

Nama :Nauval Aviv

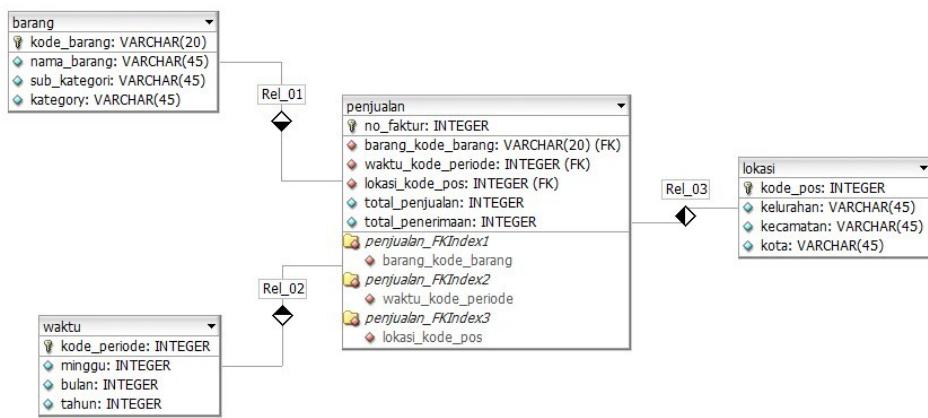
Nim :L200170060

Kelas :C

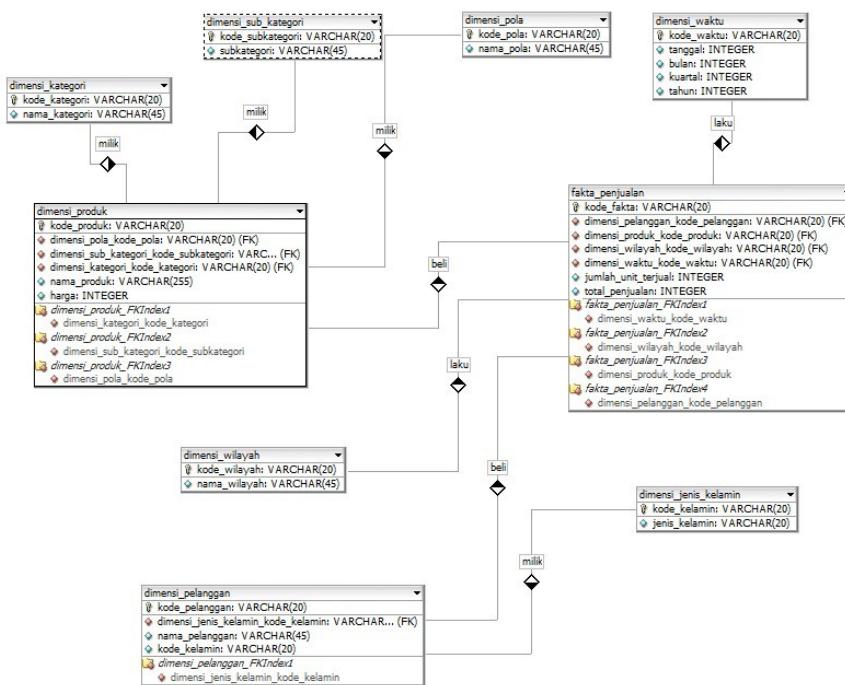
## REKAP MODUL (1-12)

### MODUL 1

Star schema penjualan.xml



Snowflake penjualan.xml



## MODUL 5

### PIVOT TABLE DAN CHART

#### D. Langkah-langkah Praktikum

##### D. 1. Kegiatan 1 : Membuat Pivot Table

1. Membuat data Excel dan menyimpannya dengan nama ‘Fakta\_Penjualan.xls’.

The screenshot shows a Microsoft Excel spreadsheet titled 'Fakta\_Penjualan - Excel'. The data is organized into several columns:

bulan	kuartal	tahun	nama_produk	nama_subkategori	nama_pola	nama_pelanggan	jenis_kelamin	nama_wilayah	jumlah	harga
1	12	4	2011 Jarik Standar Print Sogan	Standar	Print	Bapak Ketut	PRIA	Bali	2	225000
3	1	1	2012 Jarik Standar Tulis Sarimbit	kaos	Cap	Ibu Harini	WANITA	Jawa Timur	14	30000
4	4	2	2012 Rok Batik Print Kombinasi	Jarik	Tulis	Ibu Harini	WANITA	Jawa Timur	4	40000
5	4	2	2011 Batik Standar Cap Tumpal	Hem	Print	Ibu Harini	WANITA	Jawa Timur	3	70000
6	9	3	2012 Celana Standar Cap Warna	Batik	Cap	Bapak Heru	PRIA	Jawa Timur	1	150000
7	5	2	2012 Celana Standar Print Lasem	Hem	Print	Bapak Totok	PRIA	Jawa Timur	3	299000
8	12	4	2011 Kaos Batik Cap Tulis	Bolero	Cap	Ibu Hatamah	WANITA	Jawa Timur	1	255000
9	10	4	2011 Kaos Katun Print Bola	Sarimbit	Print	Ibu Hatamah	WANITA	Jawa Timur	1	150000
10	1	1	2011 Jam Standar Print Lukis	Kaos	Print	Bapak Imron	PRIA	Jawa Barat	1	60000
11	2	1	2012 Sarimbit Standar Print Lukis	Celana	Cap	Ibu Hadi Sukarni	WANITA	Jawa Barat	17	55000
12	3	1	2010 Hem Standar Cap Tumpal	Celana	Print	Ibu Hadi Sukarni	WANITA	Jawa Barat	17	55000
13	3	1	2011 Hem Sutra Print Rama	Bahan	Cap	Ibu Siti Arya	WANITA	Jawa Barat	8	120000
14	12	4	2012 Hem Standar Tulis Madura	Rok	Print	Ibu Siti Arya	WANITA	Jawa Barat	1	225000
15	1	1	2012 Hem katun Print Kelengan	Jam	Print	Ibu Siti Arya	WANITA	Jawa Barat	44	80000
16	9	3	2012 Hem katun Print Kawung	Hem	Cap	Ibu Aini Kasmaji	WANITA	Jawa Tengah	1	100000
17	6	2	2012 Bahan Standar Cap Lasem	Bahan	Tulis	Ibu Niken	WANITA	Jawa Tengah	1	130000
18	8	3	2011 Bahan Lawasan Tulis Tolet	Hem	Tulis	Ibu Atik	WANITA	Jawa Tengah	5	550000
19	4	2	2012 Bahan Belundru Cap Mahkota	Bahan	Cap	Ibu Tyas	WANITA	Jawa Tengah	7	135000
20	6	2	2010 Bahan Standar Cap Garis	Bahan	Cap	Ibu Tyas	WANITA	Jawa Tengah	1	500000
21	11	4	2010 Bolero Standar Cap Sidomukti	Hem	Print	Ibu Tyas	WANITA	Jawa Tengah	5	100000

2. Membuat Pivot Table dengan menggunakan data tersebut.

Screenshot of Microsoft Excel showing the 'Create PivotTable' dialog box. The dialog box is titled 'Create PivotTable' and contains three main sections:

- Choose the data that you want to analyze:**
  - Select a table or range: Sheet1!\$A\$1:\$K\$21
  - Use an external data source: Choose Connection...
  - Use this workbook's Data Model
- Choose where you want the PivotTable report to be placed:**
  - New Worksheet
  - Existing Worksheet: Location: [empty box]
- Choose whether you want to analyze multiple tables:**
 Add this data to the Data Model

The background shows two tables of data. The left table has columns: bulan, kuartal, tahun, nama\_produk. The right table has columns: kelamin, nama\_wilayah, jumlah, harga.

3. Menampilkan data Pivot Table dengan mencoba kombinasi **tahun**, **nama\_subkategori**, dan **jumlah**.

Screenshot of Microsoft Excel showing a PivotTable named 'PivotTable1' on Sheet2. The PivotTable is set up with the following structure:

		2010	2011	2012	Grand Total
Row Labels	Bahan	1	8	8	17
	Batik			1	1
	Bolero		1		1
	Celana	17	17		34
	Hem	5	8	4	17
	Jam		44		44
	Jarik		4		4
	kaos	1	14	15	
	Rok		1		1
	Sarimbit	1		1	
	Standar	2		2	
	<b>Grand Total</b>	<b>23</b>	<b>21</b>	<b>93</b>	<b>137</b>

The PivotTable Tools ribbon is visible at the top, showing the Analyze tab selected. The PivotTable Fields pane on the right lists fields: bulan, kuartal, tahun, nama\_produk, and nama\_subkategori, with 'tahun' and 'nama\_subkategori' checked.

## D. 2. Kegiatan 2 : Menambahkan Tipe Summary Baru

1. Menambahkan field **jumlah** ke kotak **Value** (karena sudah ada field jumlah sebelumnya, maka field dengan nama yang sama di kotak yang sama akan memiliki tambahan karakter berupa angka di akhiran kata, pada kasus ini **jumlah2** atau **sum of jumlah2**).

The screenshot shows a Microsoft Excel spreadsheet titled "Fakta\_Penjualan - Excel". The PivotTable Tools ribbon tab is selected. The PivotTable Fields pane on the right lists fields: jenis\_kelamin, nama\_wilayah, jumlah (with a checked checkbox), harga, and Pendapatan. The main area shows a PivotTable with data for items like Bahan, Batik, Bolero, Celana, Hem, Jam, Jarik, kaos, Rok, Sarimbit, and Standar across years 2010, 2011, and 2012. The table includes columns for Sum of jumlah, Count of jumlah, and their respective totals.

	2010	2011	2012	Total Sum of jumlah	Total Count of jumlah
Bahan	1	8	1	8	2
Batik				1	1
Bolero		1	1		1
Celana	17	1		17	1
Hem	5	1	2	4	2
Jam				44	1
Jarik				4	1
kaos		1	1	14	1
Rok				1	1
Sarimbit		1	1		1
Standar		2	1		2
<b>Grand Total</b>	<b>23</b>	<b>3</b>	<b>7</b>	<b>93</b>	<b>10</b>
					<b>137</b>
					<b>20</b>

2. Mengubah jumlah nilai penjualan yang terjadi (**sum**) pada **jumlah2**, menjadi jumlah transaksi yang terjadi (**count**).

The screenshot shows a Microsoft Excel spreadsheet titled "Fakta\_Penjualan - Excel". A PivotTable is active, displaying data for years 2010 and 2011 across categories like Bahan, Batik, Bolero, Celana, Hem, Jam, Jarik, kaos, Rok, Sarimbit, and Standar. The PivotTable Fields pane on the right lists fields such as 'jenis\_kelamin', 'nama\_wilayah', 'jumlah' (which is checked), and 'harga'. A 'Value Field Settings' dialog box is open, showing 'Source Name: jumlah' and 'Custom Name: Count of jumlah2'. Under 'Summarize value field by', 'Count' is selected from a dropdown menu. The 'OK' button is visible at the bottom of the dialog.

### D. 3. Kegiatan 3 : Calculated Field dan Calculated Item di Pivot Table

#### a) Calculated Field (Menambah field / kolom baru pada daftar field yang ada )

1. Membuat field baru dengan menggunakan tab **Formulas** -> **Calculated Field**, lalu membuat field baru dengan nama **Pendapatan** dengan formula, **jumlah \* harga**

This screenshot shows the 'Insert Calculated Field' dialog box in the 'PivotTable Tools' ribbon. The 'Name:' field is set to 'Pendapatan' and the 'Formula:' field contains the formula '=jumlah\*harga'. Below the formula, a 'Fields:' list shows available fields: 'nama\_produk', 'nama\_subkategori', 'nama\_pola', 'nama\_wilayah', 'jumlah', and 'harga'. The 'OK' button is visible at the bottom of the dialog. The main Excel window shows a PivotTable with columns for 'Row Labels', '2010', 'Count of jumlah2', '2011', and 'Count of jumlah3'. The PivotTable Fields pane on the right now includes the 'Pendapatan' field in the 'Values' section.

2. Setelah meng-klik tombol ‘OK’ maka akan muncul file baru dengan nama ‘Pendapatan’, cukup centang field tersebut dan data akan otomatis keluar.

Row Labels	2010			2011			Sum
	Sum of jumlah	Count of jumlah2	Sum of Pendapatan	Sum of jumlah	Count of jumlah2	Sum of Pendapatan	
Bahan	1	1	500000	8	1	960000	
Batik			0			0	
Bolero			0	1	1	255000	
Celana	17	1	935000			0	
Hem	5	1	500000	8	2	4960000	
Jam			0			0	
Jarik			0			0	
kaos			0	1	1	60000	
Rok			0			0	
Sarimbit			0	1	1	150000	
Standar			0	2	1	450000	
<b>Grand Total</b>	<b>23</b>	<b>3</b>	<b>15065000</b>	<b>21</b>	<b>7</b>	<b>30030000</b>	

Count of jumlah2	Sum of Pendapatan	Total Sum of jumlah	Total Count of jumlah2	Total Sum of Pend
2	2120000	17	4	1
1	150000	1	1	
	0	1	1	
1	935000	34	2	
2	1596000	17	5	
1	3520000	44	1	
1	160000	4	1	
1	420000	15	2	
1	225000	1	1	
	0	1	1	
	0	2	1	
<b>10</b>	<b>115692000</b>	<b>137</b>	<b>20</b>	<b>45</b>

#### D. 4. Kegiatan 4 : Operasi Roll Up dan Drill Down

- Menghilangkan beberapa field, sehingga tinggal tersisa **tahun** (Columns), **nama\_subkategori** (Rows), dan **Sum of Pendapatan** (Value).

Fakta\_Penjualan - Excel

PivotTable Tools

PivotTable Name: Fakta\_Penjualan Active Field: Sum of Pendapatam

	A	B	C	D	E	F	G	H	I	J
1										
2										
3	Sum of Pendapatam	Column Labels	2010	2011	2012	Grand Total				
4	Row Labels									
5	Bahan	500000	960000	2120000	15045000					
6	Batik	0	0	150000	150000					
7	Bolero	0	255000	0	255000					
8	Celana	935000	0	935000	3740000					
9	Hem	500000	4960000	1596000	19023000					
10	Jam	0	0	3520000	3520000					
11	Jarik	0	0	160000	160000					
12	kaos	0	60000	420000	1350000					
13	Rok	0	0	225000	225000					
14	Sarimbit	0	150000	0	150000					
15	Standar	0	450000	0	450000					
16	Grand Total	15065000	30030000	115692000	456073000					
17										
18										
19										
20										
21										
22										
23										

PivotTable Fields

Choose fields to add to report:

Search

nama\_wilayah

jumlah

harga

Pendapatam

More Tables...

Drag fields between areas below:

Filters Columns tahun

Rows nama\_subkata... Values Sum of Penda...

Defer Layout Update Update

- Menambahkan field pada kotak **Column (kuartal dan bulan)** dan **Rows (nama\_produk dan nama\_pelanggan)**, untuk menambah field-field yang bisa diperinci dan diringkas sesuai urutan kategori data yang lebih spesifik.

## Roll Up

Fakta\_Penjualan - Excel

PivotTable Tools

PivotTable Name: Fakta\_Penjualan Active Field: tahun

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1																	
2																	
3	Sum of Pendapatam	Column Labels	2010	2011	2012	Grand Total											
4	Row Labels																
5	Bahan	500000	960000	2120000	15045000												
6	Batik	0	0	150000	150000												
7	Bolero	0	255000	0	255000												
8	Celana	935000	0	935000	3740000												
9	Hem	500000	4960000	1596000	19023000												
10	Jam	0	0	3520000	3520000												
11	Jarik	0	0	160000	160000												
12	kaos	0	60000	420000	1350000												
13	Rok	0	0	225000	225000												
14	Sarimbit	0	150000	0	150000												
15	Standar	0	450000	0	450000												
16	Grand Total	15065000	30030000	115692000	456073000												
17																	
18																	
19																	
20																	
21																	
22																	
23																	

PivotTable Fields

Choose fields to add to report:

Search

bulan

kuartal

tahun

nama\_produk

nama\_subkategori

Drag fields between areas below:

Filters Columns tahun bulan kuartal

Rows nama\_subkata... Values Sum of Penda...

Defer Layout Update Update

## Drop Down

Screenshot of Microsoft Excel showing a PivotTable named "Fakta\_Penjualan". The PivotTable displays the sum of sales for various products across three years (2010, 2011, 2012) and a Grand Total. The rows are grouped by product category (Bahan, Batik, Celana, Hem, Jam). The columns represent the years and the Grand Total.

		2010	2011	2012	Grand Total
3. Sum of Pendapatan	Column Labels				
4. Row Labels		2010	2011	2012	Grand Total
5. Bahan		500000	960000	2120000	15045000
6.    Bahan Belundru Cap Mahkota		0	0	945000	945000
7.    Ibu Tyas		0	0	945000	945000
8.    Bahan Standar Cap Garis		500000	0	0	500000
9.    Ibu Tyas		500000	0	0	500000
10.    Bahan Standar Cap Lasem		0	0	130000	130000
11.    Ibu Niken		0	0	130000	130000
12.    Hem Sutra Print Rama		0	960000	0	960000
13.    Ibu Siti Arya		0	960000	0	960000
14.    Batik		0	0	150000	150000
15.    Celana Standar Cap Warna		0	0	150000	150000
16.    Bapak Heru		0	0	150000	150000
17.    Bolero		0	255000	0	255000
18.    Kaos Batik Cap Tulis		0	255000	0	255000
19.    Ibu Hatamah		0	255000	0	255000
20.    Celana		935000	0	935000	3740000
21.    Hem Standar Cap Tumpal		935000	0	0	935000
22.    Ibu Hadi Sukarni		935000	0	0	935000
23.    Sarimbit Standar Print Lukis		0	0	935000	935000
24.    Ibu Hadi Sukarni		0	0	935000	935000
25.    Hem		500000	4960000	1596000	19023000

Screenshot of Microsoft Excel showing a PivotTable named "Fakta\_Penjualan". The PivotTable displays the sum of sales for various products across three years (2010, 2011, 2012) and a Grand Total. The rows are grouped by product category (Bahan, Batik, Celana, Hem, Jam). The columns represent the years and the Grand Total.

		2010	2011	2012	Grand Total
26. Bahan Lawasan Tulis Tolet		0	2750000	0	2750000
27. Ibu Atik		0	2750000	0	2750000
28. Batik Standar Cap Tumpal		0	210000	0	210000
29. Ibu Harini		0	210000	0	210000
30. Bolero Standar Cap Sidomukti		500000	0	0	500000
31. Ibu Tyas		500000	0	0	500000
32. Celana Standar Print Lasem		0	0	897000	897000
33. Bapak Totok		0	0	897000	897000
34. Hem katun Print Kawung		0	0	100000	100000
35. Ibu Aini Kasmaji		0	0	100000	100000
36. Jam		0	0	3520000	3520000
37. Hem katun Print Kelengan		0	0	3520000	3520000
38. Ibu Siti Arya		0	0	3520000	3520000
39. Jarik		0	0	160000	160000
40. Rok Batik Print Kombinasi		0	0	160000	160000
41. Ibu Harini		0	0	160000	160000
42. kaos		0	60000	420000	1350000
43. Jam Standar Print Lukis		0	60000	0	60000
44. Bapak Imron		0	60000	0	60000
45. Jarik Standar Tulis Sarimbit		0	0	420000	420000
46. Ibu Harini		0	0	420000	420000
47. Rok		0	0	225000	225000
48. Hem Standar Tulis Madura		0	0	225000	225000

## E. Tugas

1. Dengan menggunakan **PivotTable** pada file **Fakta\_Penjualan.xls** tambahkan 2 buah field, yaitu :
  - a. **PPN** (Pajak Pertambahan Nilai) sebesar 10% dari tiap pendapatan pada Pivot Table.
  - b. **Total Penghasilan** yang dihitung dari pendapatan dikurangi dengan PPN tersebut.

