

NAMA : SRI HAJIATI

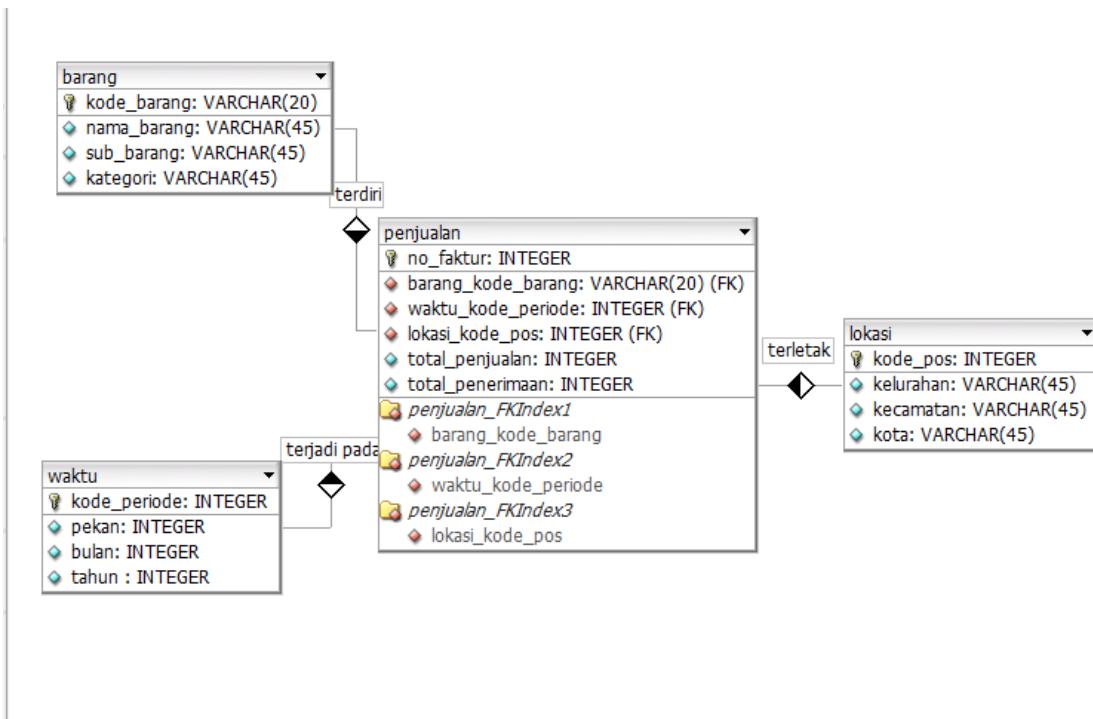
NIM : L200170103

KELAS : D

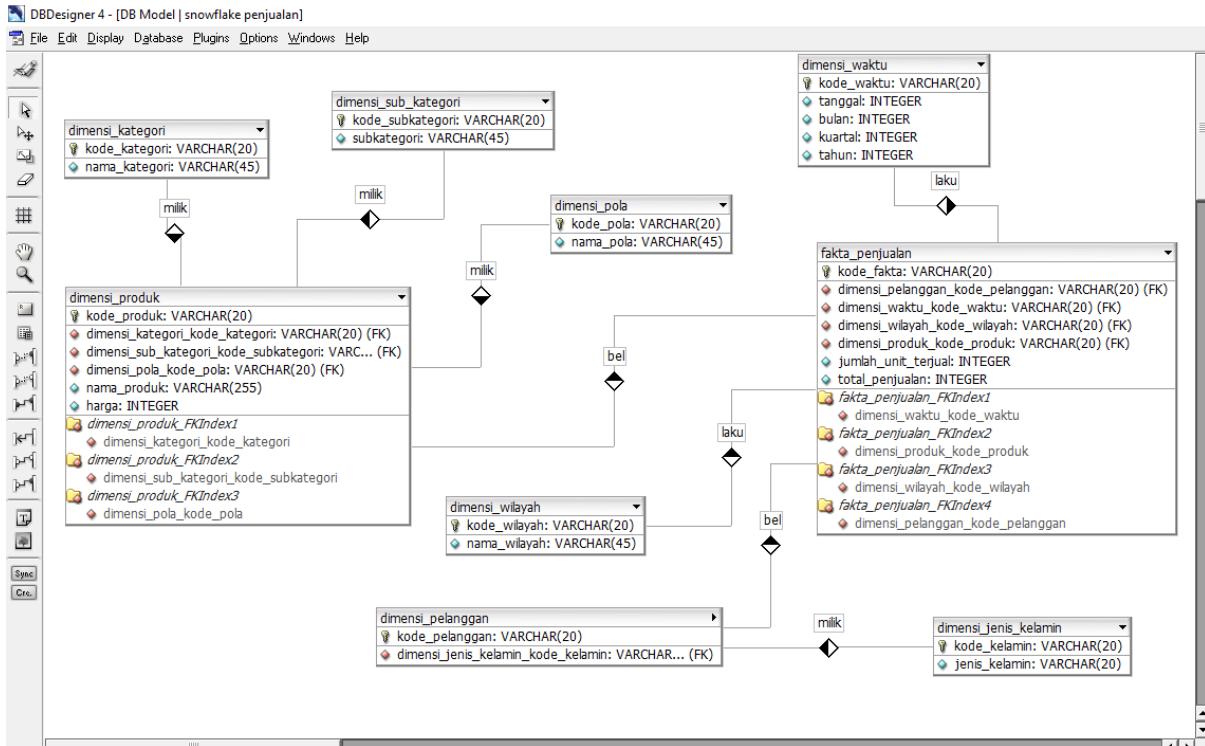
## LAPORAN MODUL PRAKTIKUM DWDM

### MODUL 1

1. star schema penjualan.xml



2. snowflake penjualan.xml



## MODUL 5

### Kegiatan 1

a. Membuat Pivot Baru dengan nama file fakta\_penjualan.xls

bulan	kuartal	tahun	nama_produk	nama_kategori	nama_subkategori	nama_pola	nama_pelanggan	jenis_kelamin	nama_wilayah	jumlah	harga
12	4	2011	Jarik Standar Print Sogar	Standar	Jarik	Print	Bapak Ketut	PRIA	Bali	2	225000
1	1	2012	Kaos Batik Cap Lukis	Batik	Kaos	Cap	Ibu Harini	WANITA	Jawa Timur	14	30000
4	2	2012	Jarik Standar Tulis Sarin	Standar	Jarik	Tulis	Ibu Harini	WANITA	Jawa Timur	4	40000
4	2	2011	Hem Katun Print Kaung	Katun	Hem	Print	Ibu Harini	WANITA	Jawa Timur	3	70000
9	3	2012	Batik Standar Cap Tumpa	Standar	Batik	Cap	Bapak Heru	PRIA	Jawa Timur	1	150000
5	2	2012	Hem Katun Print Kelenga	Katun	Hem	Print	Bapak Totok	PRIA	Jawa Timur	3	299000
12	4	2011	Bolero Standar Cap Sido	Standar	Bolero	Cap	Ibu Hatamah	WANITA	Jawa Timur	1	225000
10	4	2011	Sarimbit Standar Print Lu	Standar	Sarimbit	Print	Ibu Hatamah	WANITA	Jawa Timur	1	150000
1	1	2011	Kaos Katun Print Bola	Katun	Kaos	Print	Bapak Imron	PRIA	Jawa Barat	1	60000
2	1	2012	Celana Standar Cap War	Standar	Celana	Cap	Ibu Hadi Sukarni	WANITA	Jawa Barat	17	55000
3	1	2010	Celana Standar Cap War	Standar	Celana	Print	Ibu Hadi Sukarni	WANITA	Jawa Barat	17	55000
3	1	2011	Bahan Standar Cap Laser	Standar	Bahan	Cap	Ibu Siti Arya	WANITA	Jawa Barat	8	120000
