

Nama: Rizki Aprilia Rahman

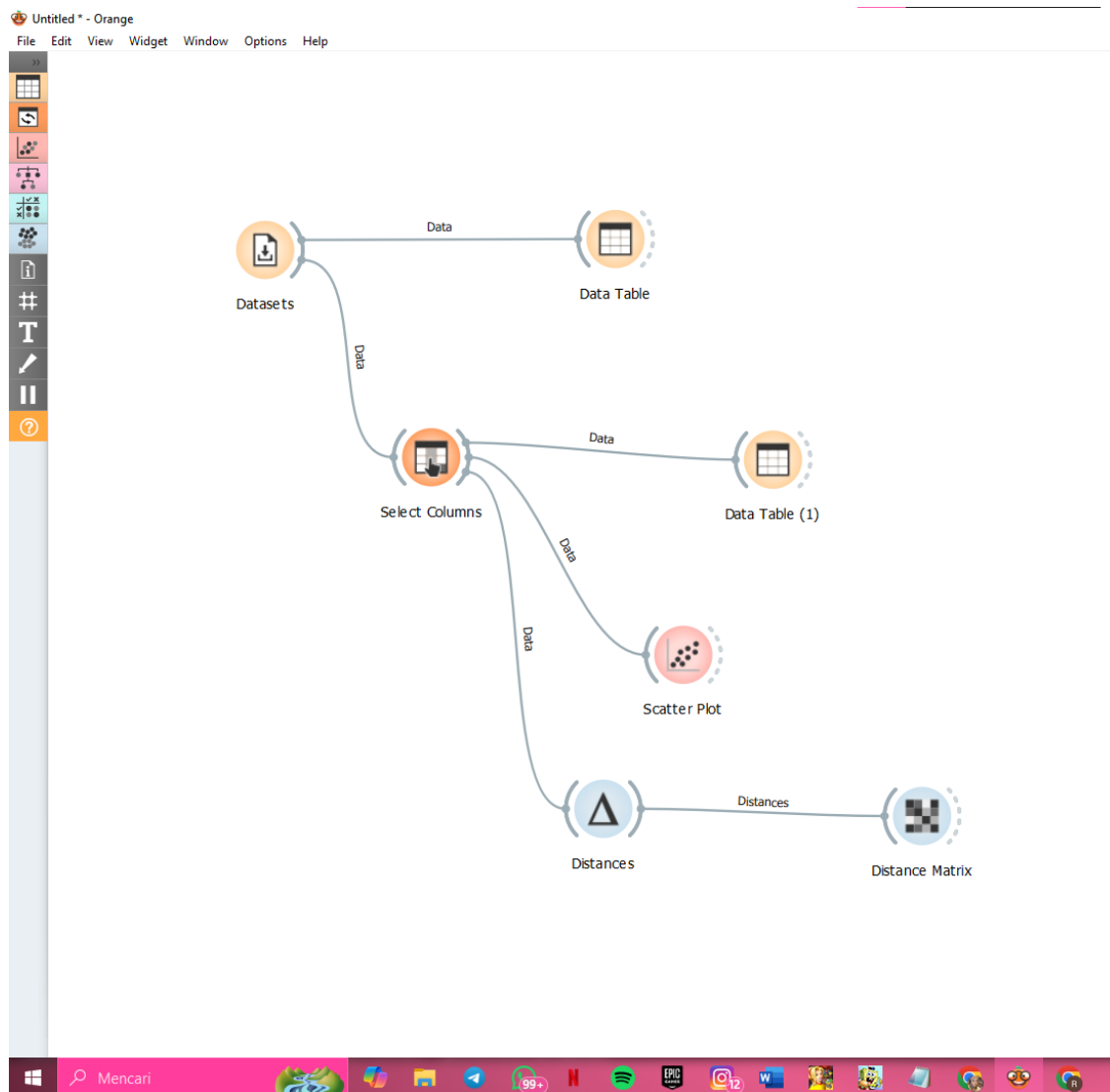
NIM: 1103213007

Kelas: TK-45-G04

TUGAS 6 PEMBELAJARAN MESIN

ORANGE DATA MINING COURSE GRADES DATASET

- Euclidean Distance in 2D



Datasets - Orange

cour Show data sets in English

Title	Size	Instances	Variables	Target	Tags
Course Grades	9.2 KB	16	8	none	synthetic, education