	Column Labels	2010	2011				
Row Labels	Sum of Pendapatan	Sum of PPN (10%)	Sum of Total Penghasilan	Sum of jumlah	Sum of kuartal	Sum of Pendapatan	Sum of PPN (10%)
Bahan	500000	50000	450000	1	2	960000	96000
Batik	0	0	0			0	0
Bolero	0	0	0			255000	25500
Celana	935000	93500	841500	17	1	0	0
Hem	500000	50000	450000	5	4	4960000	496000
Jam	0	0	0			0	0
Jarik	0	0	0			0	0
kaos	0	0	0			60000	6000
Rok	0	0	0			0	0
Sarimbit	0	0	0			150000	15000
Standar	0	0	0			450000	45000
<b>Grand Total</b>	<b>15065000</b>	<b>1506500</b>	<b>13558500</b>	<b>23</b>	<b>7</b>	<b>30030000</b>	<b>3003000</b>

2012								
Sum of Total Penghasilan	Sum of jumlah	Sum of kuartal	Sum of Pendapatan	Sum of PPN (10%)	Sum of Total Penghasilan	Sum of jumlah	Sum of kuartal	
864000	8	1	2120000	212000	1908000	8	4	
0			150000	15000	135000	1	3	
229500	1	4	0	0	0			
0			935000	93500	841500	17	1	
4464000	8	5	1596000	159600	1436400	4	5	
0			3520000	352000	3168000	44	1	
0			160000	16000	144000	4	2	
54000	1	1	420000	42000	378000	14	1	
0			225000	22500	202500	1	4	
135000	1	4	0	0	0			
405000	2	4	0	0	0			
<b>27027000</b>	<b>21</b>	<b>19</b>	<b>115692000</b>	<b>11569200</b>	<b>104122800</b>	<b>93</b>	<b>21</b>	

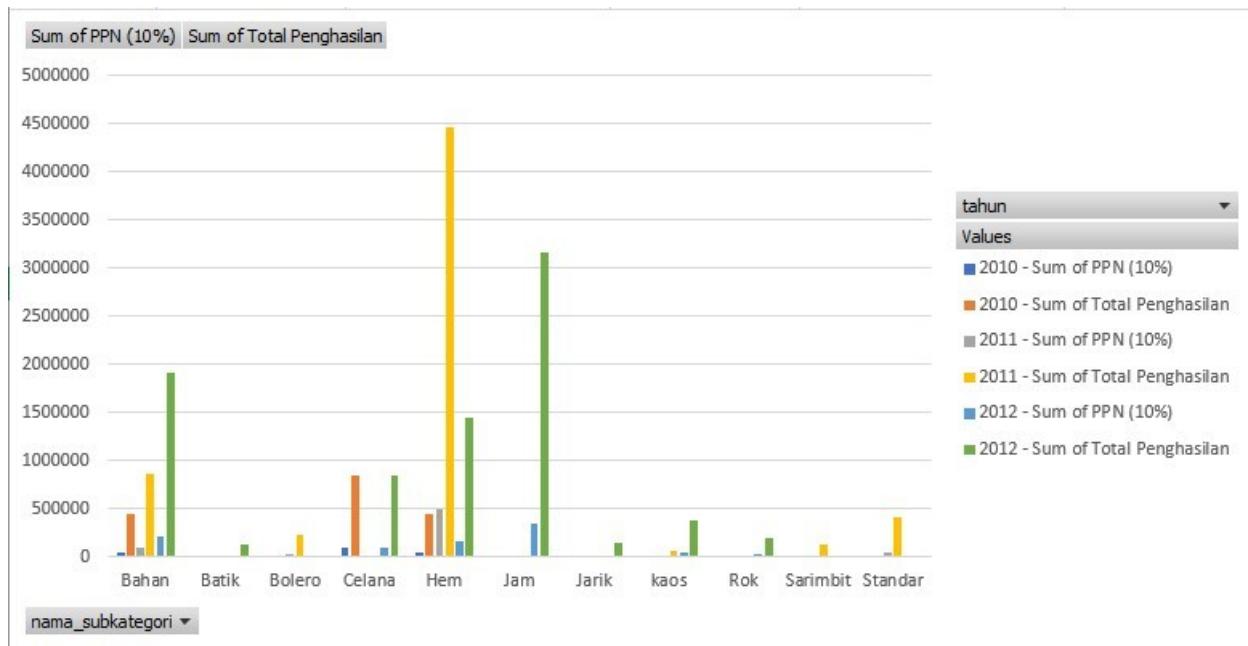
Total Sum of Pendapatan	Total Sum of PPN (10%)	Total Sum of Total Penghasilan	Total Sum of jumlah	Total Sum of kuartal
15045000	1504500	13540500	17	7
150000	15000	135000	1	3
255000	25500	229500	1	4
3740000	374000	3366000	34	2
19023000	1902300	17120700	17	14
3520000	352000	3168000	44	1
160000	16000	144000	4	2
1350000	135000	1215000	15	2
225000	22500	202500	1	4
150000	15000	135000	1	4
450000	45000	405000	2	4
<b>456073000</b>	<b>45607300</b>	<b>410465700</b>	<b>137</b>	<b>47</b>

2. Buatlah **PivotTable** dan **PivotChart** untuk melihat PPN dan Total Penghasilan tersebut selama 2010 – 2012. Kategori produk apakah yang memberikan nilai penghasilan terbanyak selama 3 tahun tersebut?

## PivotTable

Row Labels	Column Labels		2010			2011			2012			Total Sum of PPN (10%)	Total Sum of Total Penghasilan
	Sum of PPN (10%)	Sum of Total Penghasilan	Sum of PPN (10%)	Sum of Total Penghasilan	Sum of PPN (10%)	Sum of Total Penghasilan	Sum of PPN (10%)	Sum of Total Penghasilan	Sum of PPN (10%)	Sum of Total Penghasilan	Sum of PPN (10%)		
Bahan	50000	450000	96000	864000	212000	1908000	1504500	13540500	156000	1436400	1902300	17120700	13540500
Batik	0	0	0	0	0	15000	135000	135000	0	0	25500	229500	135000
Bolero	0	0	25500	229500	0	0	0	0	0	0	352000	3168000	352000
Celana	93500	841500	0	0	93500	841500	374000	3366000	374000	3168000	352000	3168000	3366000
Hem	50000	450000	496000	4464000	159600	1436400	1902300	17120700	159600	1436400	1902300	17120700	17120700
Jam	0	0	0	0	0	352000	3168000	3168000	0	0	352000	3168000	3168000
Jarik	0	0	0	0	0	16000	144000	144000	0	0	16000	144000	144000
kaos	0	0	6000	54000	42000	378000	352000	352000	0	0	135000	1215000	1215000
Rok	0	0	0	0	0	22500	202500	202500	0	0	22500	202500	202500
Sarimbit	0	0	15000	135000	0	0	0	0	0	0	15000	135000	135000
Standar	0	0	45000	405000	0	0	0	0	0	0	45000	405000	405000
<b>Grand Total</b>	<b>1506500</b>	<b>13558500</b>	<b>3003000</b>	<b>27027000</b>	<b>11569200</b>	<b>104122800</b>	<b>45607300</b>	<b>410465700</b>	<b>11569200</b>	<b>104122800</b>	<b>45607300</b>	<b>410465700</b>	<b>410465700</b>

## PivotChart



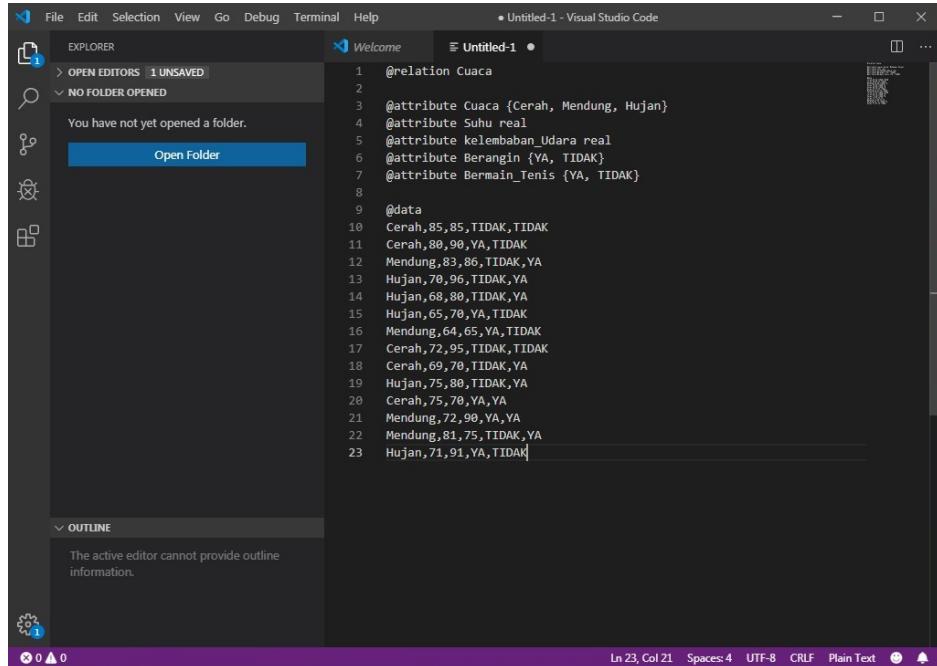
Dapat dilihat dari data di atas bahwa, kategori produk dengan nilai penghasilan terbanyak pada tahun 2010 – 2012 adalah **Hem**.

## MODUL 6

	A	B	C	D	E	F	G	H	I	J	K	L
1	Jurusan_SMA	Gender	Asal_sekolah	Rerata_SKS	Asisten	Lama_Studi		Jumlah kelas IPS	Jumlah kelas IPA	Jumlah kelas lain		
2	IPS	WANITA	SURAKARTA	18	TIDAK	TERLAMBAT		6	10	4		
3	IPA	PRIA	SURAKARTA	19	YA	TEPAT						
4	LAIN	PRIA	SURAKARTA	19	TIDAK	TERLAMBAT	Lama Studi Tempat	Lama Studi Terlambat				
5	IPA	PRIA	LUAR	17	TIDAK	TERLAMBAT		13	7			
6	IPA	WANITA	SURAKARTA	17	TIDAK	TEPAT						
7	IPA	WANITA	LUAR	18	YA	TEPAT	Rerata SKS Max	Rerata SKS Min	Rerata SKS Mean	Standar Deviation		
8	IPA	PRIA	SURAKARTA	18	TIDAK	TERLAMBAT		23	16	19	1.669384	
9	IPA	PRIA	SURAKARTA	19	TIDAK	TEPAT						
10	IPS	PRIA	LUAR	18	TIDAK	TERLAMBAT	Data Gabungan					
11	LAIN	WANITA	SURALARTA	18	TIDAK	TEPAT		3				
12	IPA	WANITA	SURAKARTA	19	TIDAK	TEPAT						
13	IPS	PRIA	SURAKARTA	20	TIDAK	TEPAT						
14	IPS	PRIA	SURAKARTA	19	TIDAK	TEPAT						
15	IPA	PRIA	SURAKARTA	19	TIDAK	TEPAT						
16	IPA	PRIA	LUAR	22	YA	TEPAT						
17	LAIN	PRIA	SURAKARTA	16	TIDAK	TERLAMBAT						
18	IPS	PRIA	LUAR	20	TIDAK	TEPAT						
19	LAIN	PRIA	LUAR	23	YA	TEPAT						
20	IPA	PRIA	SURAKARTA	21	YA	TEPAT						
21	IPS	PRIA	SURAKARTA	19	TIDAK	TERLAMBAT						

# MODUL 7

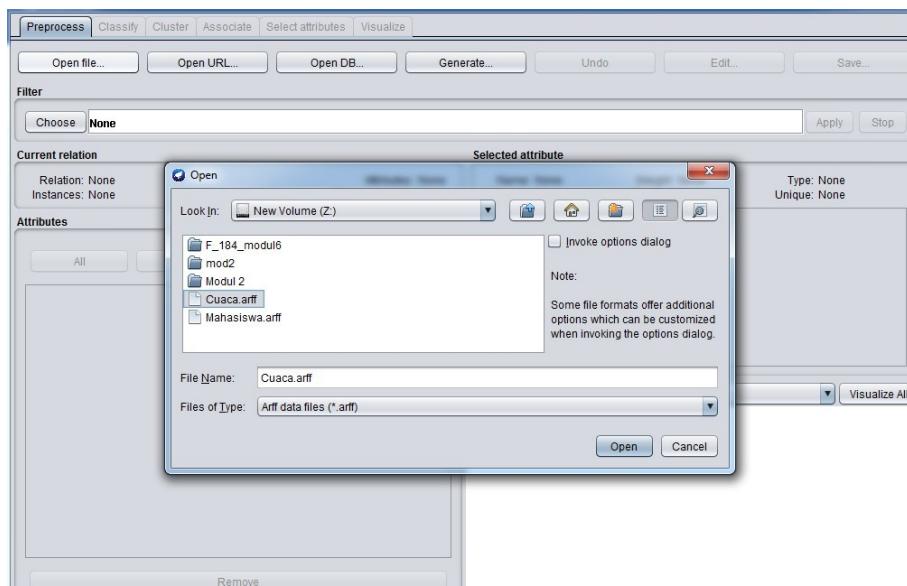
## LANGKAH-LANGKAH PRAKTIKUM

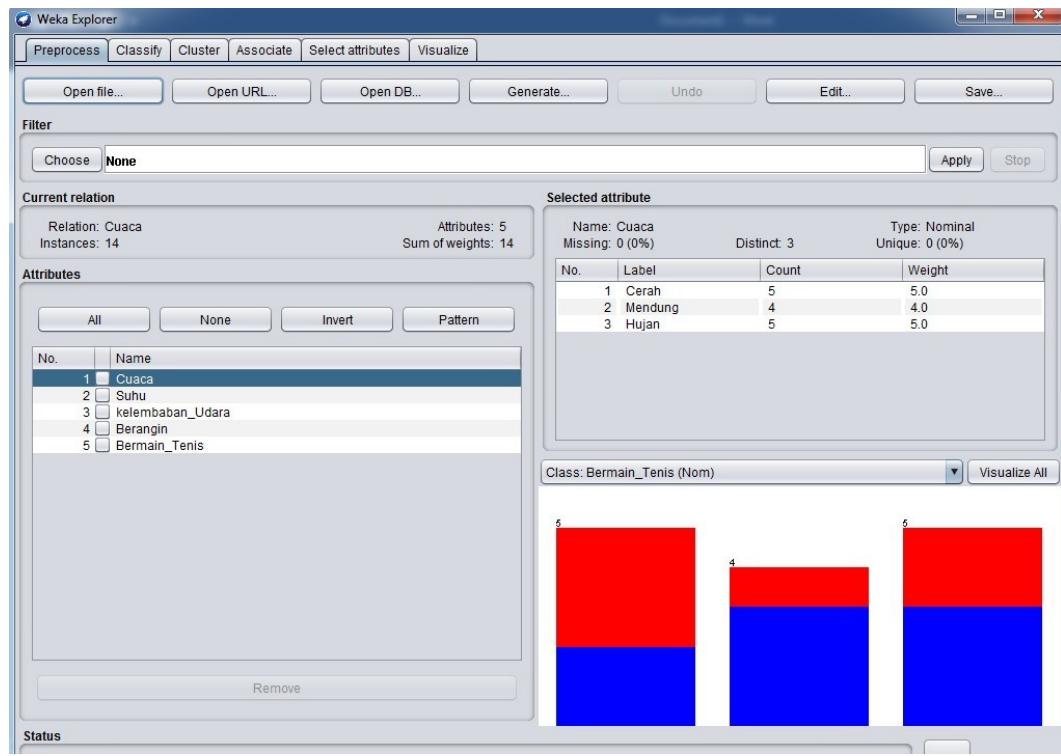


```
@relation Cuaca
@attribute Cuaca {Cerah, Mendung, Hujan}
@attribute Suhu real
@attribute kelembaban_Udara real
@attribute Berangin {YA, TIDAK}
@attribute Bermain_Tenis {YA, TIDAK}

@data
Cerah,85,85,TIDAK,TIDAK
Cerah,88,99,YA,TIDAK
Mendung,83,86,TIDAK,YA
Hujan,79,96,TIDAK,YA
Hujan,68,80,TIDAK,YA
Hujan,65,70,YA,TIDAK
Mendung,64,65,YA,TIDAK
Cerah,72,95,TIDAK,TIDAK
Cerah,69,70,TIDAK,YA
Hujan,75,88,TIDAK,YA
Cerah,75,70,YA,YA
Mendung,72,99,YA,YA
Mendung,81,75,TIDAK,YA
Hujan,71,91,YA,TIDAK
```

### Mennggunakan file ARFF dengan WEKA





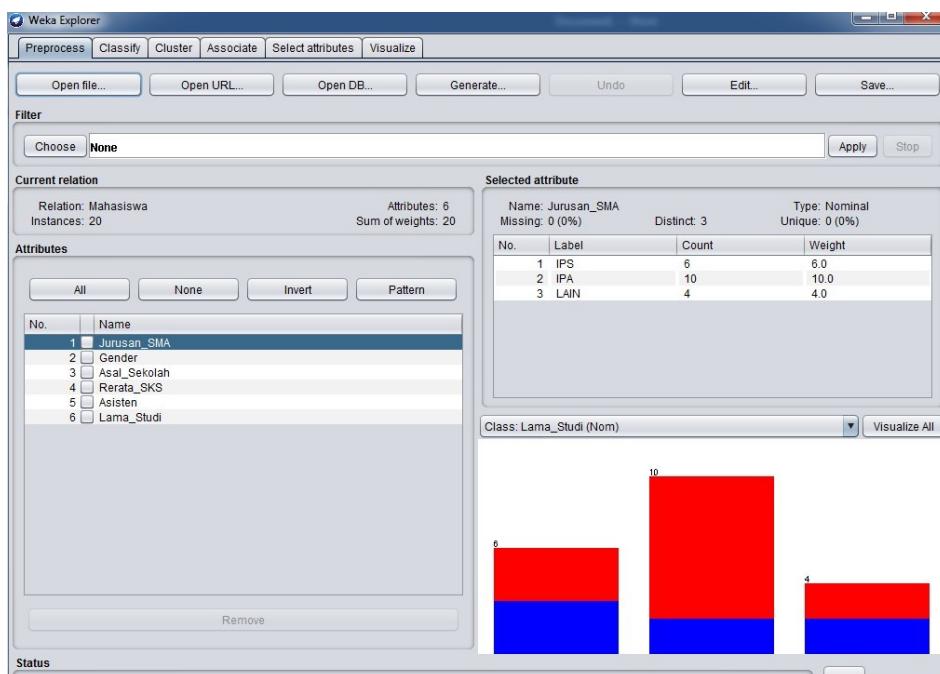
## TUGAS

1. Buat file bedasarkan tugas modul 6 soal nomor 1
2. File ARFF dan grafik

```
File Edit Selection View Go Debug Terminal Help
Mahasiswa.arff - Visual Studio Code

EXPLORER > OPEN EDITORS > Mahasiswa.arff
NO FOLDER OPENED You have not yet opened a folder.
Open Folder

Welcome Cuaca.arff Mahasiswa.arff ...
@relation Mahasiswa
1 @attribute Jurusan_SMA {IPS, IPA, LAIN}
2 @attribute Gender {WANITA, PRIA}
3 @attribute Asal_Sekolah {SURAKARTA, LUAR}
4 @attribute Rerata_SKS real
5 @attribute Asisten {TIDAK, YA}
6 @attribute Lama_Studi {TERLAMBAT, TEPAT}
7 @attribute
8 @data
9 IPS,WANITA,SURAKARTA,18,TIDAK,TERLAMBAT
10 IPA,PRIA,SURAKARTA,19,YA,TEPAT
11 LAIN,PRIA,SURAKARTA,19,TIDAK,TERLAMBAT
12 IPA,PRIA,LUAR,17,TIDAK,TERLAMBAT
13 IPA,WANITA,SURAKARTA,17,TIDAK,TEPAT
14 IPA,WANITA,LUAR,18,YA,TEPAT
15 IPA,WANITA,SURAKARTA,18,TIDAK,TERLAMBAT
16 IPA,PRIA,SURAKARTA,18,TIDAK,TERLAMBAT
17 IPA,PRIA,SURAKARTA,19,TIDAK,TEPAT
18 IPS,PRIA,LUAR,18,TIDAK,TERLAMBAT
19 LAIN,WANITA,SURAKARTA,18,TIDAK,TEPAT
20 IPA,WANITA,SURAKARTA,19,TIDAK,TEPAT
21 IPS,PRIA,SURAKARTA,20,TIDAK,TEPAT
22 IPS,PRIA,SURAKARTA,19,TIDAK,TEPAT
23 IPA,PRIA,SURAKARTA,19,TIDAK,TEPAT
24 IPA,PRIA,LUAR,22,YA,TEPAT
25 LAIN,PRIA,SURAKARTA,16,TIDAK,TERLAMBAT
26 IPS,PRIA,LUAR,28,TIDAK,TEPAT
27 LAIN,PRIA,LUAR,23,YA,TEPAT
28 IPA,PRIA,SURAKARTA,21,YA,TEPAT
29 IPS,PRIA,SURAKARTA,19,TIDAK,TERLAMBAT
30
```



3. Tulis di atas selembar kertas HVS
4. Berapa jumlah atribut yg bertipe binomial dan polynomial Bionomial : 4 (gender, Asal sekolah, Asisten, Lama studi)  
Polynomial : 1 (Jurusana SMA)
5. Berapa jumlah atribut bertipe Real  
Real : 1 (Rerata SKS)

6. Pada atribut Rerata SKS
- |                    |              |
|--------------------|--------------|
| Maximum            | : 23         |
| Minimum            | : 16         |
| Mena               | : 18,95      |
| Standard Deviation | : 1.66938375 |

