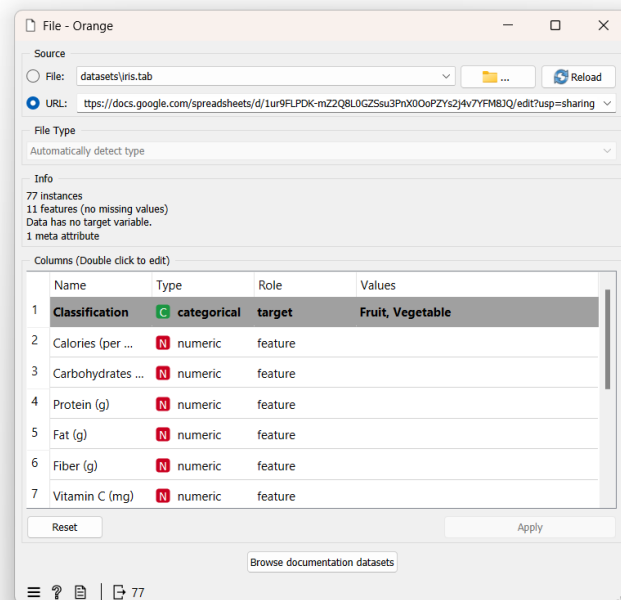
Langkah-langkah melakukan prediksi terhadap data test yang belum diketahui labelnya. Data tes boleh membuat sendiri, boleh mengambil dari sample Orange.

Kami menggunakan dataset fruit and veggie nutrition yang diambil dari Kaggle untuk melatih pengklasifikasiannya. Kumpulan data pelatihan kami terdiri dari sembilan fitur, termasuk jumlah kalori, protein, serat, serta kandungan vitamin dan mineral lainnya. Berdasarkan fitur-fitur ini, kami ingin memprediksi apakah suatu tanaman adalah fruit atau vegetable.

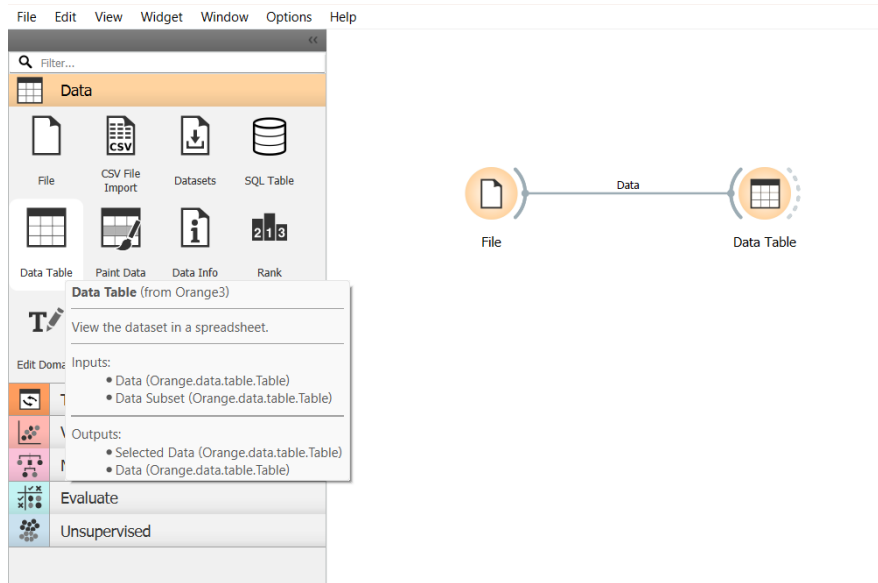
1. Pada Orange, Pilih Data yang berada pada sebelah kiri dan pada Data pilih file dengan cara klik pada file.



2. Maka file akan muncul di kanvas. Klik dua kali untuk memuat file data yang akan digunakan. Data diperoleh dengan menggunakan URL spreadsheet. Kemudian, kolom pada dataset akan digunakan secara lengkap, sehingga tidak ada perubahan yang dilakukan.



3. Untuk dapat melihat isi file, Seret garis dari sisi kanan file kemudian hubungkan ke sebelah kiri data file.



4. Kemudian klik 2 kali untuk melihat isi dari Data Table dari dataset yang digunakan.

Data Table - Orange

Info  
77 instances (no missing data)  
10 features  
Target with 2 values  
1 meta attribute