Data Table - Orange

Info
16 instances (no missing data)
7 features
No target variable.
1 meta attribute

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

Selection
☒ Select full rows

Student	English	French	History	Algebra	Biology	Physics	Physical
1 Ana	22	30	32	21	37	46	99
2 Bill	91	95	65	89	39	11	29
3 Cynthia	51	89	21	100	70	100	27
4 Demi	9	15	18	61	100	90	8
5 Eve	93	99	39	12	47	17	63
6 Fred	49	17	17	92	70	98	73
7 George	91	99	97	49	96	81	69
8 Henry	12	30	32	34	12	33	96
9 Ian	91	80	20	82	93	87	22
10 Jena	39	18	19	99	97	77	23
11 Katherine	20	50	10	71	99	78	12
12 Lea	90	100	45	45	20	15	100
13 Maya	100	98	97	32	72	22	37
14 Nash	14	4	15	61	42	51	39
15 Olga	9	22	8	100	11	92	29
16 Phill	85	90	100	45	38	92	21

Select Columns - Orange

Ignored (5)
Filter

- French
- History
- Biology
- Physics
- Physical

Features (2)
Filter

- English
- Algebra

Target

Metas (1)
Student

Reset ☐ Ignore new variables by default ☒ Send Automatically

16 | 16 | 2

Data Table (1) - Orange

Info

16 instances (no missing data)
2 features
No target variable.
1 meta attribute

Variables

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

Selection

- ☒ Select full rows

	Student	English	Algebra
1	Ana	22	21
2	Bill	91	89
3	Cynthia	51	100
4	Demi	9	61
5	Eve	93	12
6	Fred	49	92
7	George	91	49
8	Henry	12	34
9	Ian	91	82
10	Jena	39	99
11	Katherine	20	71
12	Lea	90	45
13	Maya	100	32
14	Nash	14	61
15	Olga	9	100
16	Phill	85	45

Scatter Plot - Orange

Axes

Axis x: English

Axis y: Algebra

Find Informative Projections

Attributes

Color: (Same color)

Shape: (Same shape)

Size: (Same size)

Label: Student

☐ Label only selection and subset

Symbol size:

Opacity:

Jittering:

☐ Jitter numeric values

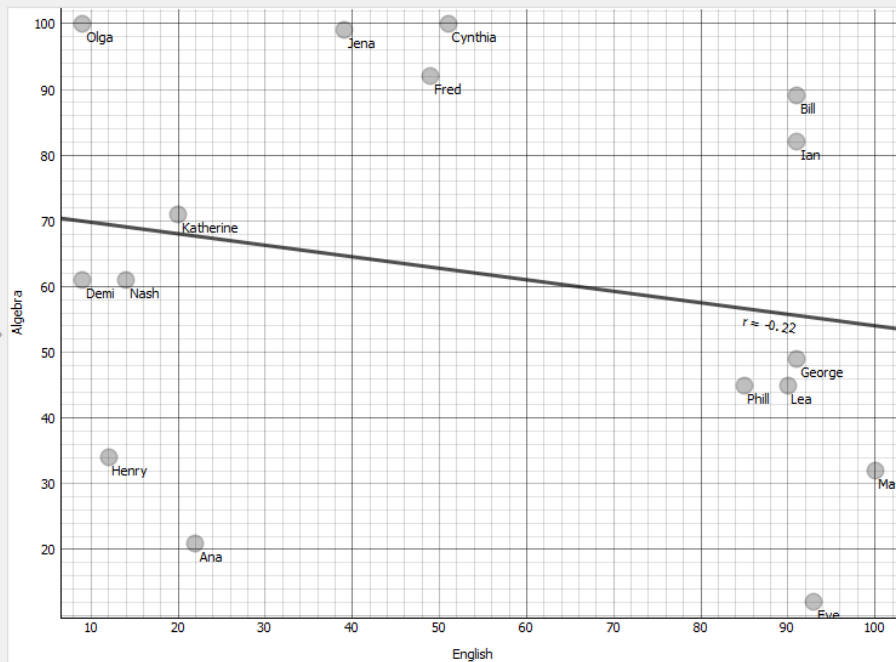
☐ Show color regions

☒ Show legend

Zoom/Select



☒ Send Automatically



Distances - Orange



Compare

☒ Rows

☐ Columns

Distance Metric

☐ Euclidean (normalized)

☐ Cosine

☒ Euclidean

☐ Pearson

☐ Manhattan (normalized)