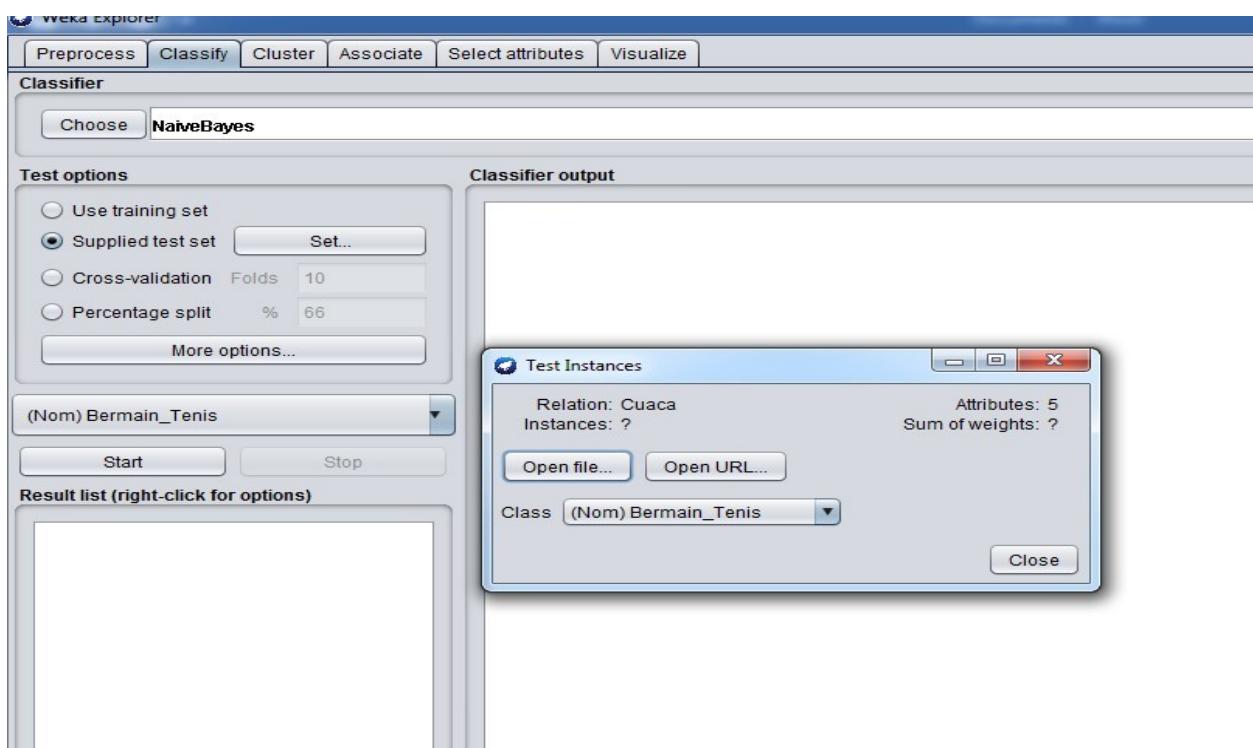
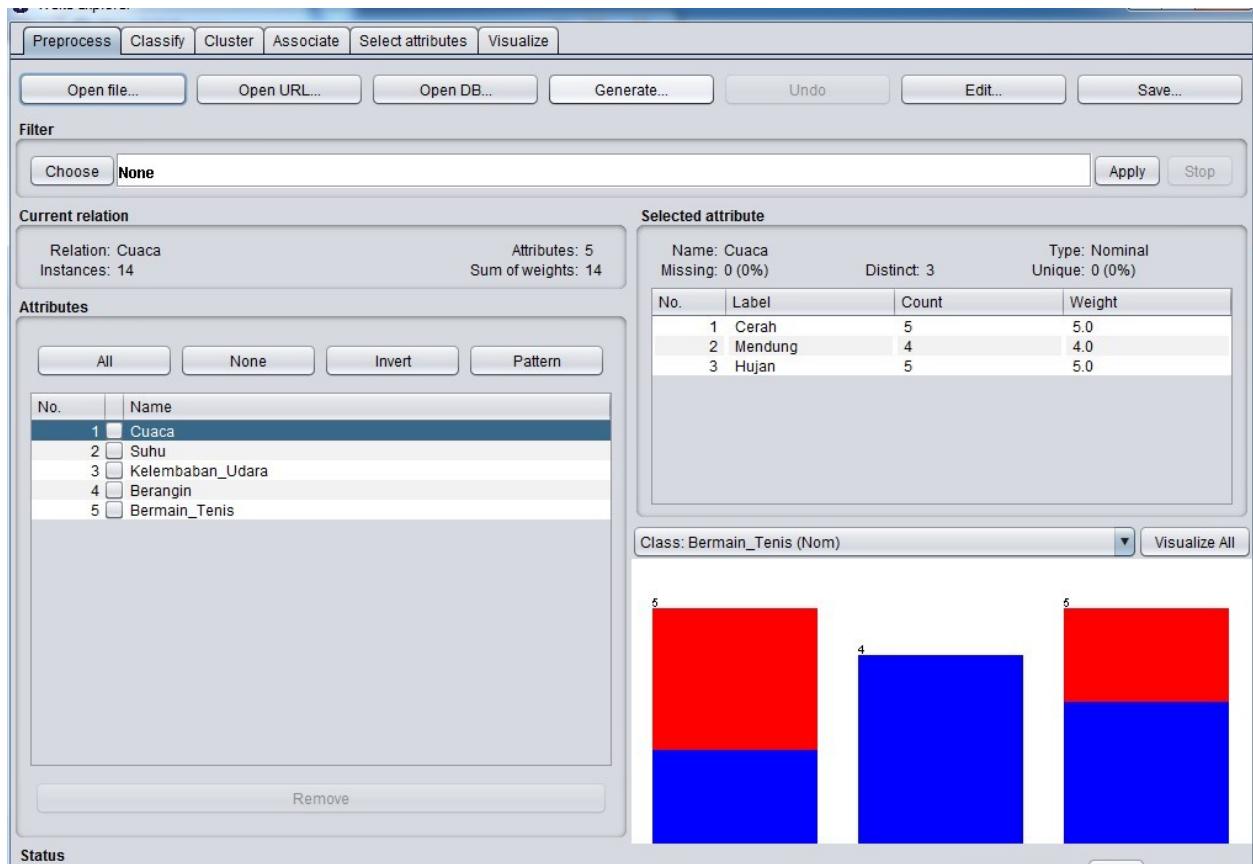
# MODUL 8

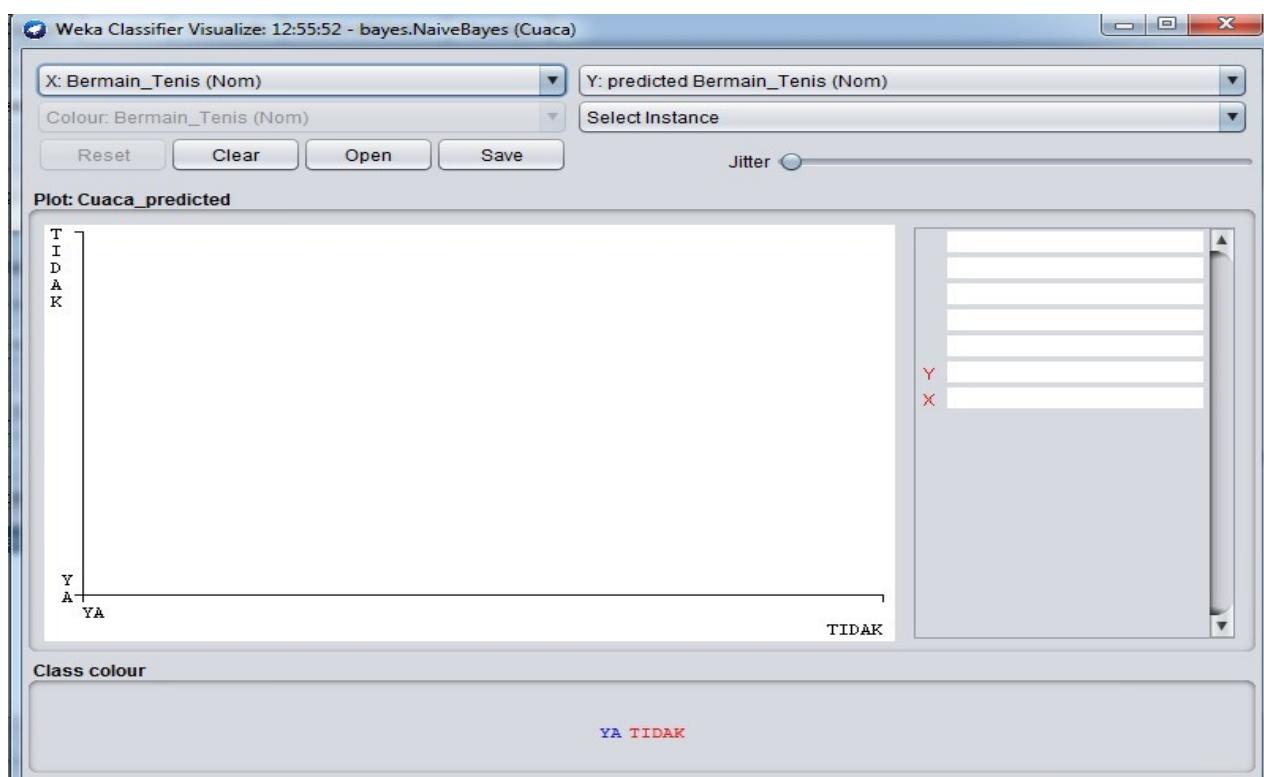
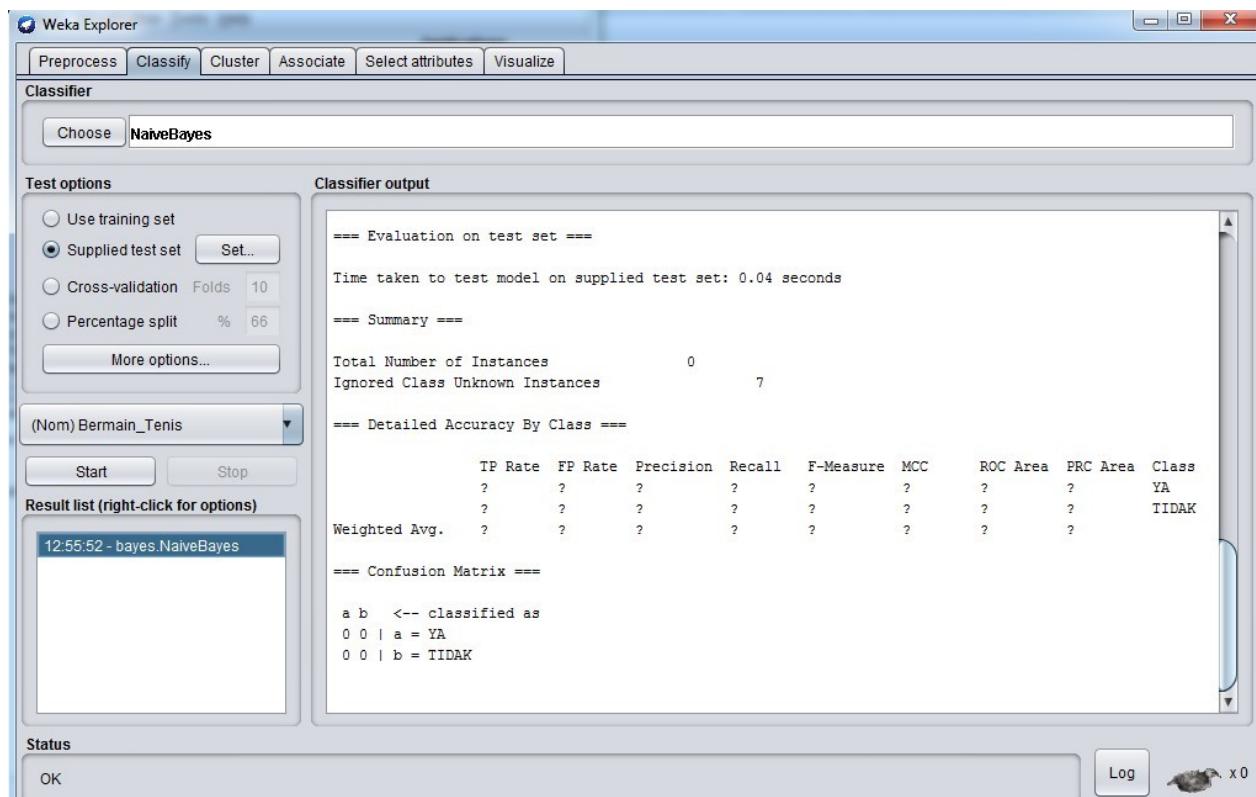
## KLASIFIKASI NAÏVE BAYES

### Langkah-langkah Praktikum

#### Implementa Naïve Bayes dengan Weka

```
Welcome Cuaca.arff CuacaTesting.arff ×
C: > Users > LABSI-20 > Documents > CuacaTesting.arff
1 @relation Cuaca
2
3 @attribute Cuaca {Cerah, Mendung, Hujan}
4 @attribute Suhu real
5 @attribute Kelembaban_Udara real
6 @attribute Berangin {YA, TIDAK}
7 @attribute Bermain_Tenis {YA, TIDAK}
8
9 @data
10 Cerah,75,65,TIDAK,?
11 Cerah,80,68,YA,?
12 Cerah,83,87,YA,?
13 Mendung,70,96,TIDAK,?
14 Mendung,68,81,TIDAK,?
15 Hujan,65,75,YA,?
16 Hujan,64,85,YA,?
```





ARFF-Viewer - Z:\HasilPrediksi.arff

File Edit View  
HasilPrediksi.arff

Relation: Cuaca\_predicted

No.	Cuaca	Suhu	Kelembaban_Udara	Berangin	prediction margin	predicted	Bermain_Tenis
	Nominal	Numeric	Numeric	Nominal	Numeric	Nominal	Nominal
1	Cerah	75.0	65.0	TIDAK	0.762765	YA	
2	Cerah	80.0	68.0	YA	0.087878	YA	
3	Cerah	83.0	87.0	YA	-0.676866	TIDAK	
4	Mend...	70.0	96.0	TIDAK	0.628523	YA	
5	Mend...	68.0	81.0	TIDAK	0.833996	YA	
6	Hujan	65.0	75.0	YA	0.253733	YA	
7	Hujan	64.0	85.0	YA	-0.160143	TIDAK	

## Implementasi Naïve Bayes dengan RapidMiner

Import Data - Select the cells to import. X

Select the cells to import.

Sheet: Training ▾ Cell range: A:E Select All  Define header row: 1

A	B	C	D	E
1 Cuaca	Suhu	Kelembaban_udara	Berangin	Bermain_tenis
2 Cerah	85.000	85.000	TIDAK	TIDAK
3 Cerah	80.000	90.000	YA	TIDAK
4 Mendung	83.000	86.000	TIDAK	YA
5 Hujan	70.000	96.000	TIDAK	YA
6 Hujan	68.000	80.000	TIDAK	YA
7 Hujan	65.000	70.000	YA	TIDAK
8 Mendung	64.000	65.000	YA	YA
9 Cerah	72.000	95.000	TIDAK	TIDAK
10 Cerah	69.000	70.000	TIDAK	YA
11 Hujan	75.000	80.000	TIDAK	YA
12 Cerah	75.000	70.000	YA	YA
13 Mendung	72.000	90.000	YA	YA
14 Mendung	81.000	75.000	TIDAK	YA
15 Hujan	74.000	84.000	YA	TIDAK

← Previous    → Next    X Cancel

Import Data - Format your columns.

### Format your columns.

Replace errors with missing values (i)

	Cuaca polynomial	Suhu integer	Kelembaban_u... integer	Berangin polynomial	Bermain_tenis binomial
1	Cerah	85	85	TIDAK	TIDAK
2	Cerah				
3	Mendung				
4	Hujan				
5	Hujan				
6	Hujan				
7	Mendung				
8	Cerah				
9	Cerah				
10	Hujan				
11	Cerah	75	70	YA	YA
12	Mendung	72	90	YA	YA
13	Mendung	81	75	TIDAK	YA

 Change role  
Please enter the new role:

 no problems.

Import Data - Format your columns.

### Format your columns.

Replace errors with missing values (i)

	Cuaca polynomial	Suhu integer	Kelembaban_u... integer	Berangin polynomial	Bermain_tenis binomial
2	Cerah	80	90	YA	
3	Mendung	83	86	TIDAK	YA
4	Hujan	70	96	TIDAK	YA
5	Hujan	68	80	TIDAK	YA
6	Hujan	65	70	YA	TIDAK
7	Mendung	64	65	YA	YA
8	Cerah	72	95	TIDAK	TIDAK
9	Cerah	69	70	TIDAK	YA
10	Hujan	75	80	TIDAK	YA
11	Cerah	75	70	YA	YA
12	Mendung	72	90	YA	YA
13	Mendung	81	75	TIDAK	YA
14	Hujan	71	91	YA	TIDAK

 no problems.

Views: Design Results Turbo Prep Auto Model Find data, operators...etc All Studio

ExampleSet //Local Repository/DataCuaca\_Training ExampleSet //Local Repository/DataCuaca\_Training ExampleSet //Local Repository/DataCuaca\_Testing

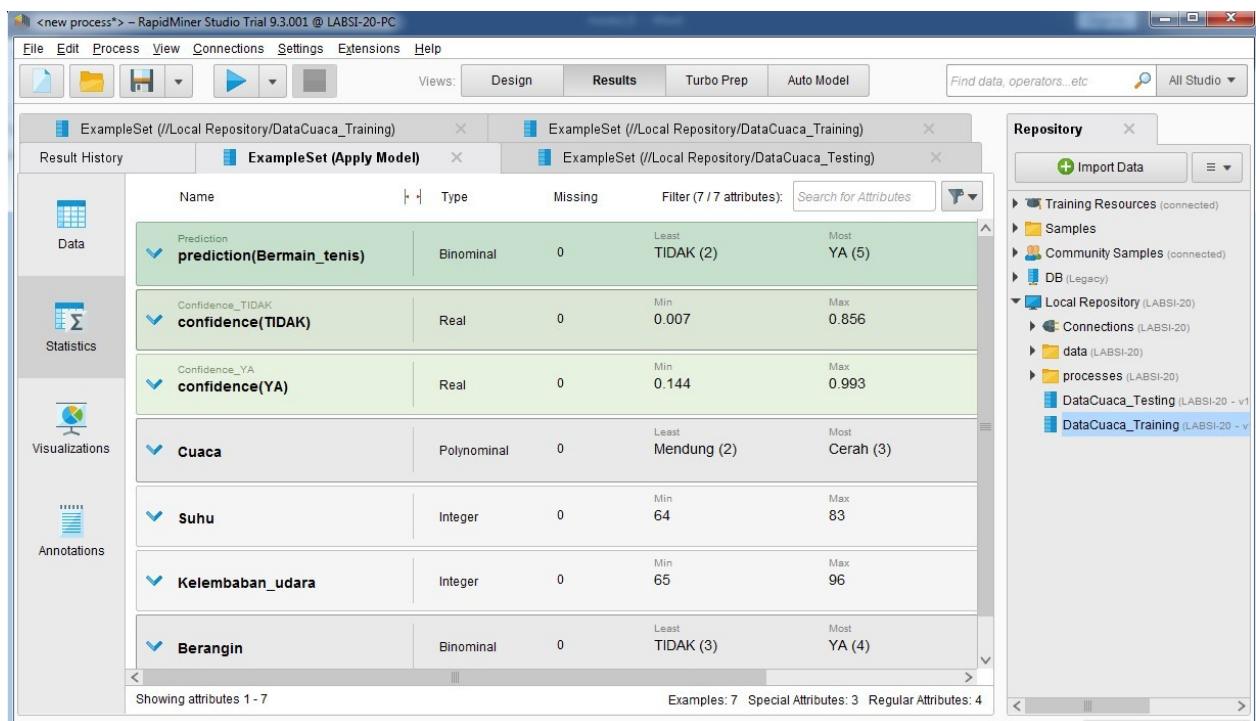
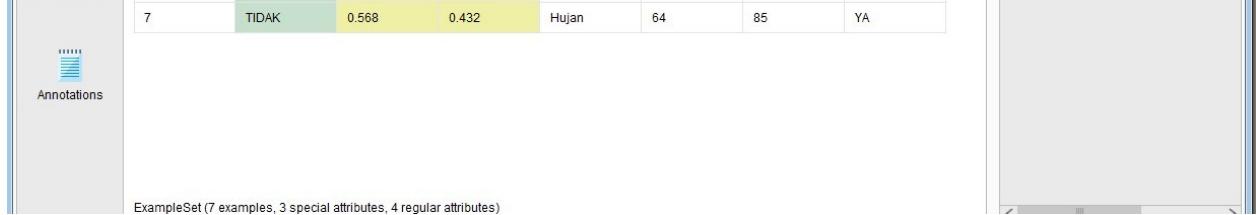
Result History ExampleSet (Apply Model) ExampleSet //Local Repository/DataCuaca\_Testing

Data Open in Turbo Prep Auto Model Filter (7 / 7 examples): all

Row No.	prediction(B...)	confidence(...)	confidence(...)	Cuaca	Suhu	Kelembaban...	Berangin
1	YA	0.154	0.846	Cerah	75	65	TIDAK
2	YA	0.498	0.502	Cerah	80	68	YA
3	TIDAK	0.856	0.144	Cerah	83	87	YA
4	YA	0.019	0.981	Mendung	70	96	TIDAK
5	YA	0.007	0.993	Mendung	68	81	TIDAK

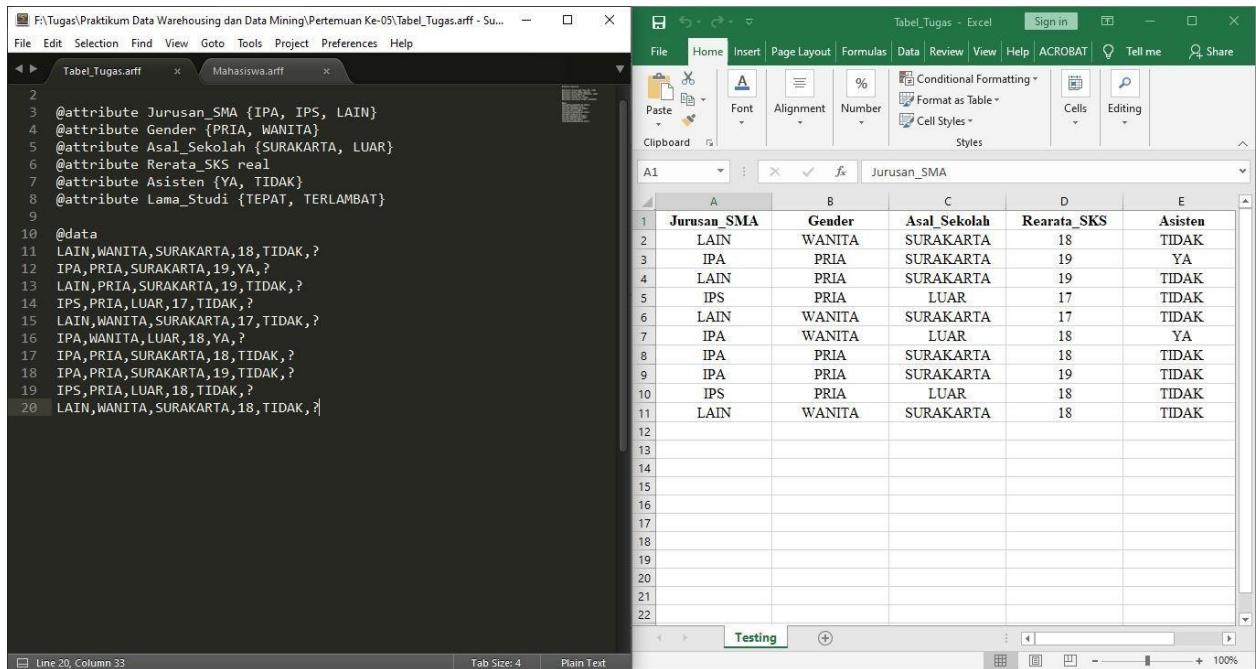
Repository + Import Data

- Training Resources (connected)
- Samples
- Community Samples (connected)
- DB (Legacy)
- Local Repository (LABSI-20)
  - Connections (LABSI-20)
  - data (LABSI-20)
  - processes (LABSI-20)
  - DataCuaca\_Testing (LABSI-20 - v1)