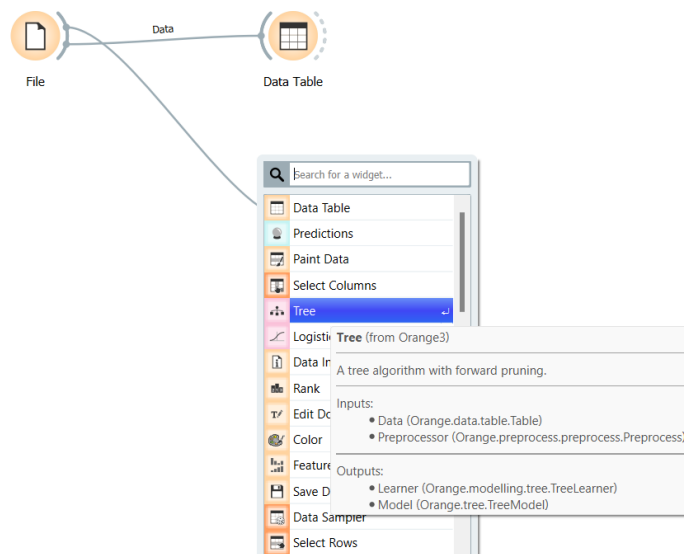
Variables  
☒ Show variable labels (if present)  
☐ Visualize numeric values  
☒ Color by instance classes

Selection  
☒ Select full rows

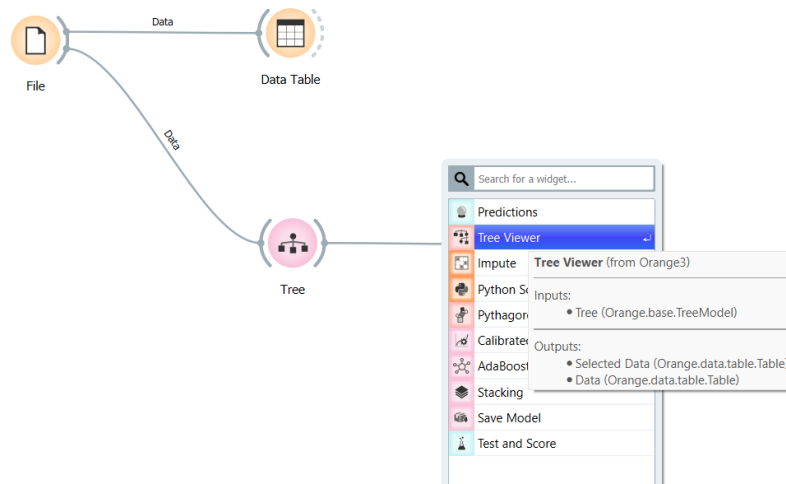
Restore Original Order  
☒ Send Automatically

	Classification	Name	Calories (per 100g)	Carbohydrates (g)	Protein (g)	Fat
1	Vegetable	Amaranth	23	4.00	2.10	
2	Fruit	Apple	52	14.00	0.30	
3	Vegetable	Apricot	48	11.00	1.40	
4	Vegetable	Asparagus	20	3.70	2.20	
5	Fruit	Avocado	160	8.50	2.00	
6	Fruit	Banana	89	23.00	1.30	
7	Vegetable	Beans	31	7.00	1.80	
8	Vegetable	Beetroot	43	9.60	1.60	
9	Fruit	Bell Pepper	20	4.60	0.90	
10	Vegetable	Bitter Gourd	19	4.30	0.80	
11	Fruit	Black Berry	43	9.60	1.40	
12	Fruit	Black Currant	63	15.00	0.90	
13	Fruit	Blueberry	57	14.50	0.70	
14	Vegetable	Bottle Gourd	15	3.40	0.60	
15	Fruit	Brinjal	25	5.90	1.00	
16	Vegetable	Broccoli	55	11.20	3.70	
17	Vegetable	Cabbage	25	6.00	1.30	
18	Fruit	Capsicum	20	4.60	0.90	
19	Vegetable	Carrot	41	10.00	0.90	
20	Fruit	Cashew	553	30.19	18.22	
21	Vegetable	Cauliflower	25	5.00	1.90	

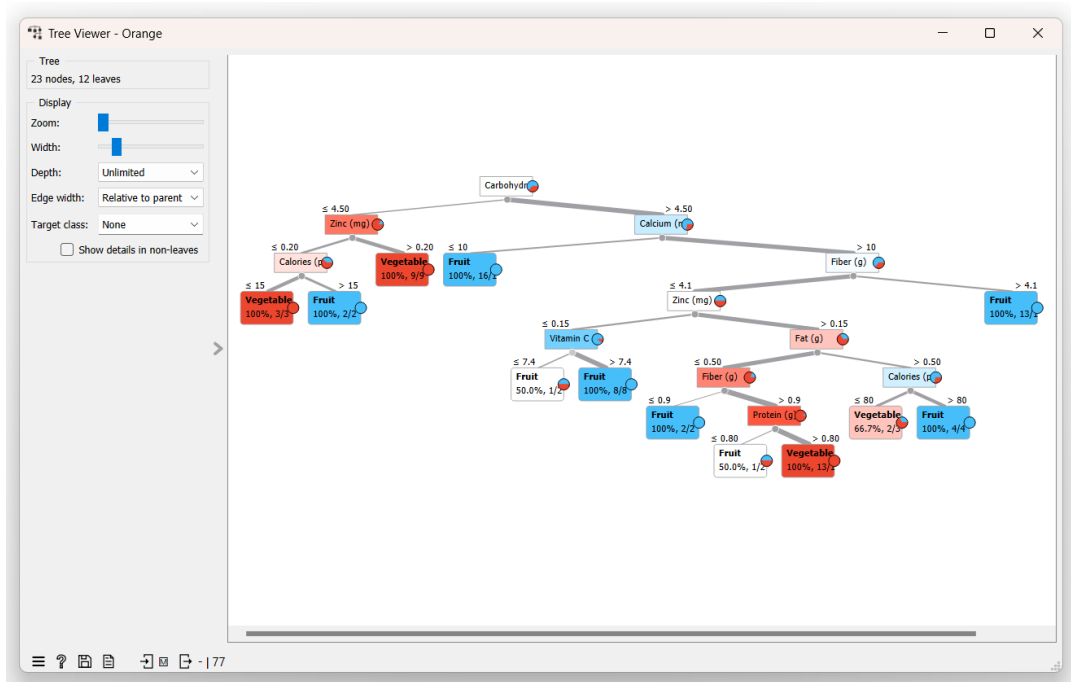
5. Selanjutnya kita perlu membangun model dari set pelatihan menggunakan widget pohon (Tree) klasifikasi.