☐ Pearson (absolute)

☐ Manhattan

☐ Spearman

☐ Mahalanobis

☐ Spearman (absolute)

☐ Hamming

☐ Jaccard



Apply Automatically



Distance Matrix - Orange



	Ana	Bill	Cynthia	Demi	Eve	Fred	George	Henry	Ian	Jena	Katherine	Lea	Maya	Nash	Olga
Ana		96.876	84.155	42.059	71.568	75.961	74.465	16.401	92.098	79.831	50.040	72.111	78.772	40.792	80.062
Bill	96.876		41.485	86.649	77.026	42.107	40.000	96.260	7.000	52.953	73.246	44.011	57.706	81.933	82.735
Cynthia	84.155	41.485		57.315	97.509	8.246	64.815	76.662	43.863	12.042	42.450	67.424	83.815	53.759	42.000
Demi	42.059	86.649	57.315		97.247	50.606	82.873	27.166	84.646	48.415	14.866	82.565	95.509	5.000	39.000
Eve	71.568	77.026	97.509	97.247		91.302	37.054	83.934	70.029	102.396	93.862	33.136	21.190	92.962	121.655
Fred	75.961	42.107	8.246	50.606	91.302		60.108	68.797	43.174	12.207	35.805	62.370	78.746	46.755	40.792
George	74.465	40.000	64.815	82.873	37.054	60.108		80.411	33.000	72.139	74.330	4.123	19.235	77.929	96.566
Henry	16.401	96.260	76.662	27.166	83.934	68.797	80.411		92.439	70.385	37.855	78.772	88.023	27.074	66.068
Ian	92.098	7.000	43.863	84.646	70.029	43.174	33.000	92.439		54.708	71.847	37.014	50.804	79.812	83.952
Jena	79.831	52.953	12.042	48.415	102.396	12.207	72.139	70.385	54.708		33.838	74.277	90.609	45.486	30.017
Katherine	50.040	73.246	42.450	14.866	93.862	35.805	74.330	37.855	71.847	33.838		74.673	89.000	11.662	31.016
Lea	72.111	44.011	67.424	82.565	33.136	62.370	4.123	78.772	37.014	74.277	74.673		16.401	77.666	97.908
Maya	78.772	57.706	83.815	95.509	21.190	78.746	19.235	88.023	50.804	90.609	89.000	16.401		90.758	113.600
Nash	40.792	81.933	53.759	5.000	92.962	46.755	77.929	27.074	79.812	45.486	11.662	77.666	90.758		39.319
Olga	80.062	82.735	42.000	39.000	121.655	40.792	96.566	66.068	83.952	30.017	31.016	97.908	113.600	39.319	
Phill	67.417	44.407	64.661	77.666	33.956	59.203	7.211	73.824	37.483	70.937	70.007	5.000	19.849	72.780	93.814