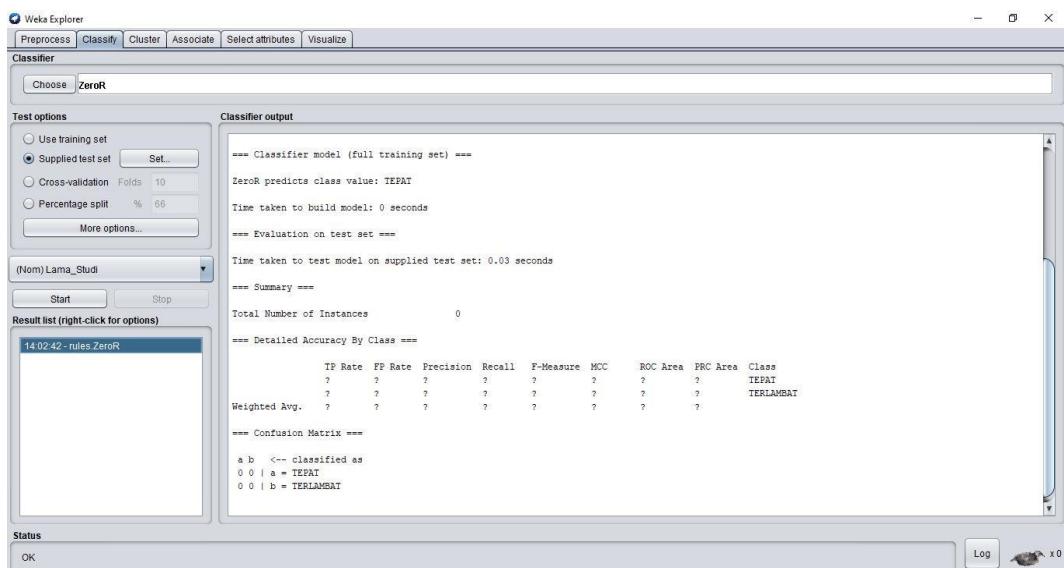
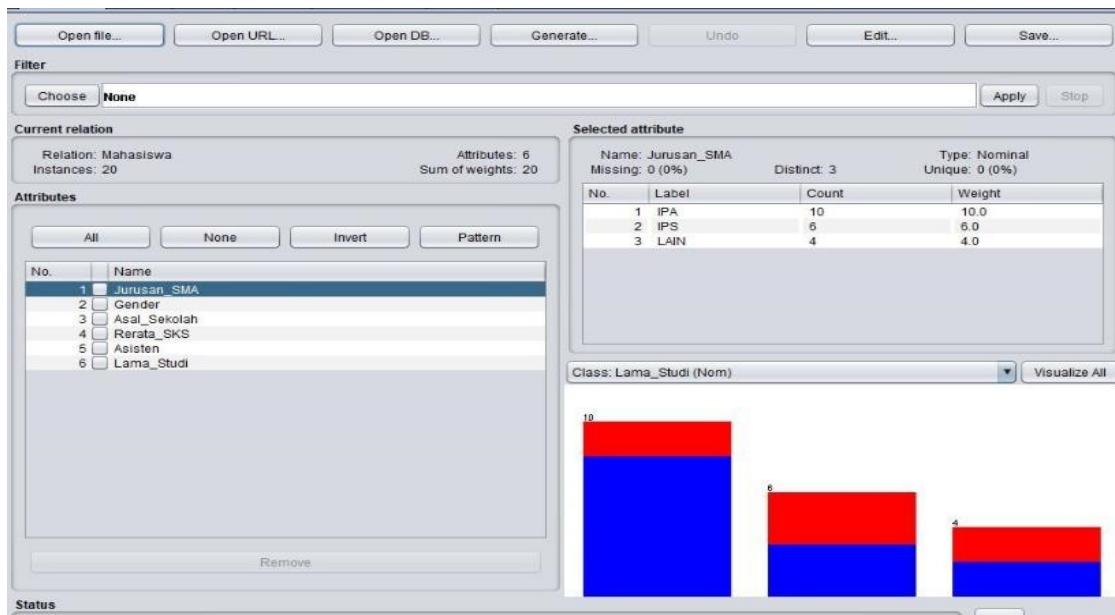


## Tugas

File format (EXCEL) dan (ARFF)



Prediksi dengan menggunakan WEKA



Relation: Mahasiswa\_predicted

No. 1: Jurusan\_SMA 2: Gender 3: Asal\_Sekolah 4: Rerata\_SKS 5: Asisten 6: prediction margin 7: predicted Lama\_Studi 8: Lama\_Studi

	Nominal	Nominal	Nominal	Numeric	Nominal	Numeric	Nominal	Nominal
1	LAIN	WANITA	SURAKARTA	18.0	TIDAK	-0.375862	TERLAMBAT	
2	IPA	PRIA	SURAKARTA	19.0	YA	0.836469	TEPAT	
3	LAIN	PRIA	SURAKARTA	19.0	TIDAK	-0.175169	TERLAMBAT	
4	IPS	PRIA	LUAR	17.0	TIDAK	-0.713206	TERLAMBAT	
5	LAIN	WANITA	SURAKARTA	17.0	TIDAK	-0.546846	TERLAMBAT	
6	IPA	WANITA	LUAR	18.0	YA	0.757815	TEPAT	
7	IPA	PRIA	SURAKARTA	18.0	TIDAK	-0.125076	TERLAMBAT	
8	IPA	PRIA	SURAKARTA	19.0	TIDAK	0.356012	TEPAT	
9	IPS	PRIA	LUAR	18.0	TIDAK	-0.588286	TERLAMBAT	
10	LAIN	WANITA	SURAKARTA	18.0	TIDAK	-0.375862	TERLAMBAT	

## Prediksi dengan Rapidminer

### Training

Import Data - Select the cells to import.

Select the cells to import.

Sheet: Training ▾ Cell range: A:F Select All  Define header row: 1

	A	B	C	D	E	F
1	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten	Lama_Studi
2	IPS	WANITA	SURAKARTA	18.000	TIDAK	TERLAMBAT
3	IPA	PRIA	SURAKARTA	19.000	YA	TEPAT
4	LAIN	PRIA	SURAKARTA	19.000	TIDAK	TERLAMBAT
5	IPA	PRIA	LUAR	17.000	TIDAK	TERLAMBAT
6	IPA	WANITA	SURAKARTA	17.000	TIDAK	TEPAT
7	IPA	WANITA	LUAR	18.000	YA	TEPAT
8	IPA	PRIA	SURAKARTA	18.000	TIDAK	TERLAMBAT
9	IPA	PRIA	SURAKARTA	19.000	TIDAK	TEPAT
10	IPS	PRIA	LUAR	18.000	TIDAK	TERLAMBAT
11	LAIN	WANITA	SURAKARTA	18.000	TIDAK	TEPAT
12	IPA	WANITA	SURAKARTA	19.000	TIDAK	TEPAT
13	IPS	PRIA	SURAKARTA	20.000	TIDAK	TEPAT
14	IPS	PRIA	SURAKARTA	19.000	TIDAK	TEPAT
15	IPA	PRIA	SURAKARTA	18.000	TIDAK	TEPAT

← Previous → Next ✖ Cancel

Import Data - Format your columns.

Format your columns.

Replace errors with missing values ⓘ

	Jurusan_SMA * polynominal	Gender * polynominal	Asal_Sekolah * polynominal	Rerata_SKS * integer	Asisten * polynominal	Lama_Studi * binominal label
1	IPS	WANITA	SURAKARTA	18	TIDAK	TERLAMBAT
2	IPA	PRIA	SURAKARTA	19	YA	TEPAT
3	LAIN	PRIA	SURAKARTA	19	TIDAK	TERLAMBAT
4	IPA	PRIA	LUAR	17	TIDAK	TERLAMBAT
5	IPA	WANITA	SURAKARTA	17	TIDAK	TEPAT
6	IPA	WANITA	LUAR	18	YA	TEPAT
7	IPA	PRIA	SURAKARTA	18	TIDAK	TERLAMBAT
8	IPA	PRIA	SURAKARTA	19	TIDAK	TEPAT
9	IPS	PRIA	LUAR	18	TIDAK	TERLAMBAT
10	LAIN	WANITA	SURAKARTA	18	TIDAK	TEPAT
11	IPA	WANITA	SURAKARTA	19	TIDAK	TEPAT
12	IPS	PRIA	SURAKARTA	20	TIDAK	TEPAT
13	IPS	PRIA	SURAKARTA	19	TIDAK	TEPAT

✓ no problems.

← Previous → Next ✖ Cancel

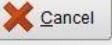
## Testing

Import Data - Select the cells to import.

**Select the cells to import.**

Sheet: Testing ▾ Cell range: A:E Select All  Define header row: 1

	A	B	C	D	E
1	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
2	LAIN	WANITA	SURAKARTA	18.000	TIDAK
3	IPA	PRIA	SURAKARTA	19.000	YA
4	LAIN	PRIA	SURAKARTA	19.000	TIDAK
5	IPS	PRIA	LUAR	17.000	TIDAK
6	LAIN	WANITA	SURAKARTA	17.000	TIDAK
7	IPA	WANITA	LUAR	18.000	YA
8	IPA	PRIA	SURAKARTA	18.000	TIDAK
9	IPA	PRIA	SURAKARTA	19.000	TIDAK
10	IPS	PRIA	LUAR	18.000	TIDAK
11	LAIN	WANITA	SURAKARTA	18.000	TIDAK

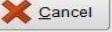
← Previous → Next 

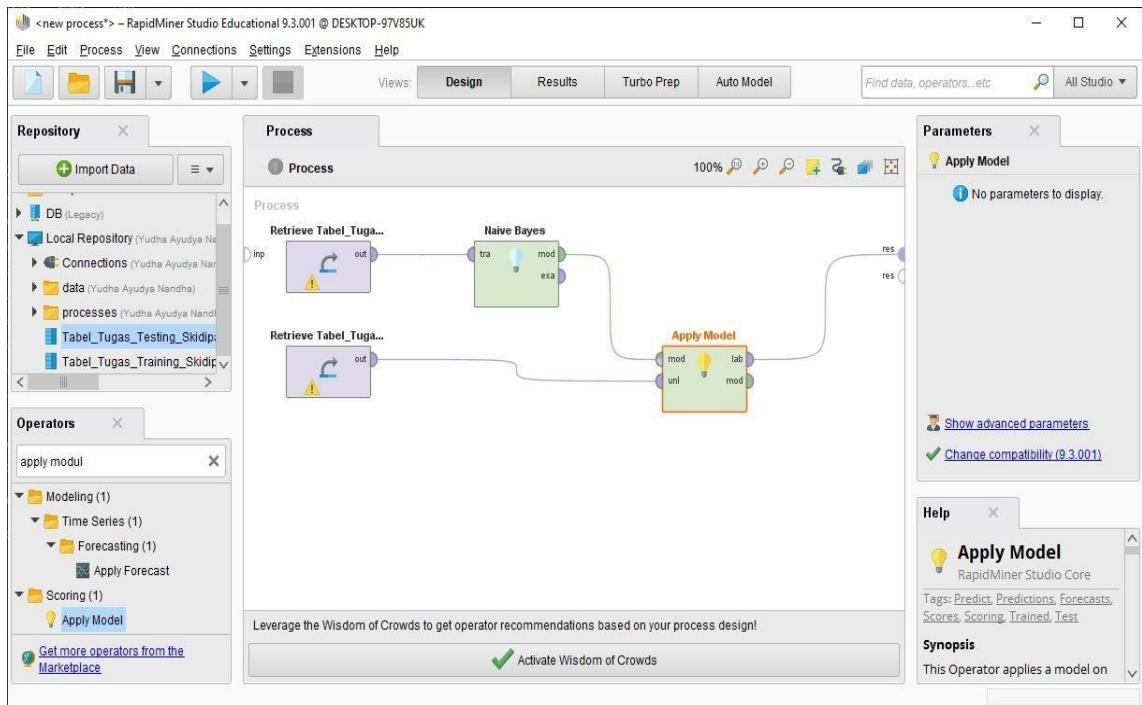
Import Data - Format your columns.

**Format your columns.**

Replace errors with missing values ⓘ

	Jurusan_SMA * polynominal	Gender * polynominal	Asal_Sekolah * polynominal	Rerata_SKS * integer	Asisten * binominal
1	LAIN	WANITA	SURAKARTA	18	TIDAK
2	IPA	PRIA	SURAKARTA	19	YA
3	LAIN	PRIA	SURAKARTA	19	TIDAK
4	IPS	PRIA	LUAR	17	TIDAK
5	LAIN	WANITA	SURAKARTA	17	TIDAK
6	IPA	WANITA	LUAR	18	YA
7	IPA	PRIA	SURAKARTA	18	TIDAK
8	IPA	PRIA	SURAKARTA	19	TIDAK
9	IPS	PRIA	LUAR	18	TIDAK
10	LAIN	WANITA	SURAKARTA	18	TIDAK

 no problems. ← Previous → Next 



The screenshot shows the "Results History" view in RapidMiner Studio. The table displays 10 examples with the following data:

Row No.	prediction(L...)	confidence(...)	confidence(...)	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	TERLAMBAT	0.648	0.352	LAIN	WANITA	SURAKARTA	18	TIDAK
2	TEPAT	0.005	0.995	IPA	PRIA	SURAKARTA	19	YA
3	TERLAMBAT	0.650	0.350	LAIN	PRIA	SURAKARTA	19	TIDAK
4	TERLAMBAT	0.868	0.132	IPS	PRIA	LUAR	17	TIDAK
5	TERLAMBAT	0.738	0.262	LAIN	WANITA	SURAKARTA	17	TIDAK
6	TEPAT	0.005	0.995	IPA	WANITA	LUAR	18	YA
7	TERLAMBAT	0.547	0.453	IPA	PRIA	SURAKARTA	18	TIDAK
8	TEPAT	0.321	0.679	IPA	PRIA	SURAKARTA	19	TIDAK
9	TERLAMBAT	0.811	0.189	IPS	PRIA	LUAR	18	TIDAK
10	TERLAMBAT	0.648	0.352	LAIN	WANITA	SURAKARTA	18	TIDAK

4. Nilai Rerata hasil dari prediksi RapidMiner untuk kolom Lama\_Studi

Rerata untuk **Tepat** : **0,476**

Rerata untuk **Tepat** : **0,524**

5. Dari hasil prediksi RapidMiner dapat disimpulkan bahwa 3 Mahasiswa akan lulus TEPAT dan 7 Mahasiswa akan lulus TERLAMBAT

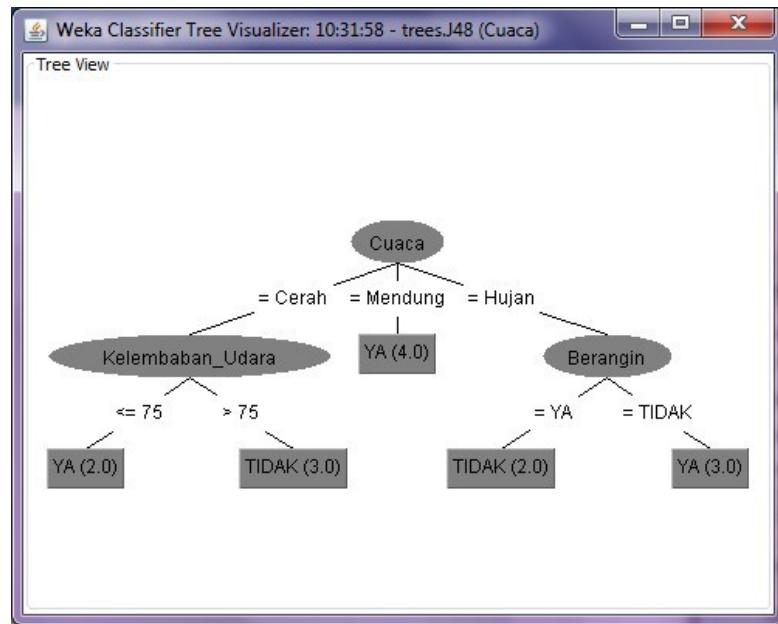
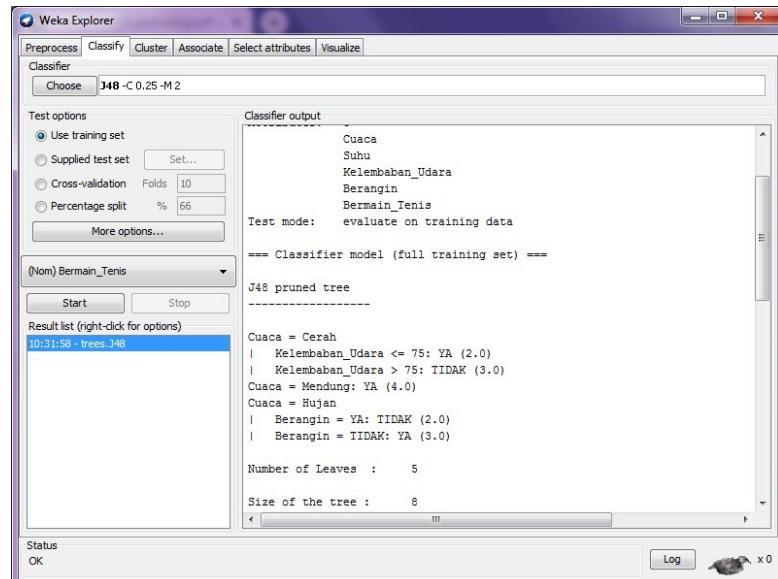
6. Dewi Row No. 1

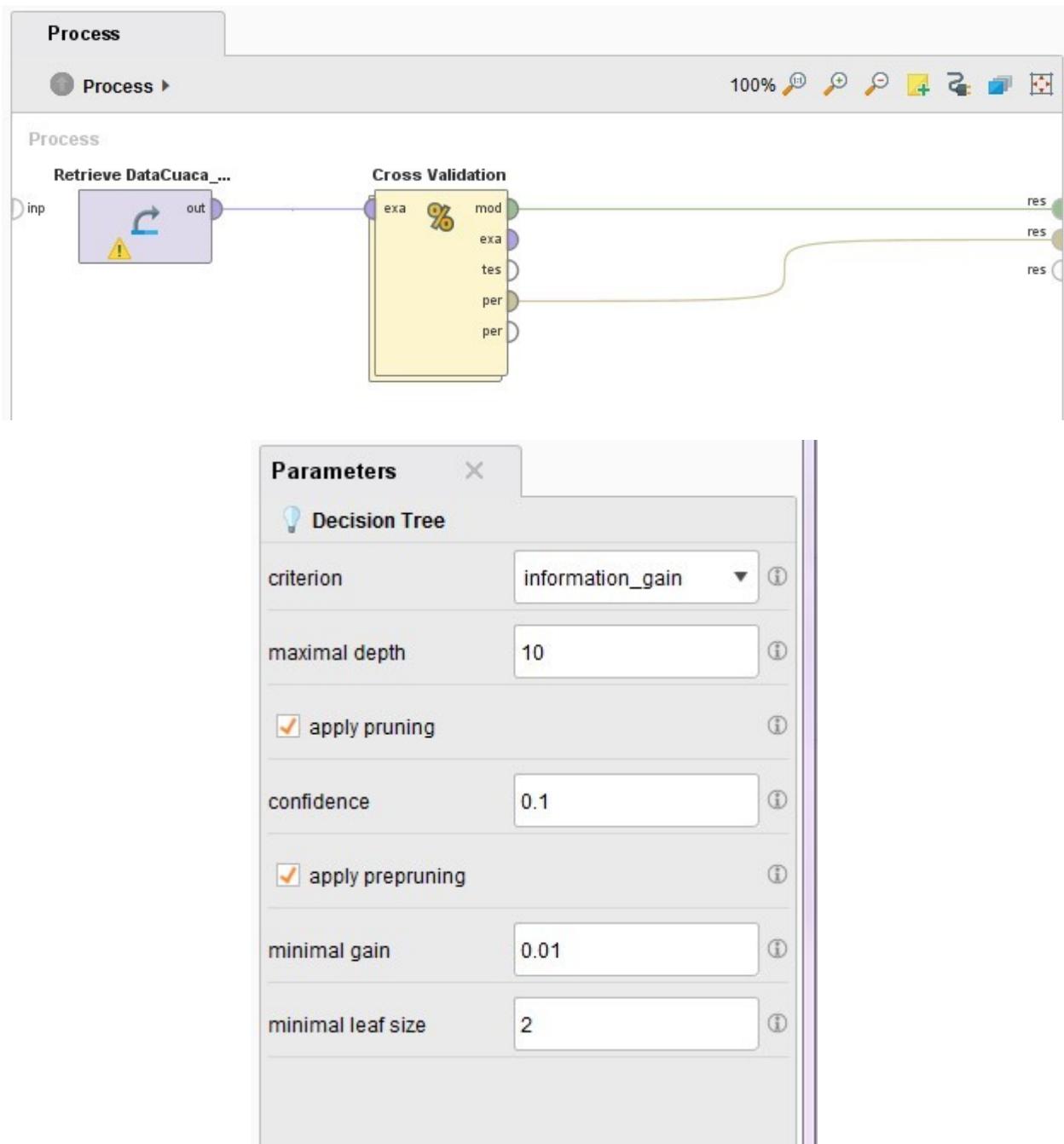
## 7. Jono Row No. 2

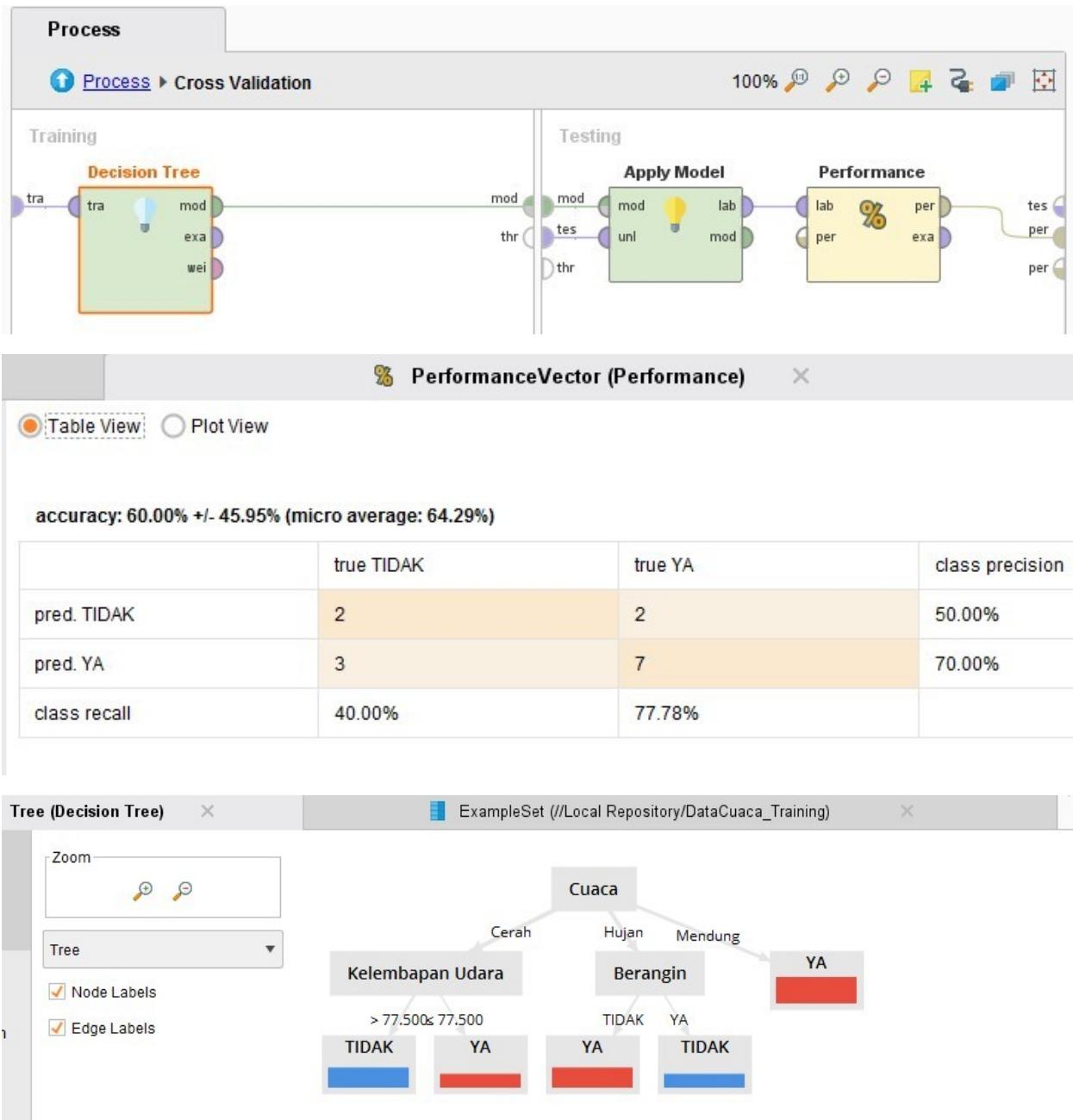
Row No.	prediction(L...	confidence(...	confidence(...	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	TEPAT	0.298	0.702	IPA	WANITA	LUAR	18	TIDAK
2	TEPAT	0.076	0.924	LAIN	PRIA	SURAKARTA	17	YA