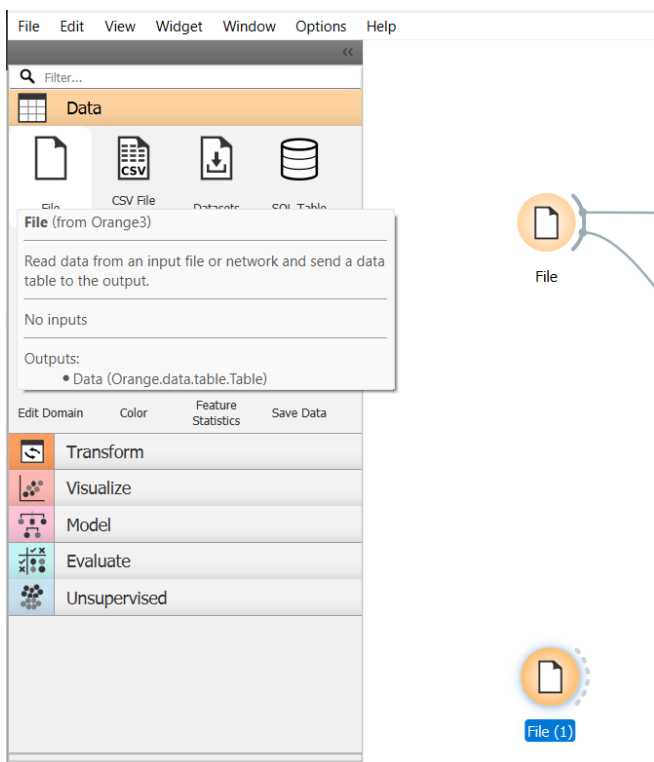
6. Apabila ingin menampilkan sebuah pohon (tree), klik pada widget Tree untuk membuka pilihan dan pilih opsi "Tree Viewer" dari daftar opsi yang tersedia.



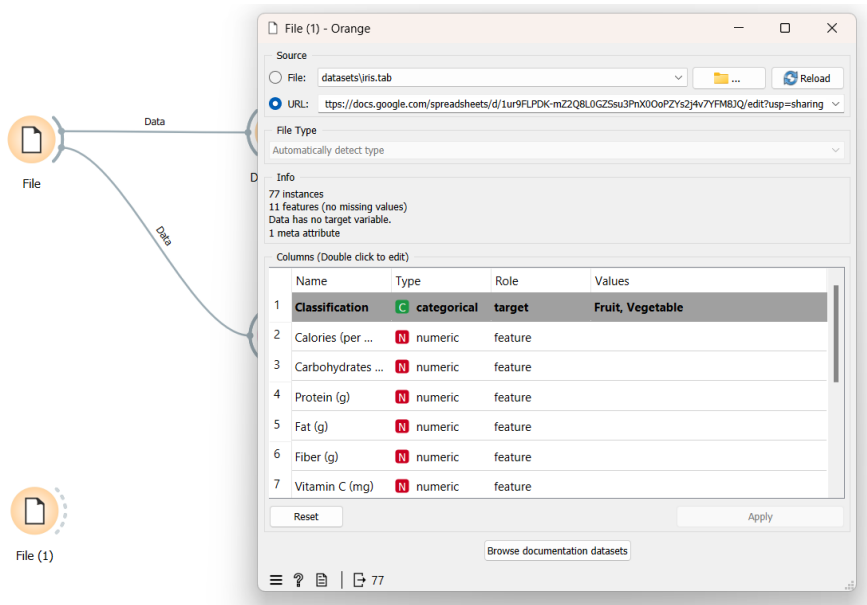
7. Setelah memilih Tree Viewer, maka akan muncul pohon (tree) yang menampilkan struktur data dengan node-node yang dapat dijelajahi.