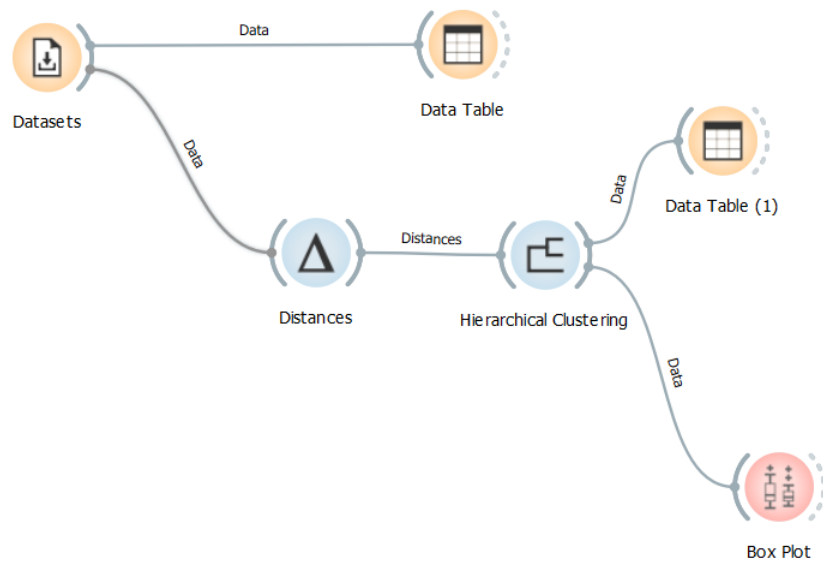
Labels: **S** Student



Send Automatically



- Explaining Clusters



Data Table - Orange

Info											
16 instances (no missing data)											
7 features											
No target variable.											
1 meta attribute											
Variables											
<input checked="" type="checkbox"/> Show variable labels (if present)											
<input checked="" type="checkbox"/> Visualize numeric values											
<input checked="" type="checkbox"/> Color by instance classes											
Selection											
<input checked="" type="checkbox"/> Select full rows											
		Student	English	French	History	Algebra	Biology	Physics	Physical		
1	Ana		22	30	32	21	37	46	99		
2	Bill		91	95	65	89	39	11	29		
3	Cynthia		51	89	21	100	70	100	27		
4	Demi		9	15	18	61	100	90	8		
5	Eve		93	99	39	12	47	17	63		
6	Fred		49	17	17	92	70	98	73		
7	George		91	99	97	49	96	81	69		
8	Henry		12	30	32	34	12	33	96		
9	Ian		91	80	20	82	93	87	22		
10	Jena		39	18	19	99	97	77	23		
11	Katherine		20	50	10	71	99	78	12		
12	Lea		90	100	45	45	20	15	100		
13	Maya		100	98	97	32	72	22	37		
14	Nash		14	4	15	61	42	51	39		
15	Olga		9	22	8	100	11	92	29		
16	Phill		85	90	100	45	38	92	21		

Distances - Orange

×

Compare

☒ Rows
 ☐ Columns

Distance Metric

☒ Euclidean (normalized)
 ☐ Cosine
 ☐ Euclidean
 ☐ Pearson
 ☐ Manhattan (normalized)
 ☐ Pearson (absolute)
 ☐ Manhattan
 ☐ Spearman
 ☐ Mahalanobis
 ☐ Spearman (absolute)
 ☐ Hamming
 ☐ Jaccard