# MODUL 9

## LANGKAH-LANGKAH PRAKTIKUM





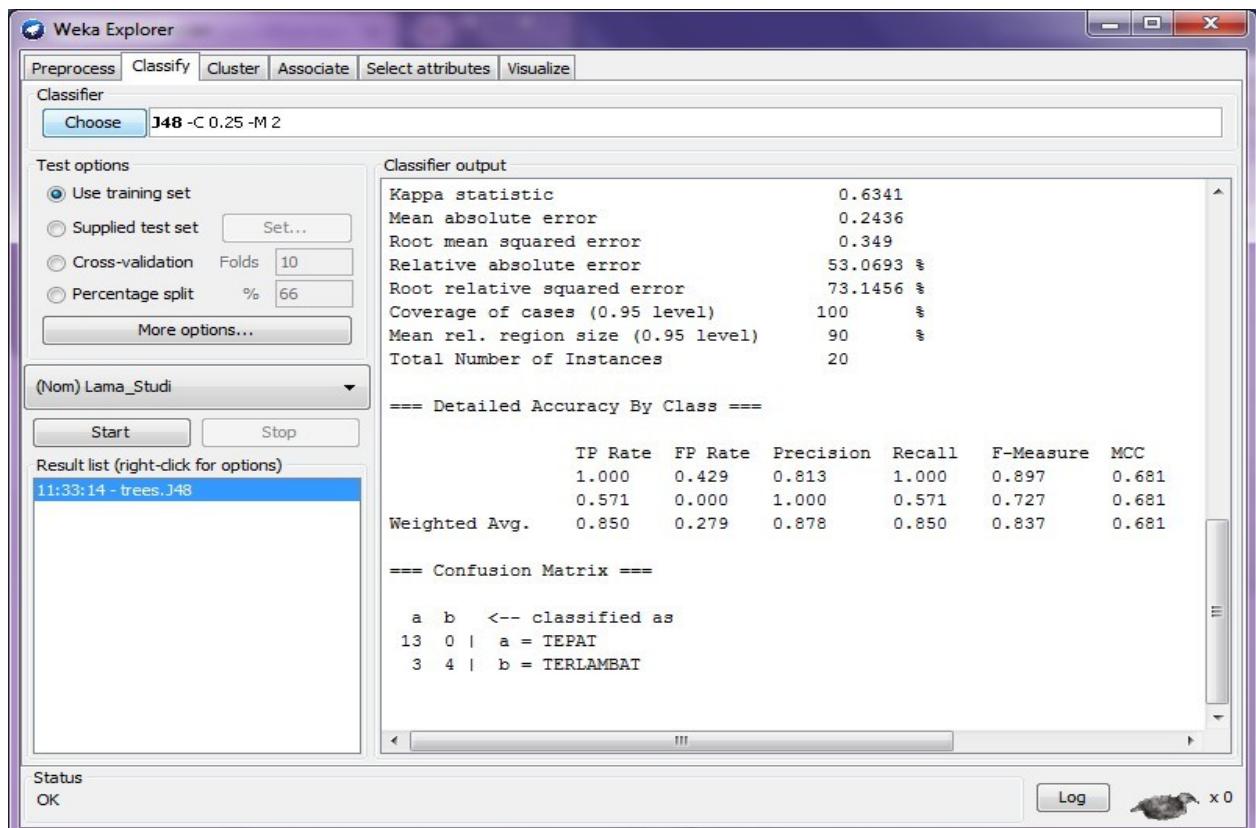


## TUGAS

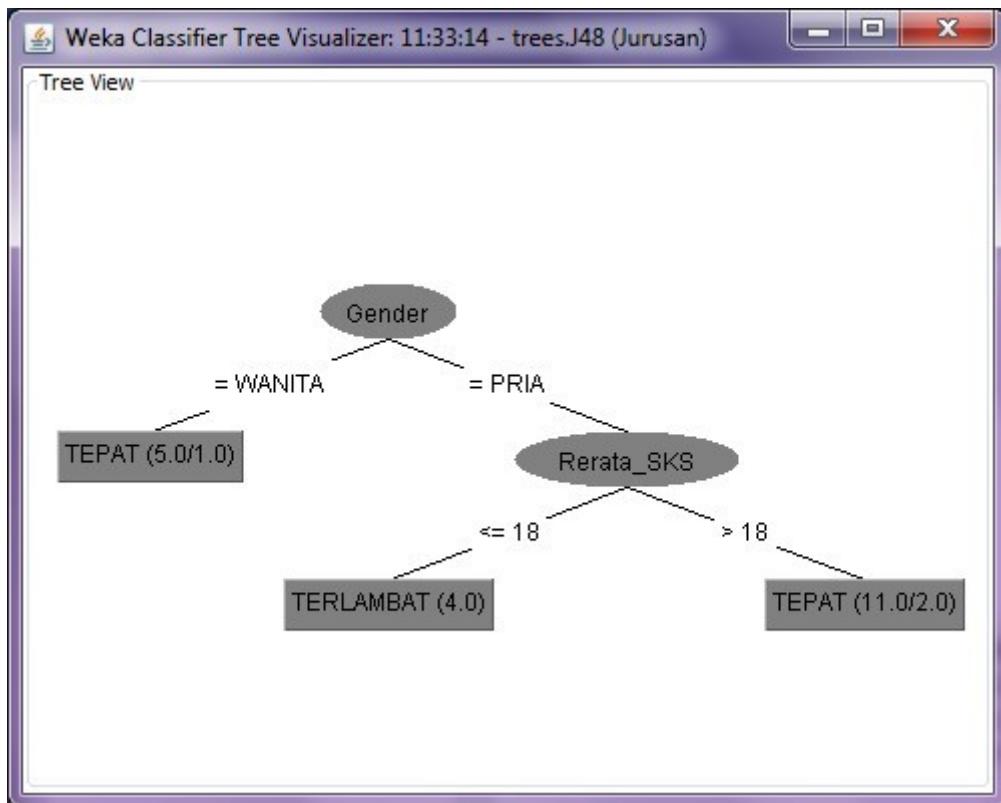
1.

Cuaca	Suhu	Kelembaban_Udar a	Berangin	Brmain_Tenis
Cerah	75	65	TIDAK	YA
Cerah	80	68	YA	YA
Cerah	83	87	YA	TIDAK
Medung	70	96	TIDAK	YA
Medung	68	81	TIDAK	YA
Hujan	65	75	TIDAK	YA
Hujan	64	85	YA	TIDAK

2.



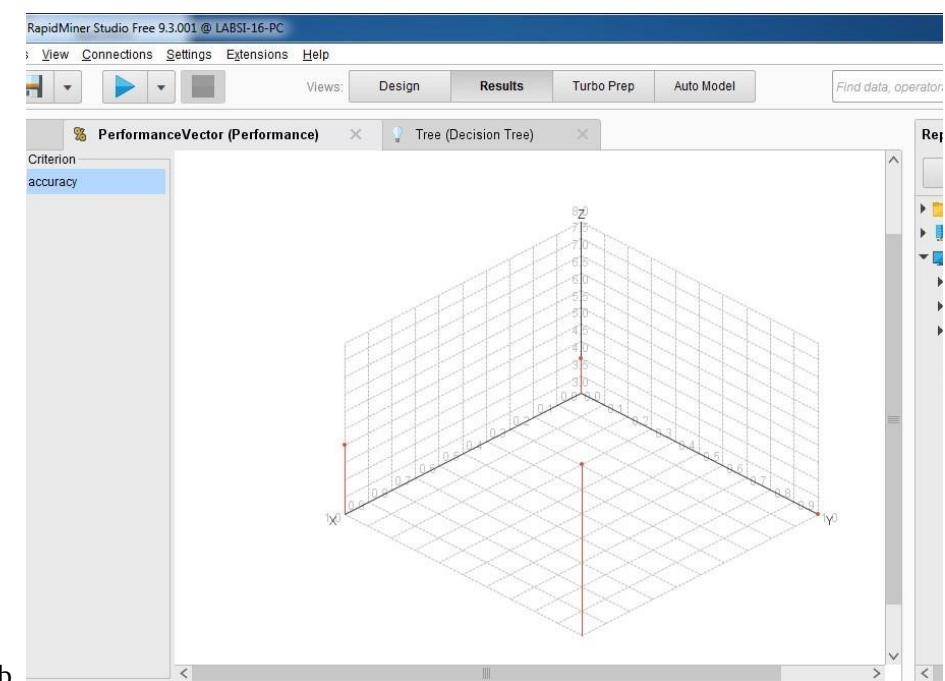
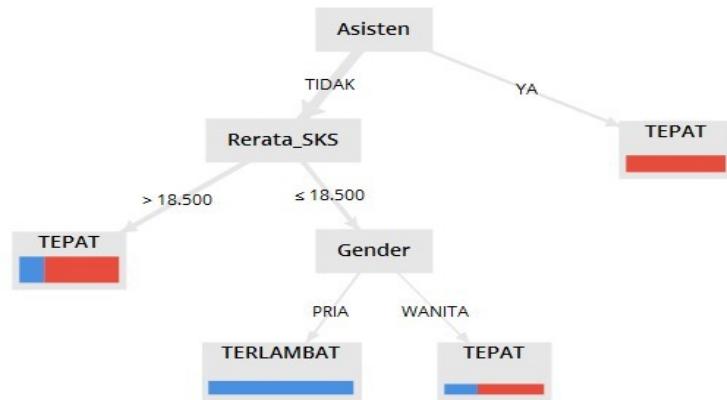
A.



B. Carilah nilai-nilai parameter berikut:

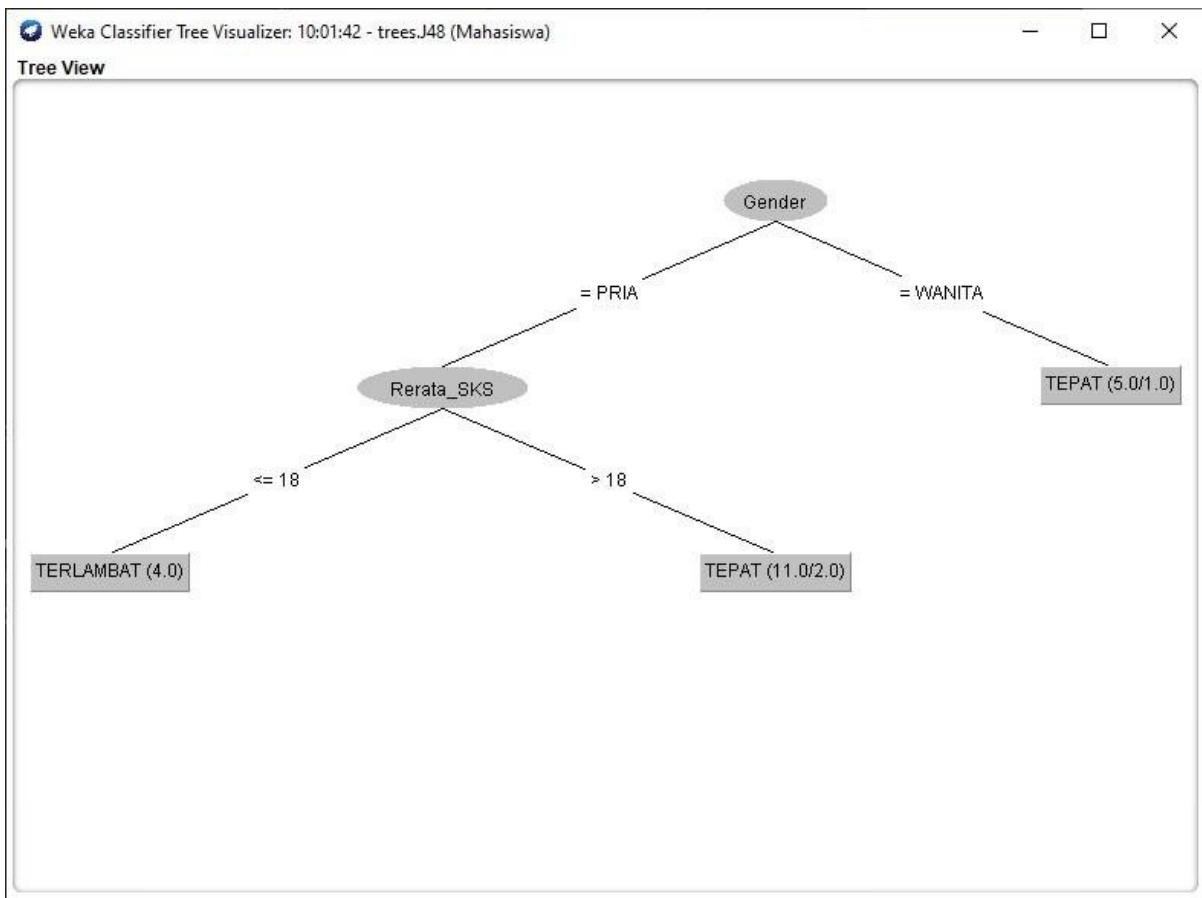
- i. Jumlah simpul daun pada pohon keputusan = 3
- ii. Jumlah simpul keseluruhan pada pohon keputusan = 5
- iii. Waktu yg di butuhkan untuk proses pelatihan = 0 detik
- iv. Tingkat ketepatan Klasifikasi = 85 %
- v. Tingkatan ketidaktepatan klasifikasi = 15 %

3. a.



b.

4. tentukan klasifikasi yang terbentuk berdasarkan kondisinya sesuai dengan simpul-simpulnya.



- a. Mahasiswa akan lulus TEPAT waktu jika:
  - i. PRIA = Rerata\_SKS > 18 (Nilai attribute lain diabaikan).
  - ii. WANITA = Semua akan lulus TEPAT waktu (Nilai attribute lain diabaikan).
- b. Mahasiswa akan lulus TERLAMBAT jika:
 

PRIA = Rerata\_SKS  $\leq$  18 (Nilai attribute lain diabaikan). ii. WANITA = Tidak ada yang akan lulus TERLAMBAT (Nilai attribute lain diabaikan).

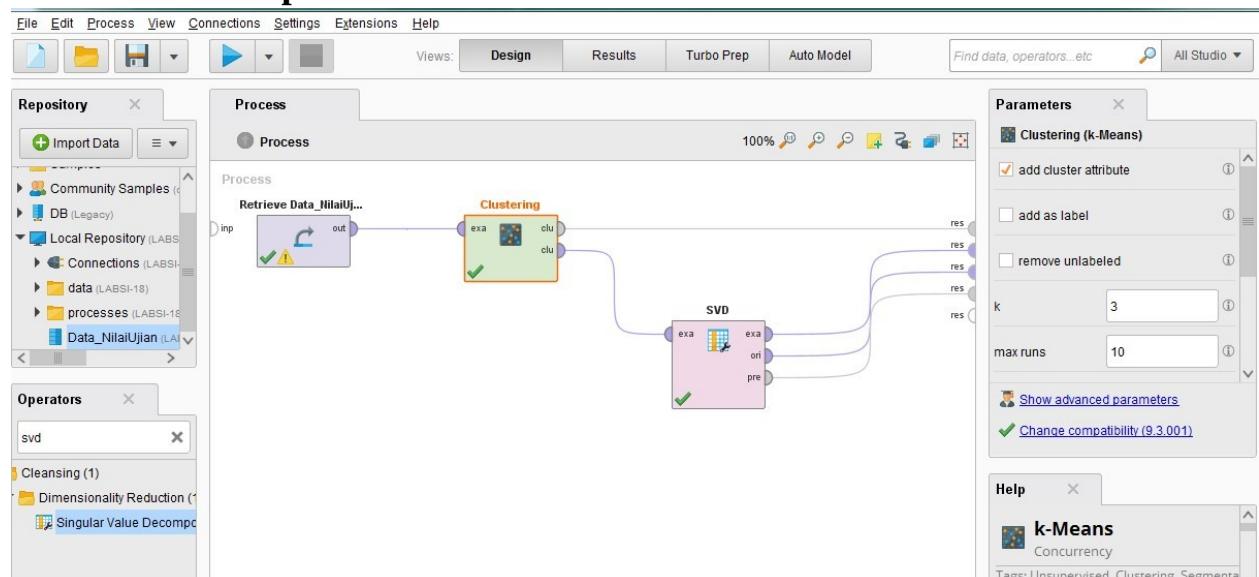
## MODUL 10

LANGKAH-LANGKAH

### Membuat tabel data nilai siswa

	A	B	C	D
1	NO_SISWA	NAMA	B.INDO	B.ING
2	S-101	JOKO	8,54	8,40
3	S-102	AGUS	9,98	6,81
4	S-103	SUSI	6,20	9,15
5	S-104	DYAH	5,24	7,26
6	S-105	WATI	5,70	5,71
7	S-106	IKA	8,57	5,87
8	S-107	EKO	7,70	7,71
9	S-108	YANTO	6,60	5,70
10	S-109	WAWAN	9,00	8,12
11	S-110	MAHMUD	9,81	9,58

### Menambahkan operator K-menas



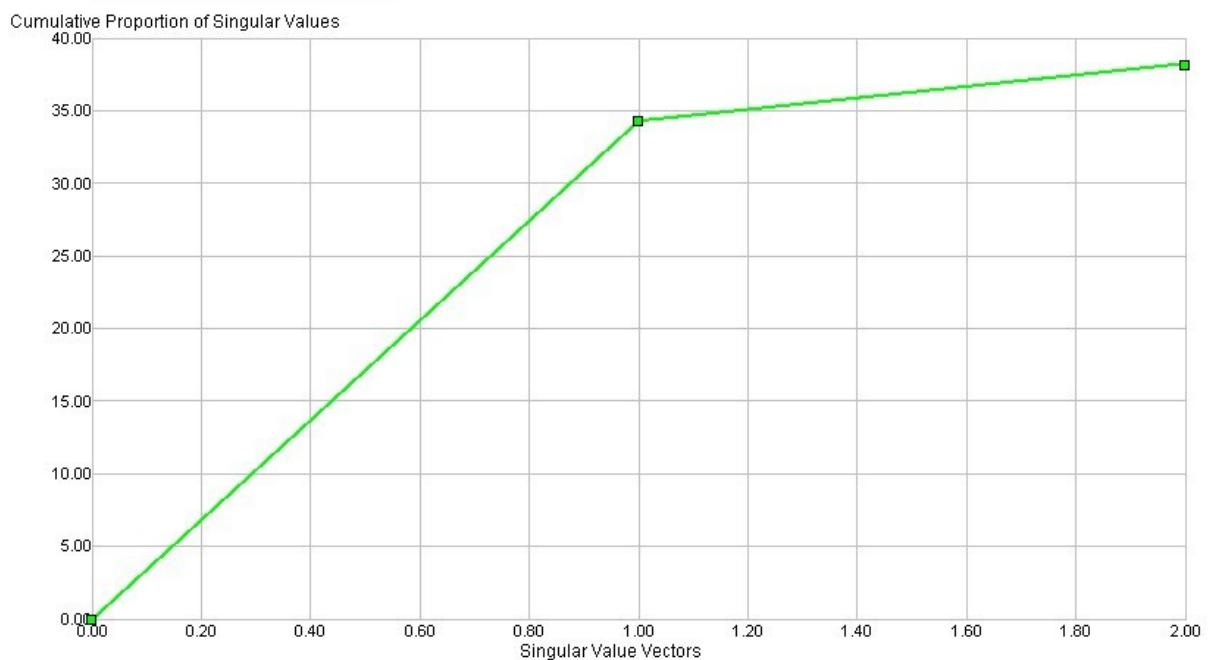
### Nilai eingenvector

Component	Singular Value	Proportion of Singular V...	Cumulative Singular Val...	Cumulative Proportion o...
SVD 1	34.340	0.898	34.340	0.898
SVD 2	3.906	0.102	38.246	1.000