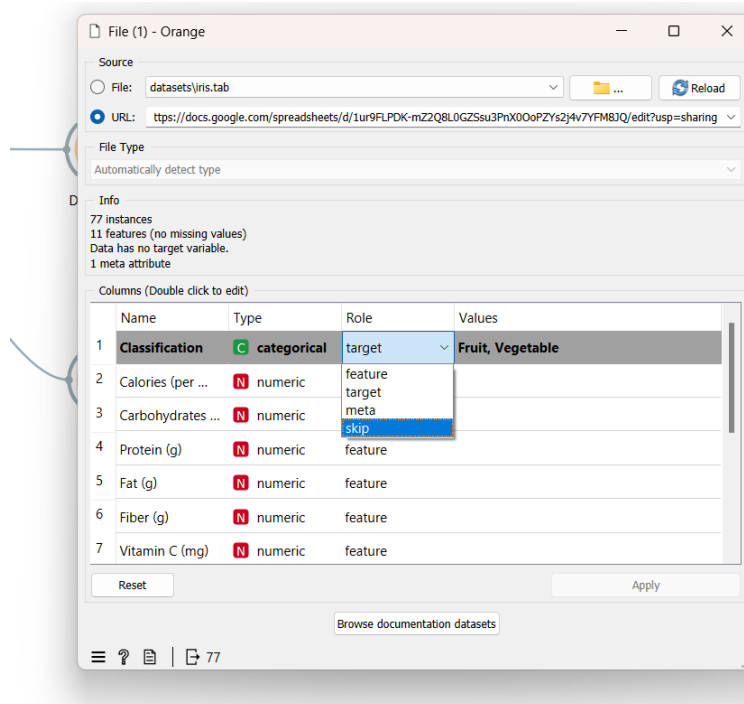
8. Dengan menggunakan dataset yang sama, pilih kembali widget File. File akan muncul di kanvas. Klik 2 kali untuk memuat file data yang akan digunakan.







9. Namun, pada bagian ini kolom Classification tidak digunakan sehingga dibuat seakan data tidak diketahui labelnya. Untuk baris Classification pada kolom Role, ubah keterangan target menjadi skip dengan cara klik 2 kali dan tahan pada tulisan target, sembari tetap menahan kemudian gulir ke bawah pada tulisan skip. Setelah berhasil klik button Apply.



10. Terlihat bahwa kolom classification tidak ada dalam tampilan tabel dataset.

**Data Table (1) - Orange**

Info  
77 instances (no missing data)  
10 features  
No target variable.  
1 meta attribute

Variables  
☒ Show variable labels (if present)  
☐ Visualize numeric values  
☒ Color by instance classes