☒ Apply Automatically

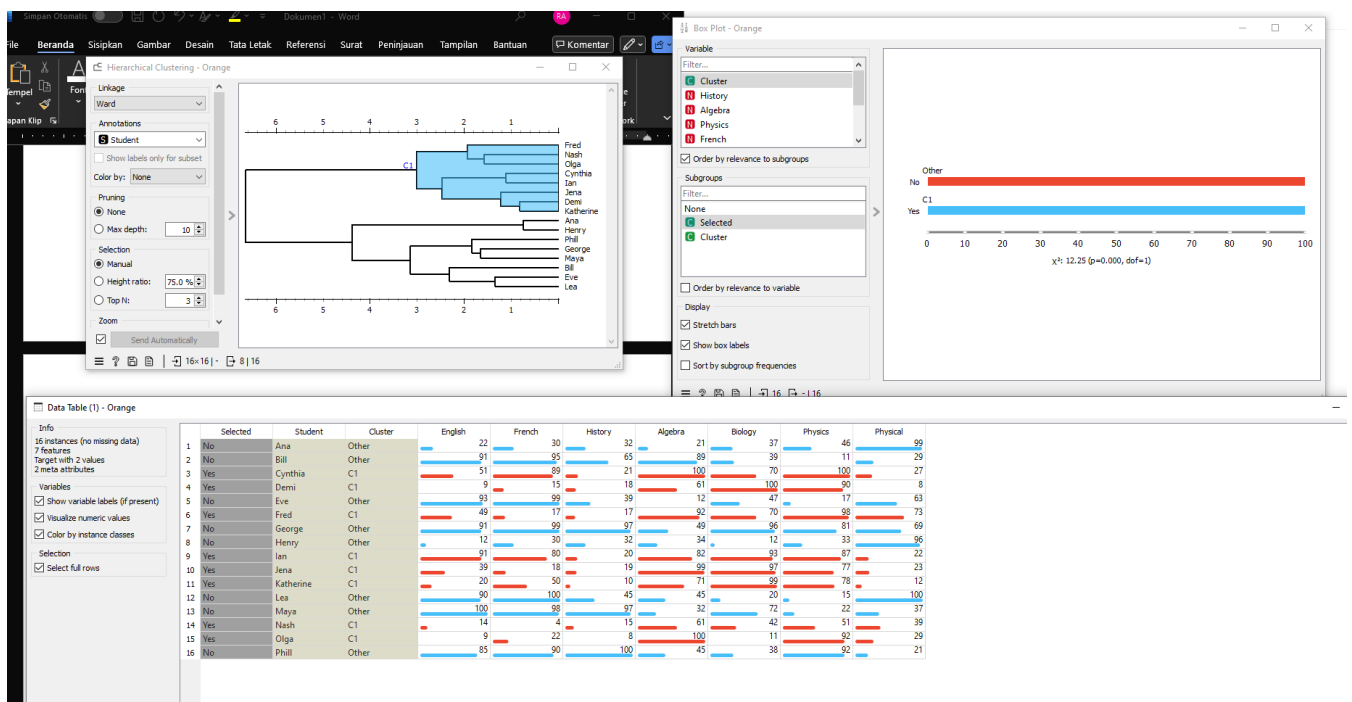
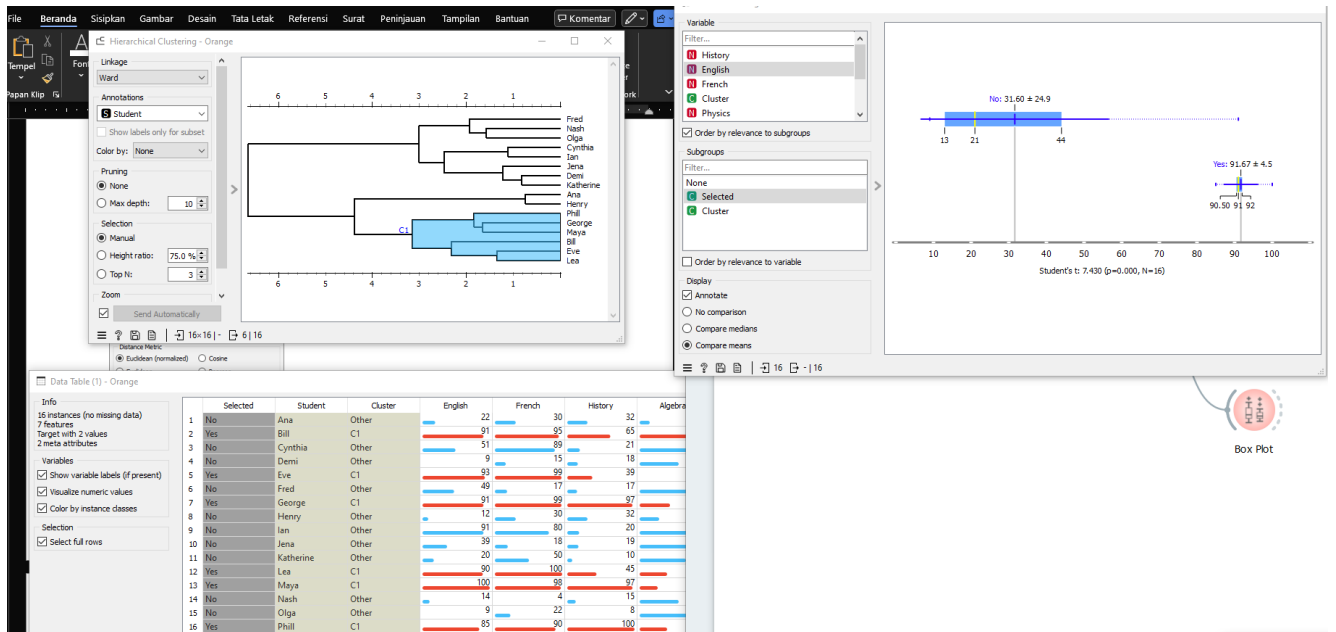
≡

?