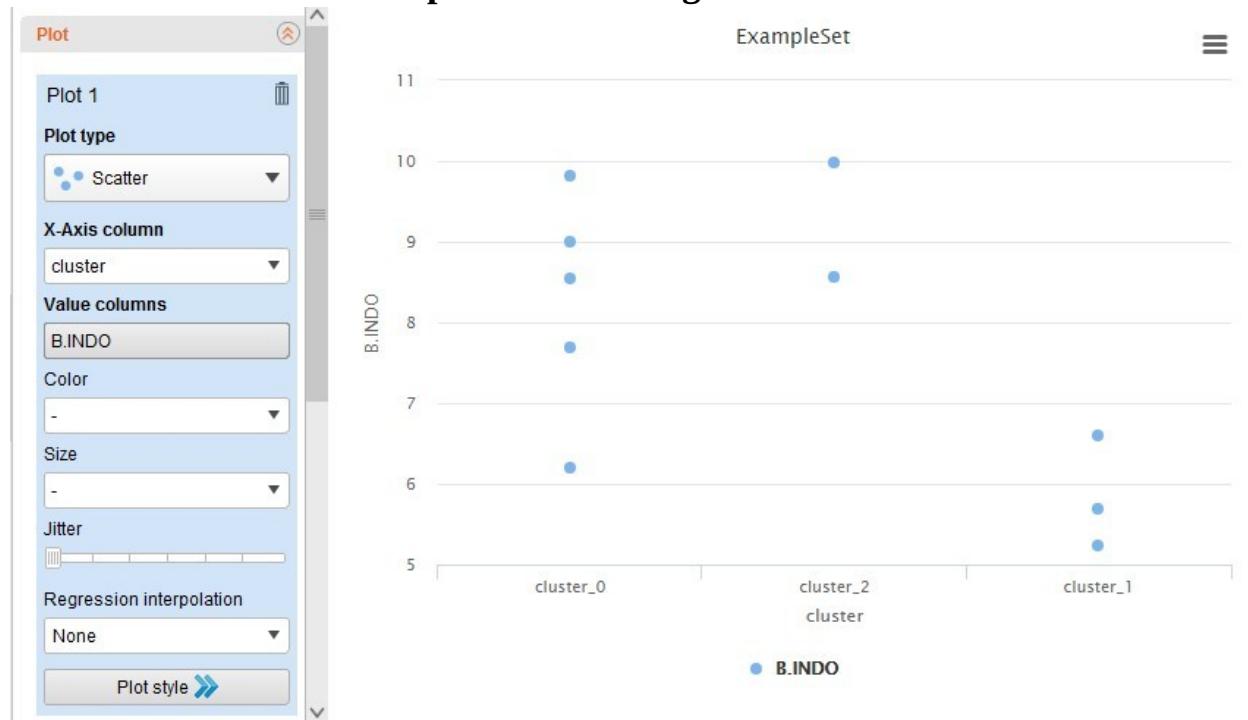
## Nilai SVD vector

Attribute	SVD Vector 1
B.INDO	0.723
B.ING	0.690

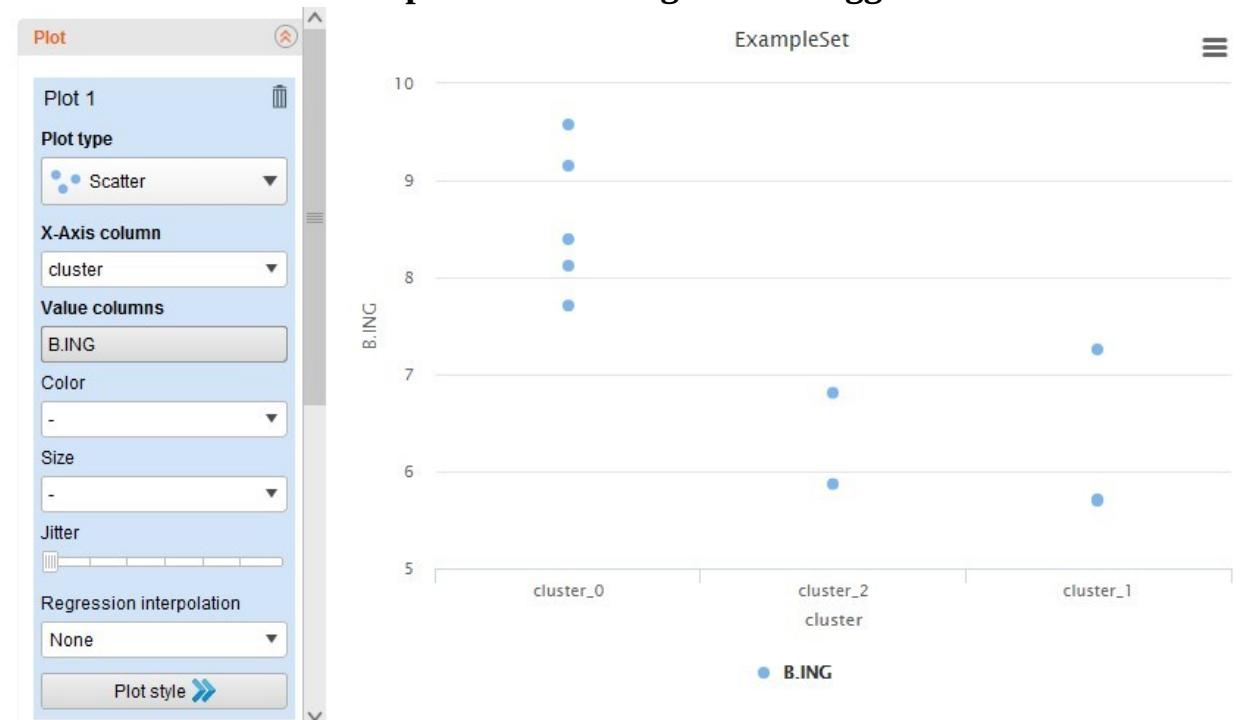
## Nilai cumulative variance

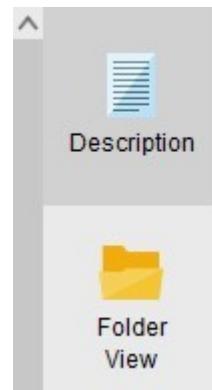


## Kelompok siswa bidang bahasa indo



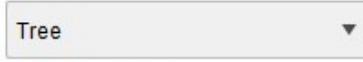
## Kelompok siswa bidang Bahasa inggris





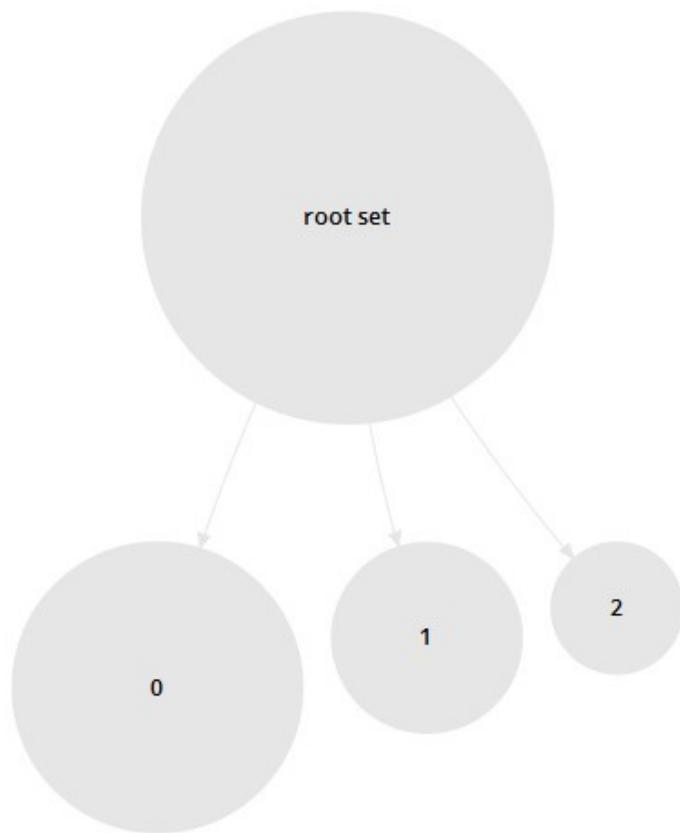
## Cluster Model

Cluster 0: 5 items  
Cluster 1: 3 items  
Cluster 2: 2 items  
Total number of items: 10



Node Labels

Edge Labels

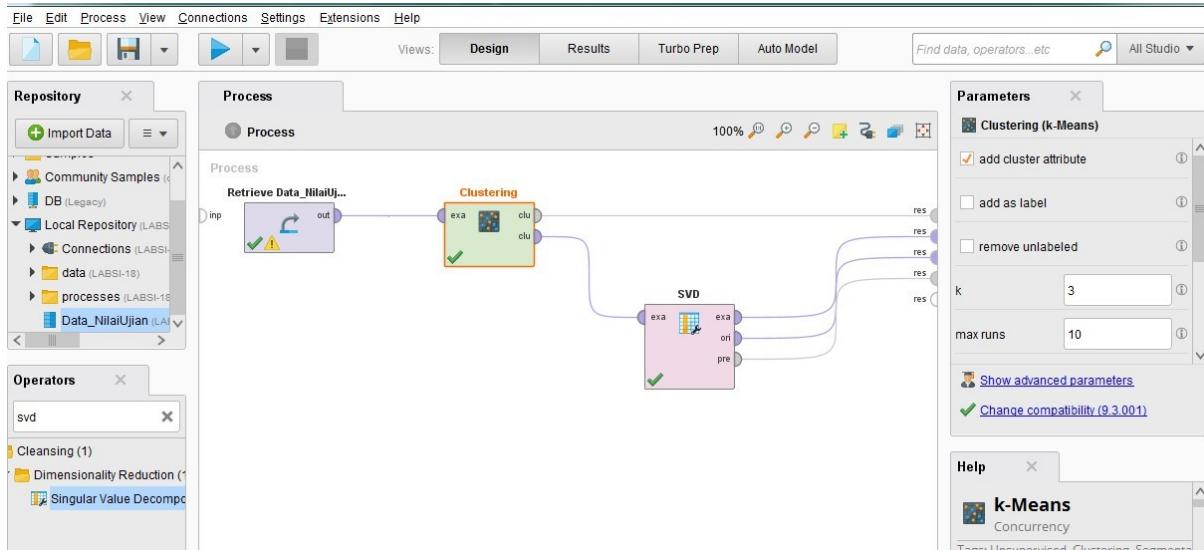


Row No.	NAMA	cluster ↑	B.INDO	B.ING
1	JOKO	cluster_0	8.540	8.400
3	SUSI	cluster_0	6.200	9.150
7	EKO	cluster_0	7.700	7.710
9	WAWAN	cluster_0	9	8.120
10	MAHMUD	cluster_0	9.810	9.580
4	DYAH	cluster_1	5.240	7.260
5	WATI	cluster_1	5.700	5.710
8	YANTO	cluster_1	6.600	5.700
2	AGUS	cluster_2	9.980	6.810
6	IKA	cluster_2	8.570	5.870

1. Cluster 2 yang diajukan untuk lomba olimpiade bidang B.Indonesia
2. Cluster 0 yang diajukan untuk lomba olimpiade bidang B.Inggris

## TUGAS

### Menambah operator k-menas menjdi 4

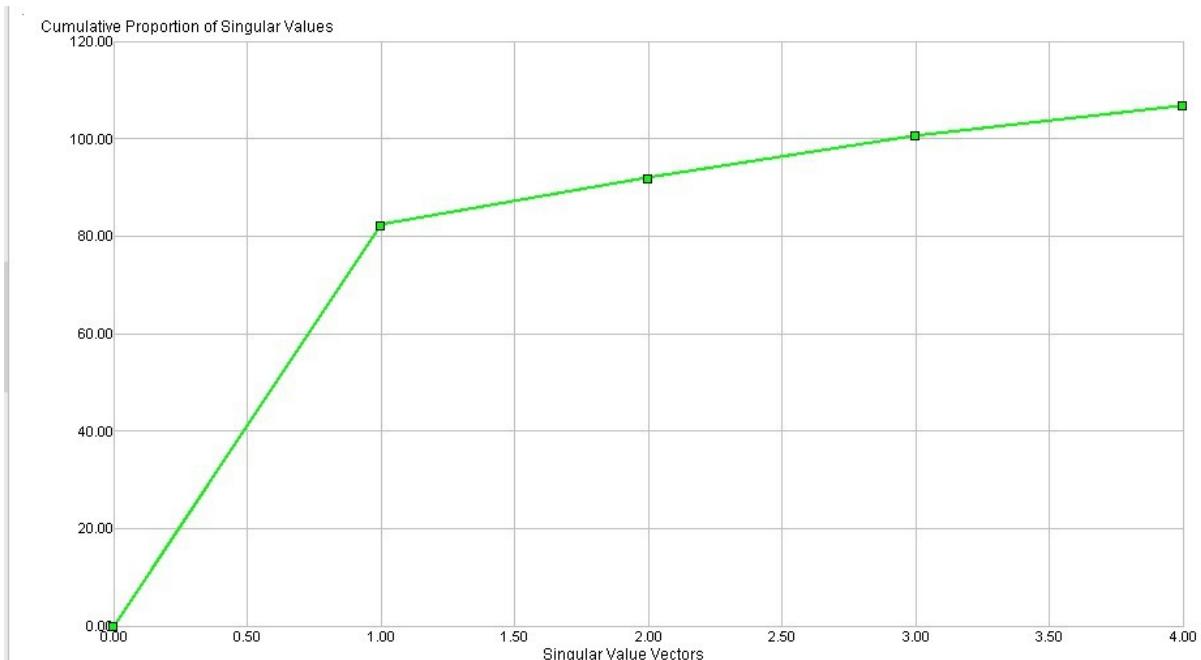


**Svd (singular value decomposition)**

Component	Singular Value	Proportion of Singular Value	Cumulative Singular Values	Cumulative Proportion of S...
SVD 1	82.317	0.770	82.317	0.770
SVD 2	9.634	0.090	91.951	0.860
SVD 3	8.717	0.082	100.668	0.942

Attribute	SVD Vector 1	SVD Vector 2	SVD Vector 3
B.INDO	0.492	0.468	-0.496
B.ING	0.484	-0.251	-0.495
MTK	0.495	0.506	0.639
IPA	0.528	-0.680	0.317

### Nilai cumulative variance



## Example set svd

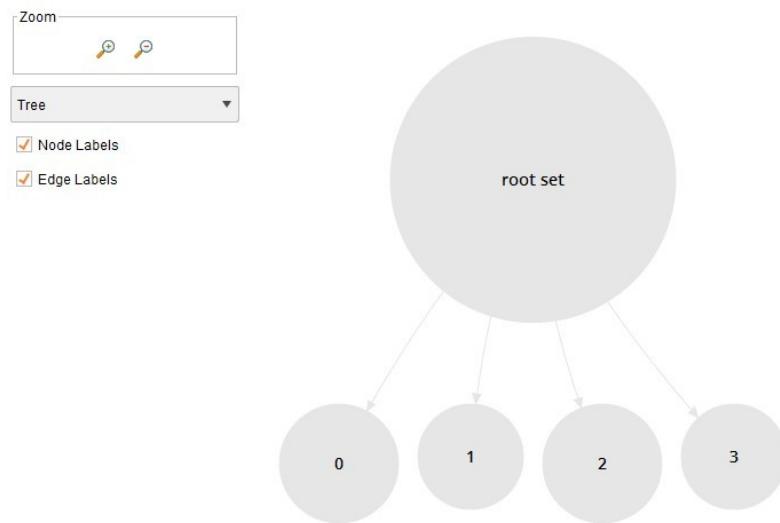
Row No.	NAMA	cluster	svd_1
1	JOKO	cluster_1	0.185
2	AGUS	cluster_2	0.194
3	SUSI	cluster_3	0.184
4	DYAH	cluster_1	0.178
5	WATI	cluster_0	0.182
6	IKA	cluster_3	0.208
7	EKO	cluster_1	0.173
8	YANTO	cluster_0	0.186
9	WAWAN	cluster_2	0.175
10	MAHMUD	cluster_2	0.170
11	BUDI	cluster_3	0.210
12	SANTI	cluster_3	0.205
13	DIAN	cluster_0	0.152
14	DANI	cluster_3	0.197
15	AHMAD	cluster_1	0.155
16	BAYU	cluster_0	0.168
17	RISA	cluster_2	0.173
18	RANI	cluster_3	0.198
19	YANI	cluster_1	0.190
20	RATIH	cluster_2	0.182
21	INDAH	cluster_1	0.210
22	JONO	cluster_2	0.197
23	SARAH	cluster_2	0.160
24	RAMAH	cluster_0	0.167
25	BAMBANG	cluster_3	0.221
26	HANIDI	cluster_0	0.166
27	NANA	cluster_0	0.147
28	FEBRI	cluster_2	0.176
29	DENI	cluster_0	0.168
30	TPNI	cluster_1	0.173

## Cluster model descriptionm

### Cluster Model

Cluster 0: 8 items  
Cluster 1: 7 items  
Cluster 2: 8 items  
Cluster 3: 7 items  
Total number of items: 30

### Graph



## Cluster

Row No.	NAMA	cluster ↑	B.INDO	B.ING	MTK	IPA
5	WATI	cluster_0	9.099	8.430	5.253	7.284
8	YANTO	cluster_0	9.454	8.876	6.046	6.417
13	DIAN	cluster_0	6.110	6.112	5.309	7.470
16	BAYU	cluster_0	8.101	6.982	5.625	6.945
24	RAMAH	cluster_0	8.118	7.696	5.816	6.014
26	HANIDI	cluster_0	5.519	9.937	5.896	6.107
27	NANA	cluster_0	7.376	5.568	5.272	6.068
29	DENI	cluster_0	8.461	6.623	7.636	5.056
1	JOKO	cluster_1	7.741	7.879	6.009	8.832
4	DYAH	cluster_1	6.114	7.376	5.619	9.996
7	EKO	cluster_1	5.007	8.559	5.286	9.461
15	AHMAD	cluster_1	5.045	6.931	5.095	8.340
19	YANI	cluster_1	8.167	7.771	5.801	9.457
21	INDAH	cluster_1	6.511	9.617	8.613	9.846
Row No.	NAMA	cluster ↑	B.INDO	B.ING	MTK	IPA
30	TPNI	cluster_1	5.949	7.400	5.872	9.092
2	AGUS	cluster_2	6.669	7.134	9.987	8.116
9	WAWAN	cluster_2	6.292	5.009	7.505	9.722
10	MAHMUD	cluster_2	6.968	5.193	6.780	8.848
17	RISA	cluster_2	6.060	5.021	9.955	7.388
20	RATIH	cluster_2	6.159	6.402	9.876	7.477
22	JONO	cluster_2	6.022	7.348	9.038	9.894
23	SARAH	cluster_2	6.547	5.055	7.785	6.977
28	FEBRI	cluster_2	5.317	7.609	8.636	7.414
3	SUSI	cluster_3	9.392	6.247	9.684	5.117
6	IKA	cluster_3	9.322	7.839	9.963	7.291
11	BUDI	cluster_3	7.599	8.790	8.560	9.530
12	SANTI	cluster_3	9.979	8.011	8.529	7.275
14	DANI	cluster_3	8.606	8.732	8.681	6.479
18	RANI	cluster_3	9.089	5.996	7.973	9.432
25	BAMBANG	cluster_3	9.772	7.573	9.113	9.894

Nama siswa dalam kelompok cluster 0

WATI YANTO DIAN BAYU RAMAH  
HANIDI NANA DENI

Nama siswa dalam kelompok cluster 1

JOKO DYAH EKO AHMAD YANI  
INDAH TPNI

Nama siswa dalam kelompok cluster 2

AGUS WAWAN MAHMUD RISA  
RATIH JONO SARAH FEBRI

Nama siswa dalam kelompok cluster 3

SUSI IKA BUDI SANTI DANI RANI BAMBANG

# MODUL 11

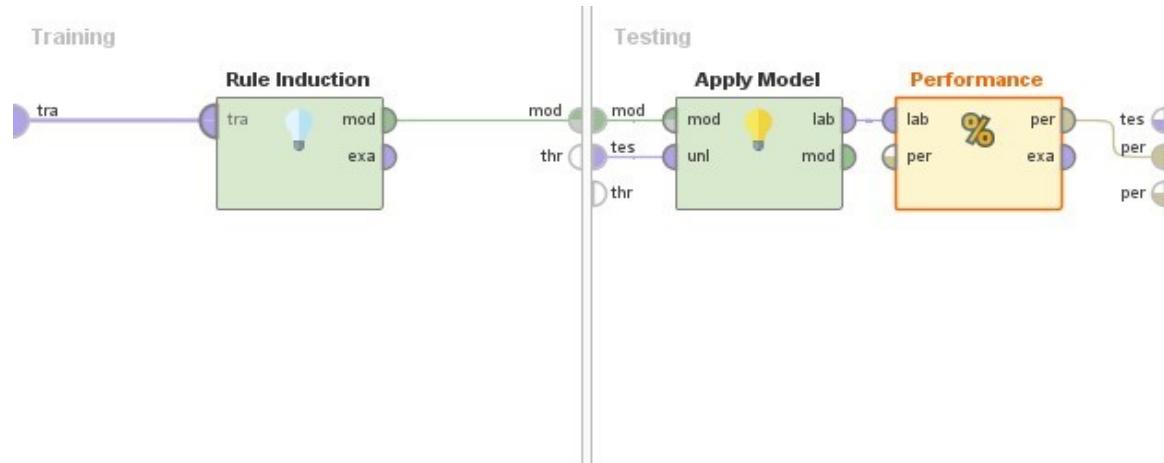
## INDUKSI DAN ATURAN ASOSIASI

### Langkah-langkah praktikum

#### Data table cuaca dalam bentuk tree



#### Proses cross validation



## Rule Model

### RuleModel

```
if Kelembaban_udara ≤ 82.500 then YA (1 / 6)
if Cuaca = Cerah then TIDAK (3 / 0)
if Cuaca = Mendung then YA (0 / 2)
if Suhu ≤ 70.500 then YA (0 / 1)
else TIDAK (0 / 0)

correct: 12 out of 13 training examples.
```

PerformanceVector (Performance)		RuleModel (Rule Induction)	
accuracy: 65.00% +/- 45.00% (micro average: 71.43%)			
	true TIDAK	true YA	class precisi
pred. TIDAK	2	1	66.67%
pred. YA	3	8	72.73%
class recall	40.00%	88.89%	

## 11.4.2 Aturan Asosiasi Data cuaca

Process

inp

Retrieve Tabel\_Cuac...



preprocessing



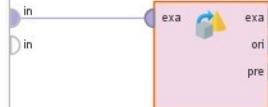
Process

Process > preprocessing

100%

preprocessing

Discretize



Parameters

Discretize (Discretize by Frequency)

attribute filter type	all
<input type="checkbox"/> invert selection	
<input type="checkbox"/> include special attributes	
number of bins	2