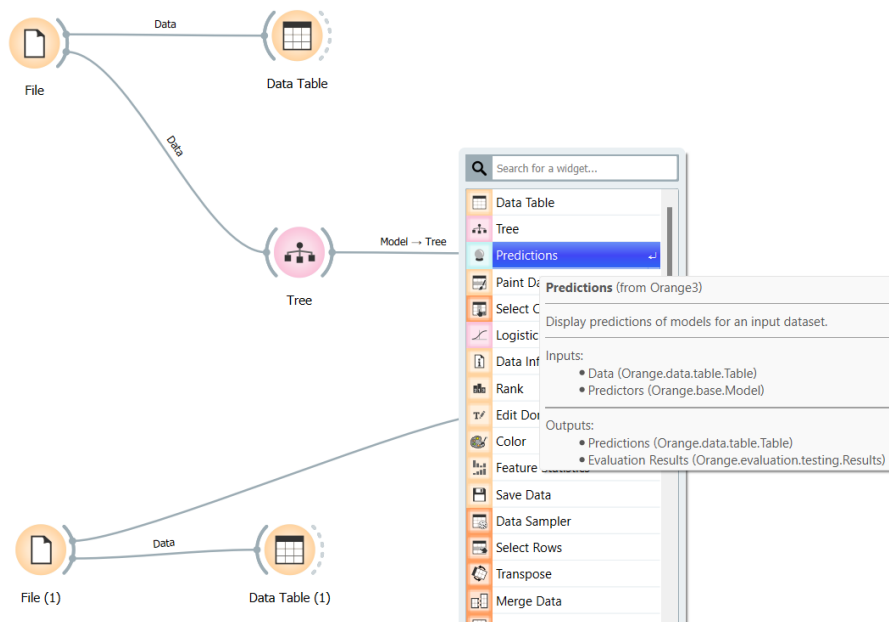
Selection  
☒ Select full rows

Restore Original Order  
☒ Send Automatically

	Name	Calories (per 100g)	Carbohydrates (g)	Protein (g)	Fat (g)	Fiber (g)
1	Amaranth	23	4.00	2.10	0.30	
2	Apple	52	14.00	0.30	0.20	
3	Apricot	48	11.00	1.40	0.40	
4	Asparagus	20	3.70	2.20	0.20	
5	Avocado	160	8.50	2.00	14.70	
6	Banana	89	23.00	1.30	0.30	
7	Beans	31	7.00	1.80	0.20	
8	Beetroot	43	9.60	1.60	0.20	
9	Bell Pepper	20	4.60	0.90	0.20	
10	Bitter Gourd	19	4.30	0.80	0.20	
11	Black Berry	43	9.60	1.40	0.40	
12	Black Currant	63	15.00	0.90	0.40	
13	Blueberry	57	14.50	0.70	0.30	
14	Bottle Gourd	15	3.40	0.60	0.10	
15	Brinjal	25	5.90	1.00	0.20	
16	Broccoli	55	11.20	3.70	0.60	
17	Cabbage	25	6.00	1.30	0.10	
18	Capsicum	20	4.60	0.90	0.20	
19	Carrot	41	10.00	0.90	0.20	
20	Cashew	553	30.19	18.22	43.85	
21	Cauliflower	25	5.00	1.90	0.30	

File (1) → Data → Data Table (1)

11. Kemudian kita perlu menghubungkannya widget File ke widget Prediksi untuk melihat hasil prediksinya.



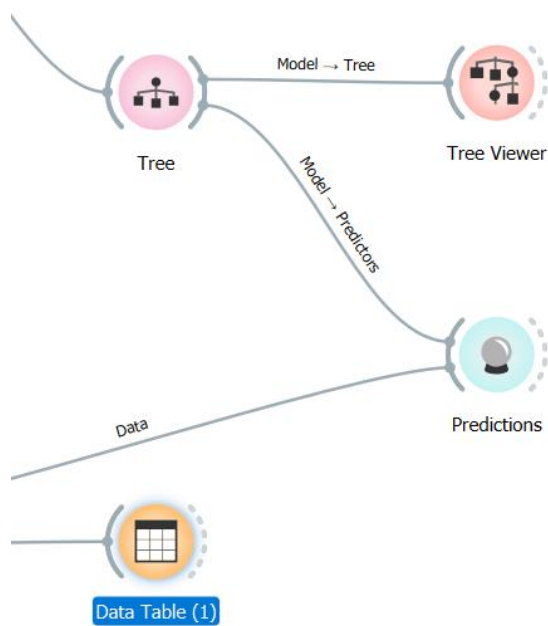
12. Pada tampilan Prediction belum terdapat hasil prediksi karena tidak ada label target yang menjadi acuan, diperlukan widget model klasifikasi.

Predictions - Orange

Restore Original Order

Name	Calories (per 100g)	Carbohydrates (g)	Protein (g)	Fat (g)	Fiber (g)
Amaranth	23	4.00	2.10	0.30	2.0
Apple	52	14.00	0.30	0.20	2.4
Apricot	48	11.00	1.40	0.40	2.0
Asparagus	20	3.70	2.20	0.20	2.1
Avocado	160	8.50	2.00	14.70	6.7
Banana	89	23.00	1.30	0.30	2.6
Beans	31	7.00	1.80	0.20	2.7
Beetroot	43	9.60	1.60	0.20	2.8
Bell Pepper	20	4.60	0.90	0.20	1.7
Bitter Gourd	19	4.30	0.80	0.20	2.6
Black Berry	43	9.60	1.40	0.40	5.3
Black Currant	63	15.00	0.90	0.40	5.8
Blueberry	57	14.50	0.70	0.30	2.4
Bottle Gourd	15	3.40	0.60	0.10	0.6
Brinjal	25	5.90	1.00	0.20	3.0
Broccoli	55	11.20	3.70	0.60	3.3
Cabbage	25	6.00	1.30	0.10	2.5
Capsicum	20	4.60	0.90	0.20	1.7
Carrot	41	10.00	0.90	0.20	2.8
Cashew	553	30.19	18.22	43.85	3.3

13. Sebelumnya sudah terdapat widget model klasifikasi (tree), maka dapat menghubungkan pohon (Tree) klasifikasi ke Prediksi.