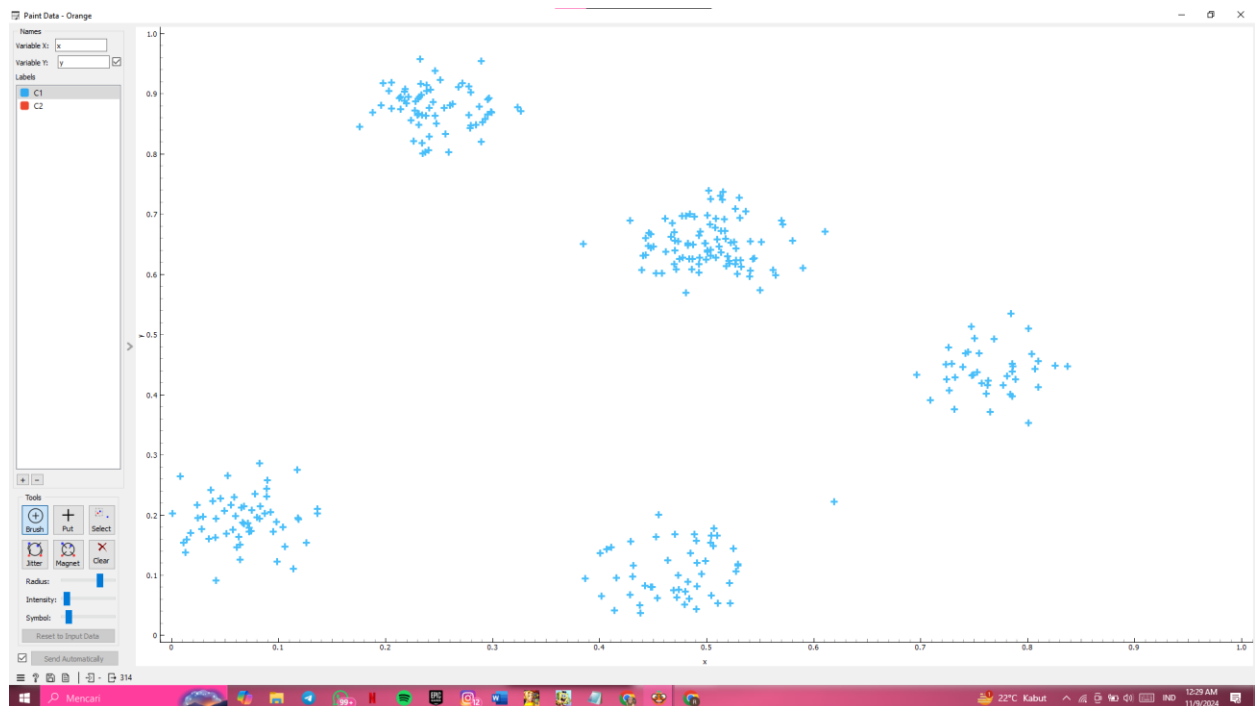
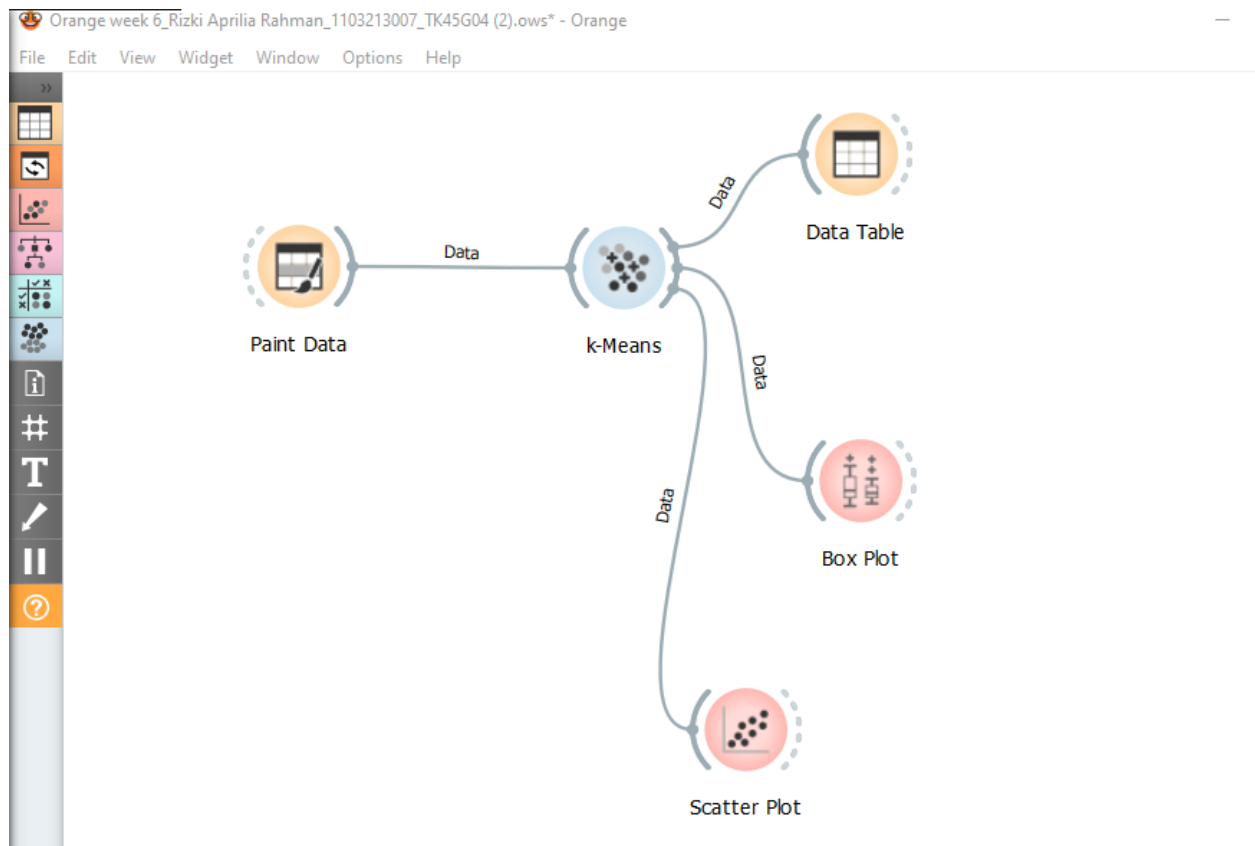
📄