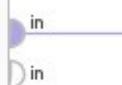
Process

Process > preprocessing >

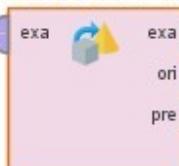
100%

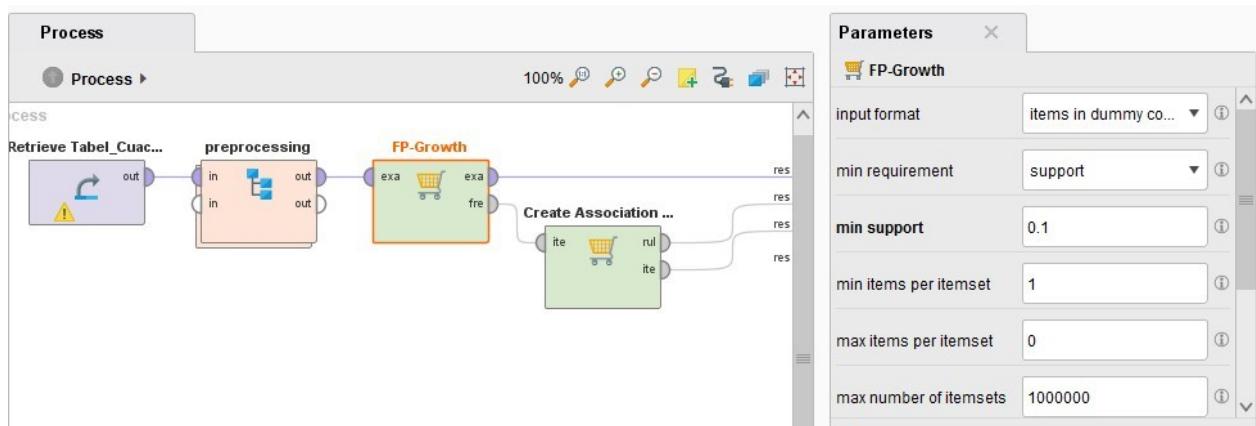
preprocessing

Discretize



Nominal to Binomial





## Frequent item set (FP-Growth)

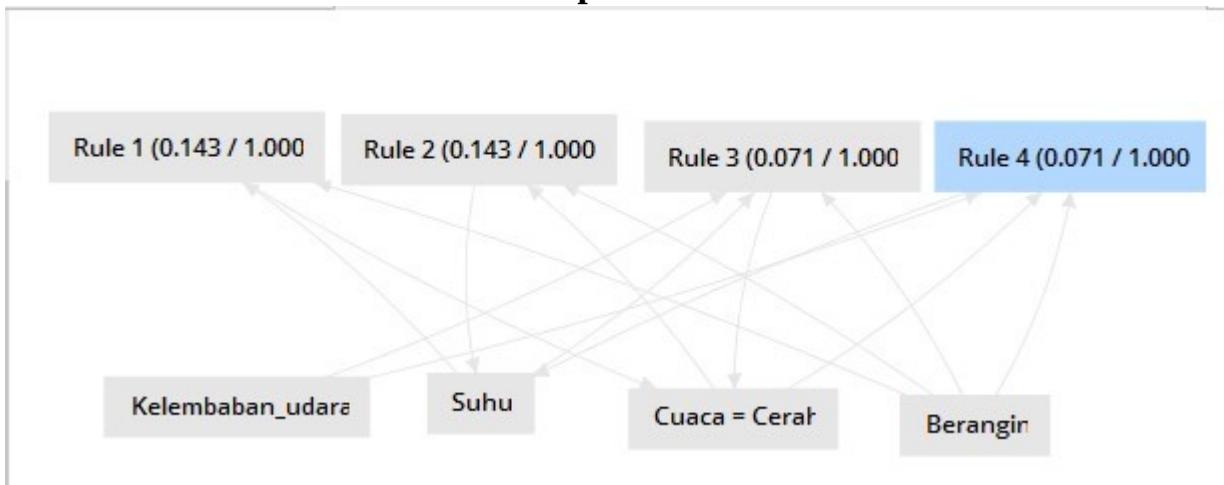
No. of Sets: 26  
Total Max. Size: 4  
Min. Size: 1  
Max. Size: 4  
Contains Item:  
  
Update View

Size	Support	Item 1	Item 2	Item 3	Item 4
2	0.143	Berangin	Cuaca = Hujan		
2	0.143	Berangin	Cuaca = Mendu...		
2	0.214	Suhu	Cuaca = Cerah		
2	0.071	Suhu	Cuaca = Hujan		
2	0.143	Suhu	Cuaca = Mendu...		
3	0.071	Kelembaban_u...	Berangin	Suhu	
3	0.071	Kelembaban_u...	Berangin	Cuaca = Cerah	
3	0.071	Kelembaban_u...	Berangin	Cuaca = Hujan	
3	0.071	Kelembaban_u...	Berangin	Cuaca = Mendu...	
3	0.143	Kelembaban_u...	Suhu	Cuaca = Cerah	
3	0.071	Kelembaban_u...	Suhu	Cuaca = Mendu...	
3	0.143	Berangin	Suhu	Cuaca = Cerah	
4	0.071	Kelembaban_u...	Berangin	Suhu	Cuaca = Cerah

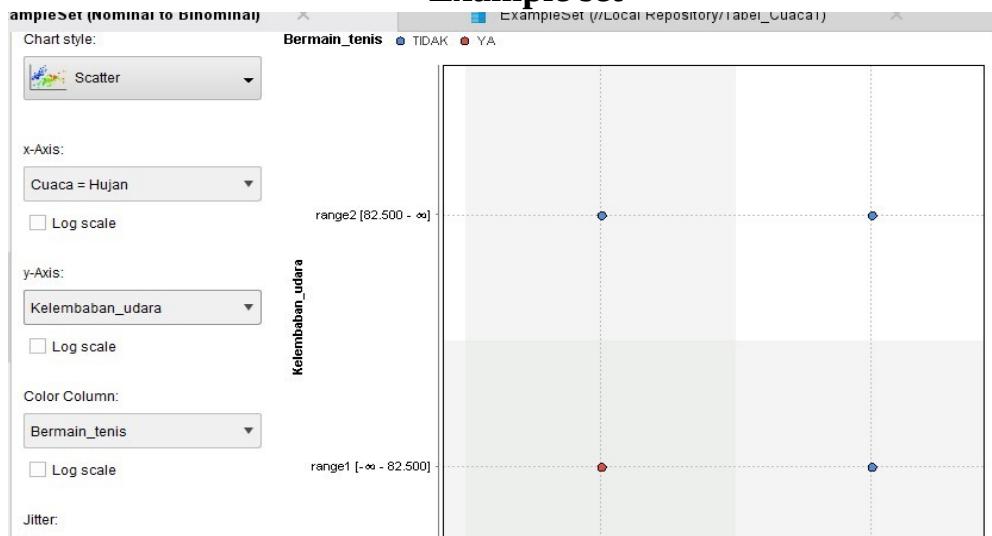
## Asosiation Rules

No.	Premises	Conclusion
1	Berangin, Suhu	Cuaca = Cerah
2	Berangin, Cuaca = Cerah	Suhu
3	Kelembaban_udara, Berangin, Suhu	Cuaca = Cerah
4	Kelembaban_udara, Berangin, Cuaca = Cerah	Suhu

## Graph view



## Example set



## Tugas

### RuleModel

```
if Rerata_SKS > 18.500 then TEPAT (2 / 10)
if Gender = PRIA then TERLAMBAT (4 / 0)
if Jurusan_SMA = IPA then TEPAT (0 / 2)
if Jurusan_SMA = IPS then TERLAMBAT (1 / 0)
else TEPAT (0 / 0)
```

correct: 17 out of 19 training examples.

accuracy: 65.00% +/- 32.02% (micro average: 65.00%)

	true TERLAMBAT	true TEPAT	class precision
pred. TERLAMBAT	4	4	50.00%
pred. TEPAT	3	9	75.00%
class recall	57.14%	69.23%	

- a) Number of bins = 2
- b) Number of bins = 3

## FP-Growth (table view)

Number of bins = 2

No. of Sets: 15 Total Max. Size: 4	Size	Support	Item 1	Item 2	Item 3	Item 4
Min. Size: <input type="text" value="1"/>	1	0.750	Gender			
Max. Size: <input type="text" value="4"/>	1	0.300	Asal_Sekolah			
Contains Item:	1	0.250	Asisten			
<input type="text"/>	1	0.250	Rerata_SKS			
<input type="button" value="Update View"/>	2	0.250	Gender	Asal_Sekolah		
	2	0.200	Gender	Asisten		
	2	0.250	Gender	Rerata_SKS		
	2	0.150	Asal_Sekolah	Asisten		
	2	0.150	Asal_Sekolah	Rerata_SKS		
	2	0.150	Asisten	Rerata_SKS		
	3	0.100	Gender	Asal_Sekolah	Asisten	
	3	0.150	Gender	Asal_Sekolah	Rerata_SKS	
	3	0.150	Gender	Asisten	Rerata_SKS	
	3	0.100	Asal_Sekolah	Asisten	Rerata_SKS	

Number of bins = 3

No. of Sets: 7 Total Max. Size: 3	Size	Support	Item 1	Item 2	Item 3
Min. Size: <input type="text" value="1"/>	1	0.750	Gender		
Max. Size: <input type="text" value="3"/>	1	0.300	Asal_Sekolah		
Contains Item:	1	0.250	Asisten		
<input type="text"/>	2	0.250	Gender	Asal_Sekolah	
<input type="button" value="Update View"/>	2	0.200	Gender	Asisten	
	2	0.150	Asal_Sekolah	Asisten	
	3	0.100	Gender	Asal_Sekolah	Asisten

## Association Rules

Number of bins 2

No.	Premises	Conclusion
2	Asal_Sekolah	Gender
3	Rerata_SKS	Gender
4	Asal_Sekolah, Rerata_SKS	Gender
5	Asisten, Rerata_SKS	Gender
6	Gender, Asal_Sekolah, Asisten	Rerata_SKS
7	Asal_Sekolah, Asisten, Rerata_SKS	Gender

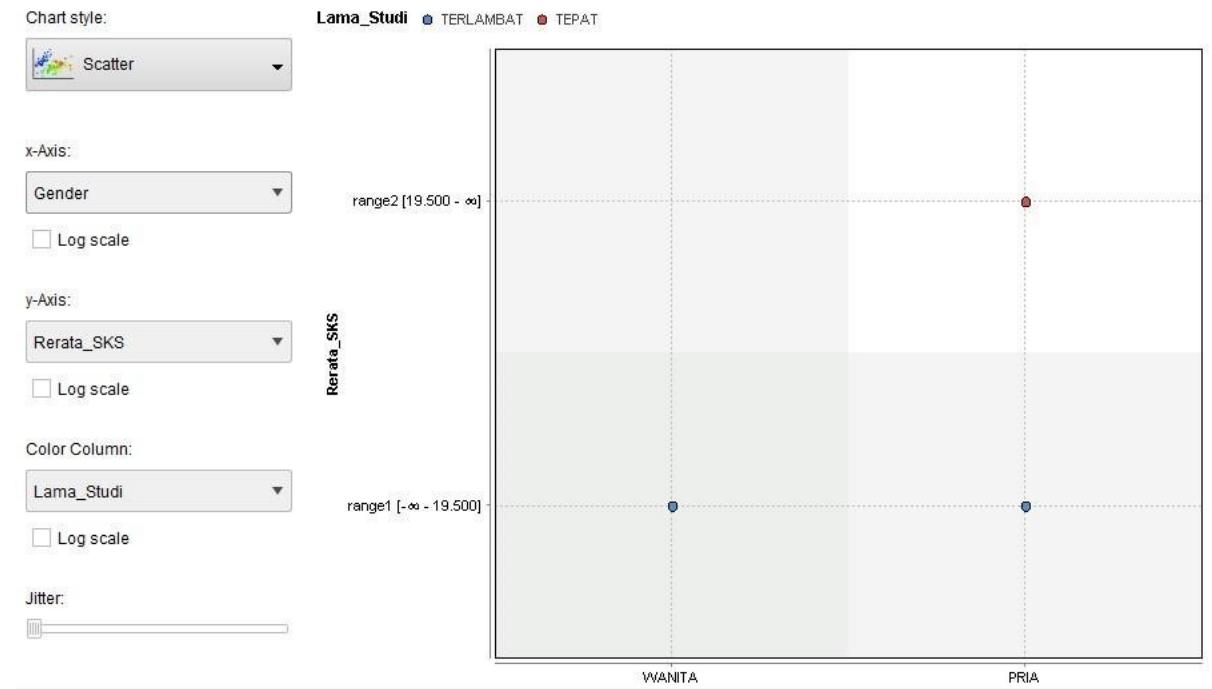
Support	Confidence	LaPlace	Gain	p-s	Lift	Convicti...
0.250	0.833	0.962	-0.350	0.025	1.111	1.500
0.250	1	1	-0.250	0.062	1.333	$\infty$
0.150	1	1	-0.150	0.038	1.333	$\infty$
0.150	1	1	-0.150	0.038	1.333	$\infty$
0.100	1	1	-0.100	0.075	4	$\infty$
0.100	1	1	-0.100	0.025	1.333	$\infty$

Number of bins 3

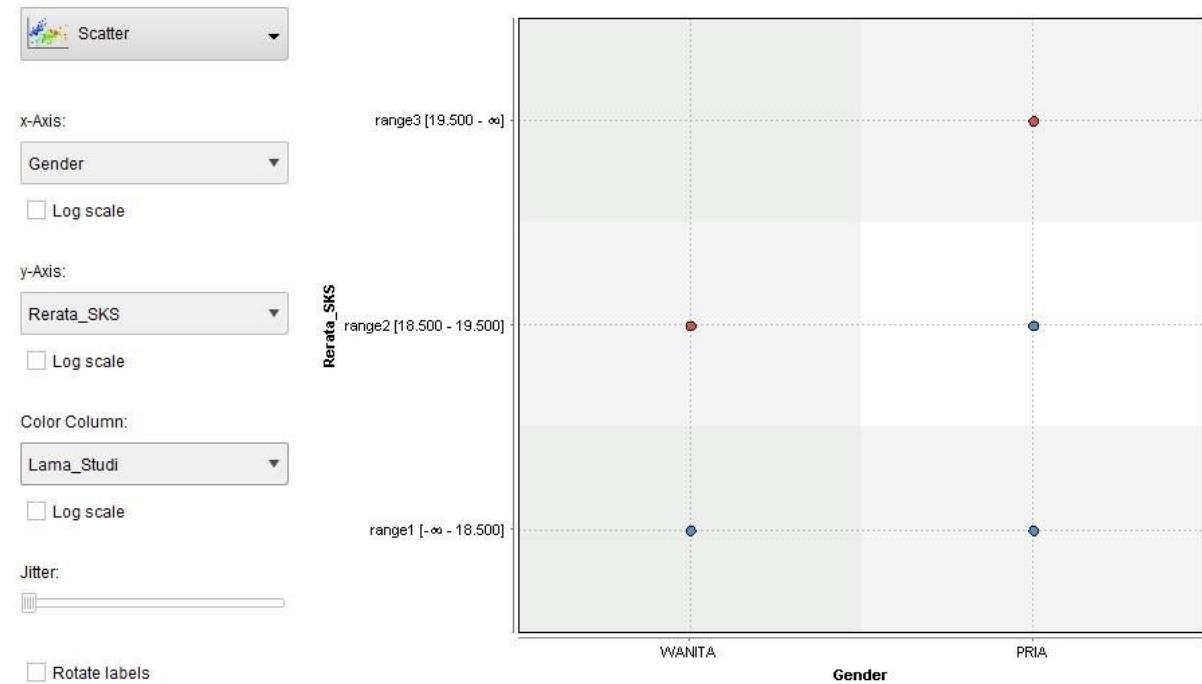
No.	Premises	Conclusion				
2	Asal_Sekolah	Gender				
Support	Confidence	LaPlace	Gain	p-s	Lift	Convicti...
0.250	0.833	0.962	-0.350	0.025	1.111	1.500

## ExampleSet

### Number of bins 2



### Number of bins 3



## MODUL 12

### REGRESI LINIER SEDERHANA

#### 12.4 Langkah-langkah Praktikum

##### Contoh Kasus :

Dalam sebuah kelas yang memiliki 10 siswa dilakukan sebuah survei terhadap lama belajar seorang siswa dan nilai hasil ujiannya. Data siswa tersebut akan kita gunakan sebagai dasar perhitungan untuk memprediksi nilai ujian terhadap siswa lain berdasarkan lama belajarnya.

##### 12.4.1 Induksi Aturan Data Cuaca Hipotesis :

Bagaimana

1. Membuat data pada Microsoft Excel dengan nama **Tabel\_LamaBelajardanNilaiUjian.xls**.

No_Siswa	Nama	Lama Belajar (Jam)	Nilai
S-101	JOKO	15	783
S-102	AGUS	18	877
S-103	SUSI	7	505
S-104	DYAH	9	860
S-105	WATI	15	968
S-106	IKA	17	793
S-107	EKO	10	752
S-108	YANTO	5	571
S-109	WAWAN	8	667
S-110	MAHMUD	15	723

2. Mengimport data table **Tabel\_LamaBelajardanNilaiUjian.xls** ke **RapidMiner** yang akan digunakan sebagai regresi Linear, dengan memberikan pengaturan tipe dan jenis data pada atribut,
  - NO\_SISWA : polynomial, id
  - NAMA : pilih Exclude column
  - LAMA JAM BELAJAR : integer
  - NILAI : integer, label

Simpan dengan nama **Data\_LamaBelajardanNilaiUjian**.

Import Data - Format your columns.

**Format your columns.**

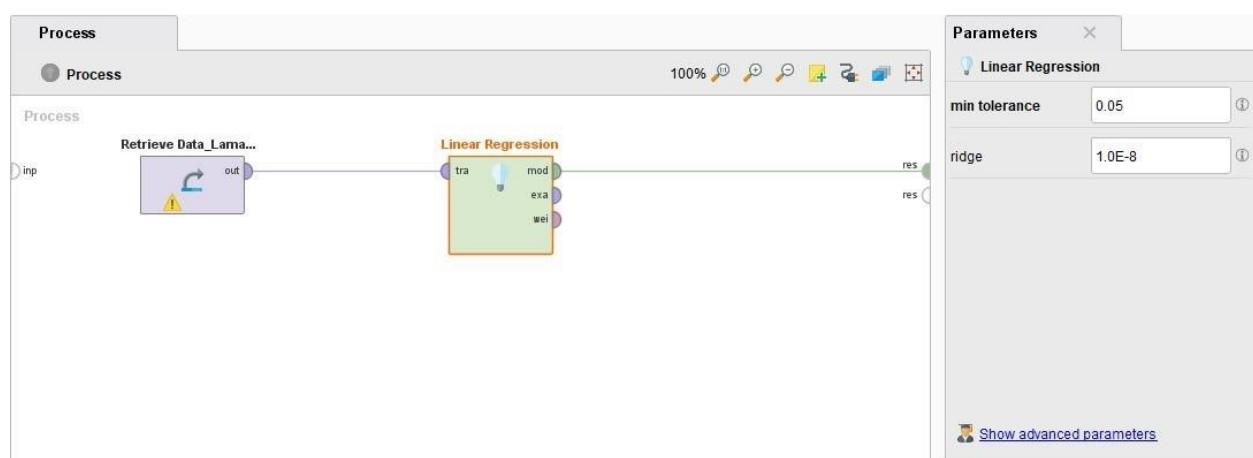
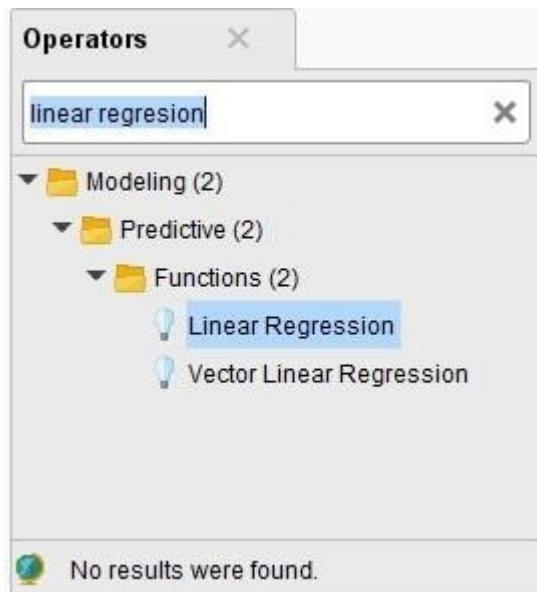
Replace errors with missing values  ⓘ

No_Siswa <i>polynomial id</i>	Nama <i>polynomial</i>	Lama Belajar (Jam) <i>integer</i>	Nilai <i>integer label</i>
1	S-101	JOKO	15
2	S-102	AGUS	18
3	S-103	SUSI	7
4	S-104	DYAH	9
5	S-105	WATI	15
6	S-106	IKA	17
7	S-107	EKO	10
8	S-108	YANTO	5
9	S-109	WAWAN	8
10	S-110	MAHMUD	15

 no problems.

 Previous  Next  Cancel

3. Masukkan **Data\_LamaBelajardanNilaiUjian** ke dalam area process, dan tambahkan **Linier Regresion**, lalu klik pada operator **Linear Regresion** dan tentukan parameter **min tolerance = 0.05** (Batas toleransi sebesar 5%).



4. Jalankan proses dengan meng-klik **Run**.

- Table View (Mencari besarnya nilai t-hitung)

Attribute	Coefficient	Std. Error	Std. Coefficient	Tolerance	t-Stat	p-Value	Code
Lama Belajar (J...	21.608	7.645	0.707	1	2.827	0.022	**
(Intercept)	492.769	96.909	?	?	5.085	0.001	****

- Text View (Mencari model regresi)

## LinearRegression

```
21.608 * Lama Belajar (Jam)
+ 492.769
```

#### **12.4.2 Mencari Nilai t dan Model Regresi Linier Menggunakan RapidMiner**

##### **Hipotesis :**

Bagaimana memprediksi nilai ujian siswa berdasarkan lama belajarnya menggunakan model regresi linier yang telah dihasilkan menggunakan RapidMiner.