14. Setelah diberikan data label target (klasifikasi), akan tampak seperti pada gambar di bawah ini

Predictions - Orange

Show probabilities for: (None) Restore Original Order

	Tree	Name	Calories (per 100g)	Carbohydrates (g)	Protein (g)	Fat (g)	Fiber (g)	Vitamin C (mg)	Zinc (mg)
1	Vegetable	Amaranth	23	4.00	2.10	0.30	2.0	70.0	0.90
2	Fruit	Apple	52	14.00	0.30	0.20	2.4	0.5	0.04
3	Vegetable	Apricot	48	11.00	1.40	0.40	2.0	3.9	0.20
4	Vegetable	Asparagus	20	3.70	2.20	0.20	2.1	5.6	0.50
5	Fruit	Avocado	160	8.50	2.00	14.70	6.7	8.8	1.30
6	Fruit	Banana	89	23.00	1.30	0.30	2.6	8.7	0.20
7	Vegetable	Beans	31	7.00	1.80	0.20	2.7	16.3	0.20
8	Vegetable	Beetroot	43	9.60	1.60	0.20	2.8	1.6	0.40
9	Fruit	Bell Pepper	20	4.60	0.90	0.20	1.7	128.0	0.20
10	Vegetable	Bitter Gourd	19	4.30	0.80	0.20	2.6	19.8	0.80
11	Fruit	Black Berry	43	9.60	1.40	0.40	5.3	21.0	0.53
12	Fruit	Black Currant	63	15.00	0.90	0.40	5.8	181.0	0.27
13	Fruit	Blueberry	57	14.50	0.70	0.30	2.4	9.7	0.16
14	Vegetable	Bottle Gourd	15	3.40	0.60	0.10	0.6	10.0	0.20
15	Fruit	Brinjal	25	5.90	1.00	0.20	3.0	1.0	0.20
16	Vegetable	Broccoli	55	11.20	3.70	0.60	3.3	89.2	0.40
17	Vegetable	Cabbage	25	6.00	1.30	0.10	2.5	36.6	0.20
18	Fruit	Capsicum	20	4.60	0.90	0.20	1.7	128.0	0.20
19	Vegetable	Carrot	41	10.00	0.90	0.20	2.8	3.6	0.20
20	Fruit	Cashew	553	30.19	18.22	43.85	3.3	0.5	5.78

77 | 77 | -  
Data Table (1)

15. Kita dapat memberikan informasi lebih detail pada label Tree nya dengan cara mengubah pilihan dropdown “None” kemudian kita pilih opsi “Classes known to the model”.

Predictions - Orange

Show probabilities for: (None) Classes in data Classes known to the model Classes in data and model

	Tree	Name	Calories (per 100g)	Carbohydrates (g)	Protein (g)
1	Vegetable	Am		4.00	2.10
2	Fruit	Apple	52	14.00	0.30
3	Vegetable	Apricot	48	11.00	1.40
4	Vegetable	Asparagus	20	3.70	2.20
5	Fruit	Avocado	160	8.50	2.00
6	Fruit	Banana	89	23.00	1.30
7	Vegetable	Beans	31	7.00	1.80
8	Vegetable	Beetroot	43	9.60	1.60
9	Fruit	Bell Pepper	20	4.60	0.90

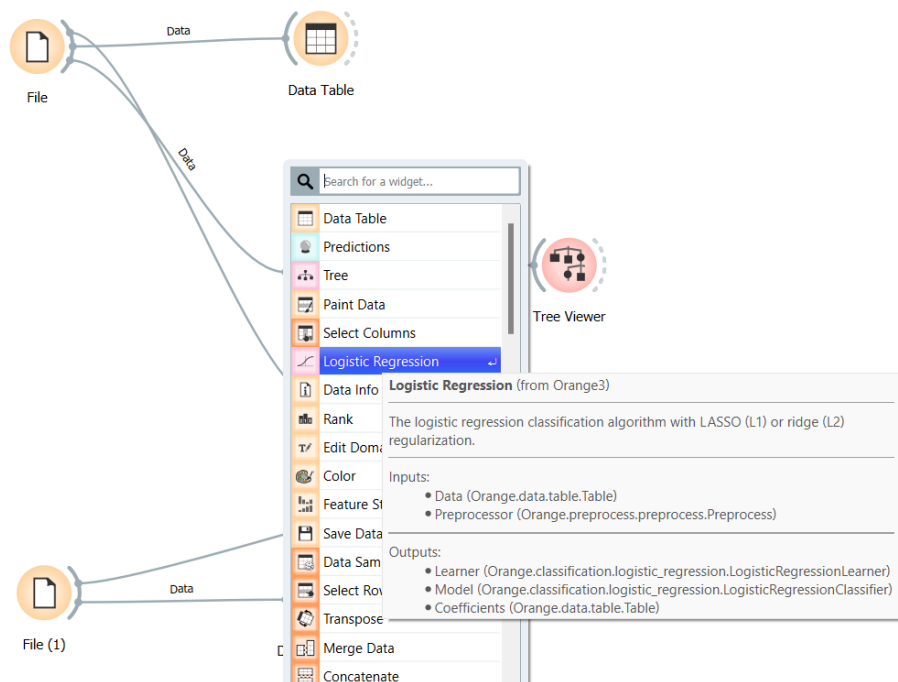
Maka kategori di Tree nya akan berubah menjadi representasi numerik, contohnya “1.00-0.00”, di mana “1.00” mewakili “tree” dan “0.00” mewakili “fruit”.