↩️ 16

📏 16×16



- How to choose k for k-Means?



k-Means - Orange

Number of Clusters

☒ Fixed:

☐ From to

Preprocessing

☒ Normalize columns

Initialization

Initialize with KMeans++

Re-runs:

Maximum iterations:

☒ Apply Automatically

314 | 5

Data Table - Orange

Info

314 instances (no missing data)

2 features

No target variable.

2 meta attributes

Variables

☒ Show variable labels (if present)

☒ Visualize numeric values

☒ Color by instance classes

Selection

☒ Select full rows

	Cluster	Silhouette	x	y
1	C2	0.711606	0.295602	0.890909
2	C2	0.711055	0.290363	0.852962
3	C2	0.725949	0.213821	0.873793
4	C2	0.710972	0.296523	0.892272
5	C2	0.728864	0.231346	0.894651
6	C2	0.727655	0.245996	0.863496
7	C2	0.726113	0.217758	0.903241
8	C2	0.710792	0.292844	0.858431
9	C2	0.725395	0.23287	0.916826
10	C2	0.727975	0.229491	0.865045
11	C2	0.718941	0.279384	0.902726
12	C2	0.710618	0.295054	0.865176
13	C2	0.725054	0.247187	0.85035
14	C2	0.722091	0.202735	0.90417
15	C2	0.716169	0.287625	0.878873
16	C2	0.727338	0.217737	0.888626
17	C2	0.71559	0.234248	0.818586
18	C2	0.710157	0.23981	0.806355
19	C2	0.71411	0.232428	0.958036
20	C2	0.728077	0.237737	0.863136
21	C2	0.725929	0.262596	0.883248
22	C2	0.728455	0.227501	0.871733
23	C2	0.716004	0.226002	0.821028
24	C2	0.727643	0.254462	0.876871
25	C2	0.715361	0.279295	0.847027
26	C2	0.725983	0.23801	0.914412
27	C2	0.708811	0.236799	0.803107
28	C2	0.719631	0.27766	0.86452
29	C2	0.719444	0.246187	0.938495
30	C2	0.727648	0.221083	0.89481
31	C2	0.718718	0.197841	0.91806
32	C2	0.724817	0.230904	0.848262
33	C2	0.728654	0.233374	0.897652
34	C2	0.70768	0.234501	0.800652
35	C2	0.718627	0.255602	0.832739
36	C2	0.709244	0.175593	0.845254
37	C2	0.719334	0.271823	0.917805
38	C2	0.729053	0.244415	0.88605
39	C2	0.713363	0.284284	0.848646
40	C2	0.708395	0.299139	0.869055
41	C2	0.725626	0.223675	0.856033
42	C4	0.726502	0.498717	0.651654
43	C4	0.720304	0.527495	0.642538
44	C4	0.717876	0.518521	0.613655
45	C4	0.71271	0.48072	0.697086
46	C4	0.725753	0.509858	0.657485
47	C4	0.715501	0.530801	0.693892
48	C4	0.711478	0.477068	0.696694
49	C4	0.716466	0.471607	0.607845
50	C4	0.717453	0.46975	0.616941

Restore Original Order

☒ Send Automatically

Mencari