1. Membuat table data baru di **Tabel\_LamaBelajardanNilaiUjian.xls** Pada Sheet 2.

No_Siswa	Nama	Lama Belajar (Jam)
S-111	BUDI	12
S-112	SANTI	13
S-113	DIAN	14
S-114	DANI	11
S-115	AHMAD	5
S-116	BAYU	13
S-117	RISA	9
S-118	RANI	10
S-119	YANI	10
S-120	RATIH	9

2. Mengimport data tersebut ke **RapidMiner** yang akan digunakan sebagai regresi Linear (Sama seperti percobaan **12.4.1**), dengan memberikan pengaturan tipe dan jenis data pada atribut,

- NO\_SISWA : polynomial, id
- NAMA : pilih Exclude column
- LAMA JAM BELAJAR : integer

Simpan dengan nama **Data\_PrediksiNilaiUjian**.

Import Data - Format your columns.

**Format your columns.**

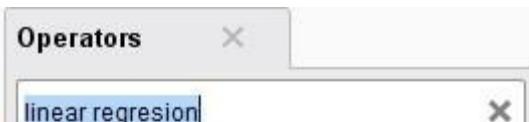
Replace errors with missing values  ⓘ

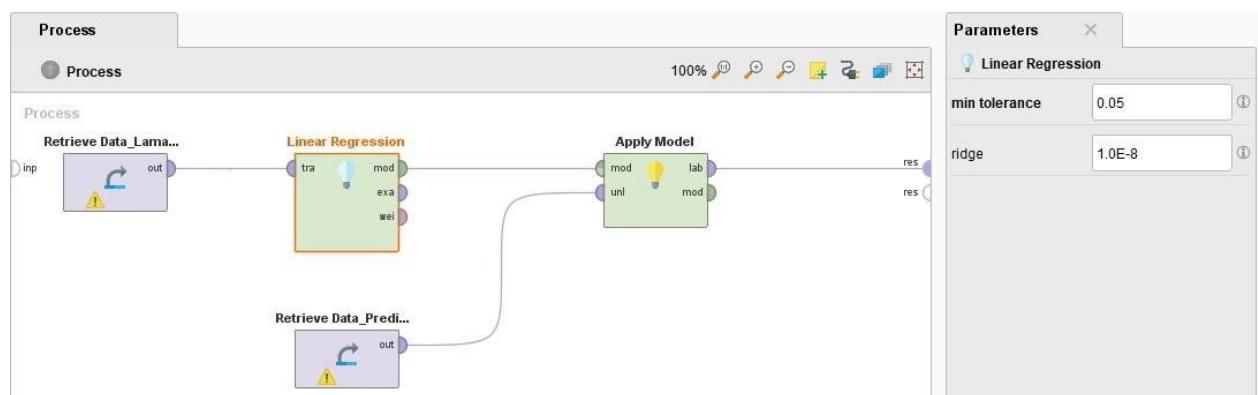
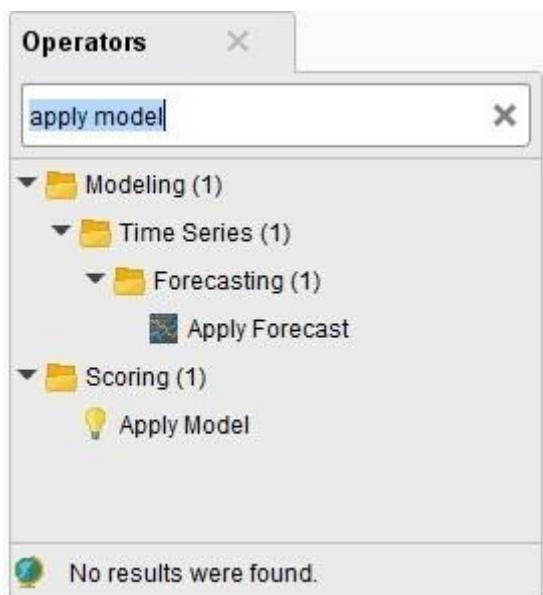
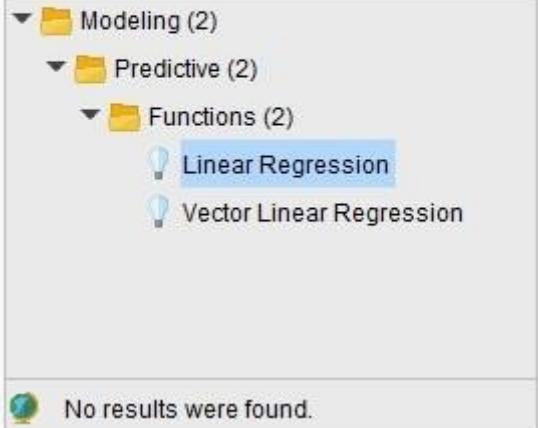
No_Siswa <small>polynominal id</small>	Nama <small>polynominal</small>	Lama Belajar (Jam) <small>integer</small>
1 S-111	BUDI	12
2 S-112	SANTI	13
3 S-113	DIAN	14
4 S-114	DANI	11
5 S-115	AHMAD	5
6 S-116	BAYU	13
7 S-117	RISA	9
8 S-118	RANI	10
9 S-119	YANI	10
10 S-120	RATIH	9

 no problems.

 [Previous](#)  [Next](#)  [Cancel](#)

3. Masukkan **Data\_PrediksiNilaiUjian** ke dalam area process, dan tambahkan **Linier Regresion** dan **Apply Model**, lalu klik pada operator Linear Regresion dan tentukan parameter **min tolerance = 0.05** (Batas toleransi sebesar 5%).



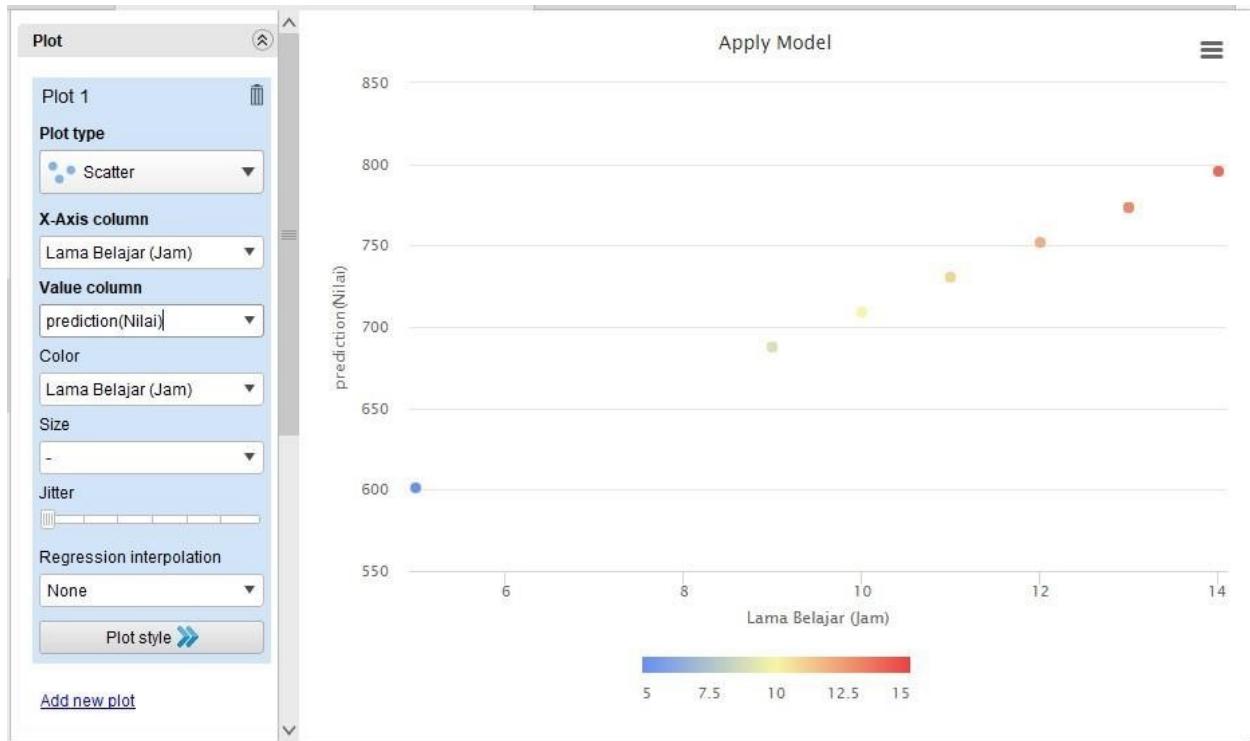


4. Jalankan proses dengan meng-klik **Run**.

- Data View (hasil Prediksi nilai ujian)

Row No.	No_Siswa	prediction(Nilai)	Lama Belaja...
1	S-111	752.061	12
2	S-112	773.668	13
3	S-113	795.276	14
4	S-114	730.453	11
5	S-115	600.807	5
6	S-116	773.668	13
7	S-117	687.238	9
8	S-118	708.845	10
9	S-119	708.845	10
10	S-120	687.238	9

- Chart View (Scatter Plot)



### 12.4.3 Pembuktian Model Regresi

Melakukan pembuktian hasil prediksi menggunakan **RapidMiner** dengan hasil prediksi menggunakan model regresi yang telah dihasilkan pada kegiatan 12.4.1

$$Y = 21,608X_1 + 492,769$$

Memasukkan nilai variable  $X_1$  ke dalam model regresi :

1. No. SISWA = S-112, NAMA = SANTI, memiliki nilai  $X_1 = 13$  Jam. Sehingga,  $Y = (21,608 \times 13) + 492,769 = 773,672$
2. No. SISWA = S-113, NAMA = DIAN, memiliki nilai  $X_1 = 14$  Jam. Sehingga,  $Y = (21,608 \times 14) + 492,769 = 795,28$
3. No. SISWA = S-114, NAMA = DANI, memiliki nilai  $X_1 = 11$  Jam. Sehingga,  $Y = (21,608 \times 11) + 492,769 = 730,456$
4. No. SISWA = S-116, NAMA = BAYU, memiliki nilai  $X_1 = 13$  Jam. Sehingga,  $Y = (21,608 \times 13) + 492,769 = 773,672$
5. No. SISWA = S-117, NAMA = RISA, memiliki nilai  $X_1 = 9$  Jam.

Sehingga,  $Y = (21,608 \times 9) + 492,769 = 687,24$

6. No. SISWA = S-118, NAMA = RANI, memiliki nilai  $X_1 = 10$  Jam.

Sehingga,  $Y = (21,608 \times 10) + 492,769 = 708,848$

7. No. SISWA = S-119, NAMA = YANI, memiliki nilai  $X_1 = 10$  Jam.

Sehingga,  $Y = (21,608 \times 10) + 492,769 = 708,848$

No_Siswa	Nama	Lama Belajar (Jam)		
S-111	BUDI	12		752,064
S-112	SANTI	13		773,672
S-113	DIAN	14		795,28
S-114	DANI	11		730,456
S-115	AHMAD	5		600,808
S-116	BAYU	13		773,672
S-117	RISA	9		687,24
S-118	RANI	10		708,848
S-119	YANI	10		708,848
S-120	RATIH	9		687,24

## 12.5 Tugas Kasus

:

Dalam sebuah survei terhadap 15 kepala keluarga telah diperoleh variable pendapatan rata-rata perbulan, jumlah anggota keluarga yang tinggal serumah, dan daya beli rata-rata perbulan.

### Hipotesis :

Bagaimanaka model regresi linier yang terbentuk, dan lakukan prediksi terhadap 10 data yang belum diketahui nilai daya belinya.

1. Buatlah table berikut dengan menggunakan Microsoft Excel.

NO. RESPONDEN	PENDAPATAN (RUPIAH)	JUMLAH ANGGOTA KELUARGA	DAYA BELI (RUPIAH)
1	1000000	6	834000
2	1400000	7	1200000
3	200000	3	134000
4	1400000	6	1167000
5	500000	3	334000
6	1700000	5	1360000
7	400000	3	267000
8	1900000	5	1520000
9	300000	3	200000
10	500000	4	375000
11	700000	7	600000
12	1900000	3	1267000
13	800000	4	600000
14	1500000	4	1125000
15	1300000	7	1115000

2. Buatlah proses Regresi Linier Sederhana menggunakan RapidMiner dengan ketentuan sebagai berikut.

- a) Variable bebas (X) = Pendapatan (X<sub>1</sub>), Jumlah Anggota Keluarga (X<sub>2</sub>)
- b) Variable terikat (Y) = Daya Beli
- c) Toleransi yang digunakan = 5%

Pendapatan (X<sub>1</sub>)

Import Data - Format your columns.

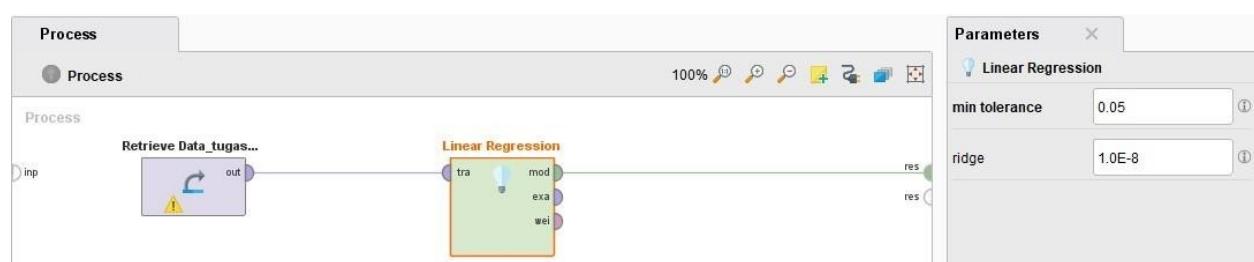
Format your columns.

Replace errors with missing values  ⓘ

NO. RESPONDEN <small>polynomial id</small>	PENDAPATAN (RUPIAH) <small>integer</small>	JUMLAH ANGGOTA K... <small>integer</small>	DAYA BELI (RUPIAH) <small>integer label</small>
1	1000000	6	834000
2	1400000	7	1200000
3	200000	3	134000
4	1400000	6	1167000
5	500000	3	334000
6	1700000	5	1360000
7	400000	3	267000
8	1900000	5	1520000
9	300000	3	200000
10	500000	4	375000
11	700000	7	600000
12	1900000	3	1267000
13	800000	4	600000

✓ no problems.

← Previous → Next ✖ Cancel



Attribute	Coefficient	Std. Error	Std. Coefficient	Tolerance	t-Stat	p-Value	Code
PENDAPATAN (RUPIAH)	0.788	0.038	0.985	1	20.731	0.000	***
(Intercept)	-7467.579	44843.156	?	?	-0.167	0.870	

## LinearRegression

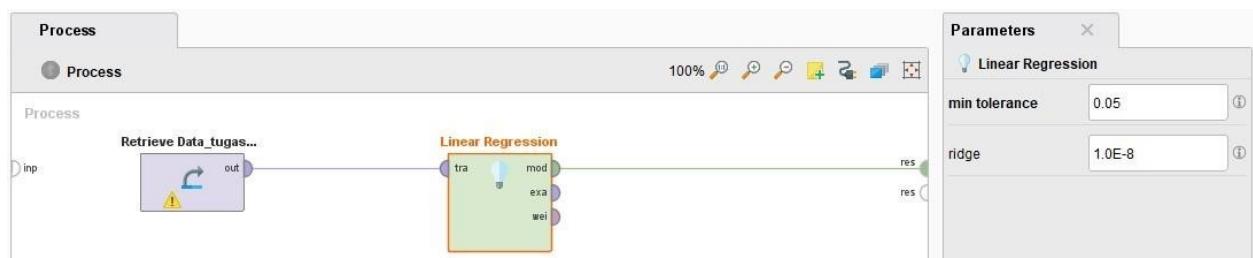
0.788 \* PENDAPATAN (RUPIAH)  
 - 7467.579

Jumlah Anggota Keluarga (X<sub>2</sub>)

	NU. RESPONDEN polynomial id	PENDAPATAN (RUPIAH) integer	JUMLAH ANGGOTA K... integer	DATA BELI (RUPIAH) integer label
1	1	1000000	6	834000
2	2	1400000	7	1200000
3	3	200000	3	134000
4	4	1400000	6	1167000
5	5	500000	3	334000
6	6	1700000	5	1360000
7	7	400000	3	267000
8	8	1900000	5	1520000
9	9	300000	3	200000
10	10	500000	4	375000
11	11	700000	7	600000
12	12	1900000	3	1267000
13	13	800000	4	600000

✓ no problems.

← Previous → Next ✖ Cancel



Attribute	Coefficient	Std. Error	Std. Coefficient	Tolerance	t-Stat	p-Value	Code
JUMLAH ANGG...	151688.679	70801.481	0.511	?	2.142	0.052	*
(Intercept)	98652.830	347817.155	?	?	0.284	0.781	

## Linear Regression

$$151688.679 * \text{JUMLAH ANGGOTA KELUARGA} \\ + 98652.830$$

3. Tentukan apakah variable  $X_1$  dan  $X_2$  mempengaruhi secara signifikan terhadap nilai variable Y berdasarkan besarnya nilai t-stat?

➤ Semisal ambil contoh data **NO. RESPONDEN** = 1 :

$$X_1 \Rightarrow Y = 0,788 X 1.000.000 - 7.463,579 = 780.536,421$$

$$X_2 \Rightarrow Y = 151.688,679 X 6 + 98.652,830 = 811.479,244$$

Jadi, melihat dari hasil yang di dapat variable  $X_1$  dan  $X_2$  memang mempengaruhi secara signifikan terhadap nilai Y.

4. Tuliskan model persamaan regresi linier sederhana yang terbentuk!

➤  $X_1 \Rightarrow Y = 0,788 X \text{ Pendapatan (Rupiah)} - 7.463,579$   
 ➤  $X_2 \Rightarrow Y = 151.688,679 X \text{ Jumlah Anggota Keluarga} + 98.652,830$

5. Gunakan data testing untuk menjawab perintah berikut :

NO. RESPONDEN	PENDAPATAN (RUPIAH)	JUMLAH ANGGOTA KELUARGA
1	900000	5
2	800000	3
3	500000	2
4	1900000	6
5	600000	2
6	800000	5
7	1000000	6
8	1100000	4
9	1000000	4
10	500000	3

- a) Lakukan prediksi Daya Beli (Y) dengan menggunakan Model Persamaan Regresi Linier dari hasil pertanyaan nomor 4!

NO. RESPONDEN	PENDAPATAN (RUPIAH)	JUMLAH ANGGOTA KELUARGA	X1	X2
1	900000	5	701731,421	857096,225
2	800000	3	622931,421	553718,867
3	500000	2	386531,421	402030,188
4	1900000	6	1489731,421	1008784,904
5	600000	2	465331,421	402030,188
6	800000	5	622931,421	857096,225
7	1000000	6	780531,421	1008784,904
8	1100000	4	859331,421	705407,546
9	1000000	4	780531,421	705407,546
10	500000	3	386531,421	553718,867

b) Lakukan prediksi Daya Beli (Y) menggunakan RapidMiner!

Pendapatan (X<sub>1</sub>)

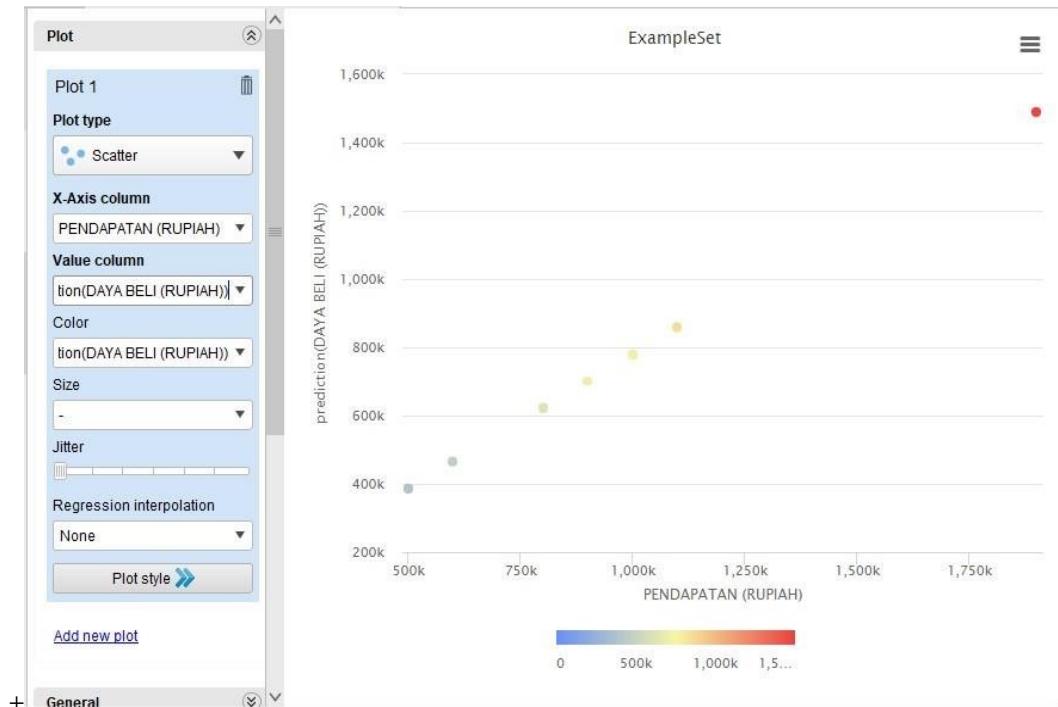
Row No.	NO. RESPON...	prediction(D...	PENDAPATA...
1	1	701500.958	900000
2	2	622726.676	800000
3	3	386403.830	500000
4	4	1489243.776	1900000
5	5	465178.112	600000
6	6	622726.676	800000
7	7	780275.239	1000000
8	8	859049.521	1100000
9	9	780275.239	1000000
10	10	386403.830	500000

Jumlah Anggota Keluarga (X<sub>2</sub>)

Row No.	NO. RESPON...	prediction(D...	JUMLAH AN...
1	1	857096.226	5
2	2	553718.868	3
3	3	402030.189	2
4	4	1008784.906	6
5	5	402030.189	2
6	6	857096.226	5
7	7	1008784.906	6
8	8	705407.547	4
9	9	705407.547	4
10	10	553718.868	3

6. Gambarkan pola sebaran data menggunakan Plot View (Scatter) dengan ketentuan berikut :

- a) x-Axis = Pendapatan (Rupiah), y-Axis = Prediction (Daya Beli (Rupiah)), Color Column = Prediction (Daya Beli (Rupiah))



- b) x-Axis = Jumlah Anggota Keluarga, y-Axis = Prediction (Daya Beli (Rupiah)), Color Column = Prediction (Daya Beli (Rupiah))