Predictions - Orange

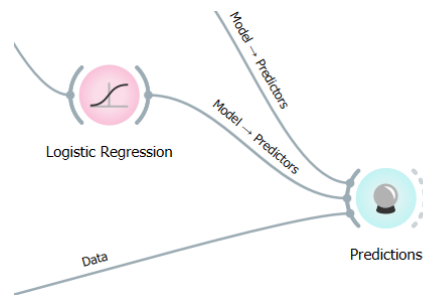
Show probabilities for: Classes known to the model

	Tree	Name
1	0.00 : 1.00 → Vegetable	Amaranth
2	1.00 : 0.00 → Fruit	Apple
3	0.00 : 1.00 → Vegetable	Apricot
4	0.00 : 1.00 → Vegetable	Asparagus
5	1.00 : 0.00 → Fruit	Avocado
6	1.00 : 0.00 → Fruit	Banana
7	0.00 : 1.00 → Vegetable	Beans
8	0.00 : 1.00 → Vegetable	Beetroot
9	1.00 : 0.00 → Fruit	Bell Pepper
10	0.00 : 1.00 → Vegetable	Bitter Gourd
11	1.00 : 0.00 → Fruit	Black Berry
12	1.00 : 0.00 → Fruit	Black Currant
13	1.00 : 0.00 → Fruit	Blueberry
14	0.00 : 1.00 → Vegetable	Bottle Gourd
15	1.00 : 0.00 → Fruit	Brinjal
16	0.33 : 0.67 → Vegetable	Broccoli
17	0.00 : 1.00 → Vegetable	Cabbage

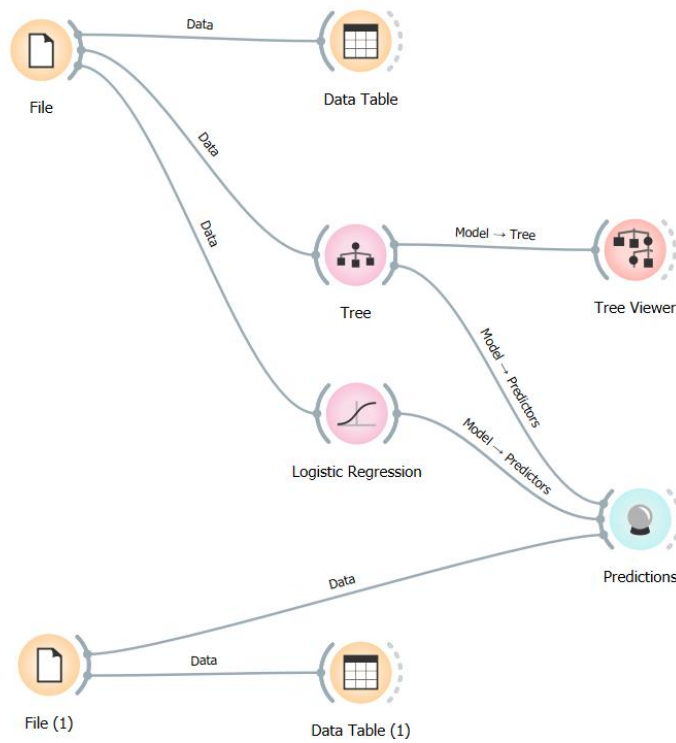
16. Lakukan prediksi kembali untuk memperoleh perbandingan hasil prediksi dengan menggunakan model klasifikasi lainnya, sebagai contoh yaitu Logistic Regression.



17. Hubungkan widget Logistic Regression ke widget Prediction untuk mendapatkan hasil prediksinya



Berikut adalah tampilan lengkap alur klasifikasi - prediksi



18. Perbandingan nilai prediksi antara tree dengan Logistic Regression

Predictions - Orange

Show probabilities for: Classes known to the model Restore Original Order

	Tree	Logistic Regression	Name	Calories (per 100g)	Carbohydrates (g)	Protein (g)	Fat (g)	Fiber (g)
1	0.00 : 1.00 → Vegetable	0.00 : 1.00 → Vegetable	Amaranth	23	4.00	2.10	0.30	2.0
2	1.00 : 0.00 → Fruit	0.98 : 0.02 → Fruit	Apple	52	14.00	0.30	0.20	2.4
3	0.00 : 1.00 → Vegetable	0.63 : 0.37 → Fruit	Apricot	48	11.00	1.40	0.40	2.0
4	0.00 : 1.00 → Vegetable	0.31 : 0.69 → Vegetable	Asparagus	20	3.70	2.20	0.20	2.1
5	1.00 : 0.00 → Fruit	1.00 : 0.00 → Fruit	Avocado	160	8.50	2.00	14.70	6.7
6	1.00 : 0.00 → Fruit	0.94 : 0.06 → Fruit	Banana	89	23.00	1.30	0.30	2.6
7	0.00 : 1.00 → Vegetable	0.30 : 0.70 → Vegetable	Beans	31	7.00	1.80	0.20	2.7
8	0.00 : 1.00 → Vegetable	0.40 : 0.60 → Vegetable	Beetroot	43	9.60	1.60	0.20	2.8
9	1.00 : 0.00 → Fruit	0.55 : 0.45 → Fruit	Bell Pepper	20	4.60	0.90	0.20	1.7
10	0.00 : 1.00 → Vegetable	0.09 : 0.91 → Vegetable	Bitter Gourd	19	4.30	0.80	0.20	2.6
11	1.00 : 0.00 → Fruit	0.90 : 0.10 → Fruit	Black Berry	43	9.60	1.40	0.40	5.3
12	1.00 : 0.00 → Fruit	0.89 : 0.11 → Fruit	Black Currant	63	15.00	0.90	0.40	5.8
13	1.00 : 0.00 → Fruit	0.98 : 0.02 → Fruit	Blueberry	57	14.50	0.70	0.30	2.4
14	0.00 : 1.00 → Vegetable	0.42 : 0.58 → Vegetable	Bottle Gourd	15	3.40	0.60	0.10	0.6
15	1.00 : 0.00 → Fruit	0.69 : 0.31 → Fruit	Brinjal	25	5.90	1.00	0.20	3.0
16	0.33 : 0.67 → Vegetable	0.22 : 0.78 → Vegetable	Broccoli	55	11.20	3.70	0.60	3.3
17	0.00 : 1.00 → Vegetable	0.35 : 0.65 → Vegetable	Cabbage	25	6.00	1.30	0.10	2.5
18	1.00 : 0.00 → Fruit	0.55 : 0.45 → Fruit	Capsicum	20	4.60	0.90	0.20	1.7
19	0.00 : 1.00 → Vegetable	0.27 : 0.73 → Vegetable	Carrot	41	10.00	0.90	0.20	2.8
20	1.00 : 0.00 → Fruit	1.00 : 0.00 → Fruit	Cashew	553	30.19	18.22	43.85	3.3

77 | - Data Table (1)

19. Adapun, terdapat hasil yang berbeda dalam memprediksi Classification, seperti yang terlihat pada gambar di bawah ini.

Predictions - Orange

Show probabilities for: Classes known to the model Restore Original Order

	Tree	Logistic Regression	Name	Calories (per 100g)	Carbohydrates (g)	Protein (g)	Fat (g)	Fiber (g)
21	0.00 : 1.00 → Vegetable	0.21 : 0.79 → Vegetable	Cauliflower	25	5.00	1.90	0.30	2.0
22	1.00 : 0.00 → Fruit	0.99 : 0.01 → Fruit	Chikoo	83	19.90	0.90	0.40	5.3
23	0.00 : 1.00 → Vegetable	0.79 : 0.21 → Fruit	Chilli	40	9.00	1.90	0.20	1.6
24	1.00 : 0.00 → Fruit	1.00 : 0.00 → Fruit	Coconut	354	15.20	3.30	33.50	9.0
25	1.00 : 0.00 → Fruit	0.96 : 0.04 → Fruit	Corn	96	21.82	3.22	1.18	2.4
26	1.00 : 0.00 → Fruit	0.99 : 0.01 → Fruit	Cranberry	46	12.20	0.40	0.10	4.6
27	1.00 : 0.00 → Fruit	0.33 : 0.67 → Vegetable	Cucumber	16	3.63	0.65	0.11	0.5
28	1.00 : 0.00 → Fruit	0.84 : 0.16 → Fruit	Custard Apple	101	23.60	2.06	0.60	3.3
29	1.00 : 0.00 → Fruit	1.00 : 0.00 → Fruit	Dates	277	74.97	2.45	0.39	6.7
30	1.00 : 0.00 → Fruit	0.89 : 0.11 → Fruit	Dragon Fruit	60	11.00	1.80	0.60	3.0
31	1.00 : 0.00 → Fruit	1.00 : 0.00 → Fruit	Elderberry	73	18.40	0.70	0.50	7.0
32	0.50 : 0.50 → Fruit	0.83 : 0.17 → Fruit	Fig	74	19.20	0.80	0.30	2.9
33	0.00 : 1.00 → Vegetable	0.00 : 1.00 → Vegetable	Garlic	149	33.10	6.36	0.50	2.2
34	0.33 : 0.67 → Vegetable	0.45 : 0.55 → Vegetable	Ginger	80	17.80	1.80	0.80	2.0
35	1.00 : 0.00 → Fruit	0.94 : 0.06 → Fruit	Gooseberry	44	10.20	0.90	0.60	5.4
36	1.00 : 0.00 → Fruit	0.91 : 0.09 → Fruit	Grapes	69	18.00	0.70	0.20	0.9
37	1.00 : 0.00 → Fruit	0.97 : 0.03 → Fruit	Guava	68	14.00	2.60	1.00	5.4
38	1.00 : 0.00 → Fruit	0.49 : 0.51 → Vegetable	Jackfruit	95	23.50	2.30	0.60	2.5
39	1.00 : 0.00 → Fruit	0.75 : 0.25 → Fruit	Kiwi	61	14.70	1.10	0.50	3.0
40	1.00 : 0.00 → Fruit	0.93 : 0.07 → Fruit	Lemon	29	9.30	1.10	0.30	2.8

77 | - Data Table (1)