12	4	2012	Rok Batik Print Kombina	Batik	Rok	Print	Ibu Siti Arya	WANITA	Jawa Barat	1	225000
1	1	2012	Jam Standar Print Lukis	Standar	Jam	Print	Ibu Siti Arya	WANITA	Jawa Barat	44	80000
9	3	2012	Hem Standar Cap Tumpa	Standar	Hem	Cap	Ibu Aini Kasmaj	WANITA	Jawa Tengah	1	100000
6	2	2012	Bahan Lawasan Tulis Toi Lawasan	Bahan	Tulis	Ibu Niken	WANITA	Jawa Tengah	1	130000	
8	3	2011	Hem Standar Tulis Madu	Standar	Hem	Tulis	Ibu Atik	WANITA	Jawa Tengah	5	550000
4	2	2012	Bahan Standar Cap Gari	Standar	Bahan	Cap	Ibu Tyas	WANITA	Jawa Tengah	7	135000
6	2	2010	Bahan Beludru Cap Mahl Beludru	Bahan	Bahan	Cap	Ibu Tyas	WANITA	Jawa Tengah	1	500000
11	4	2010	Hem Sutra Print Rama	Sutra	Hem	Print	Ibu Tyas	WANITA	Jawa Tengah	5	100000

b. tekan tombol ctrl+shift+\*

c. lalu klik tab insert pada ribbon, pilih menu pivotTable| InsertPivotTable

d.lalu pilih new Worksheet klik tombol OK. Hasilnya gambar dibawah ini :

The screenshot shows the Microsoft Excel interface with the PivotTable Tools ribbon tab selected. On the left, there is a PivotTable named 'PivotTable1' located at cell A3. The PivotTable Field List pane on the right lists fields such as bulan, kuartal, tahun, nama\_produk, nama\_kategori, and others. The Row Labels and Values sections are currently empty.

e. lalu centang yang ada di PivotTable Field List yaitu :

- tahun

-nama\_sub kategori

- jumlah

Book1 - Microsoft Excel

PivotTable... Options Design

PivotTable Active Field Group Sort Refresh Change Data Source Sort Data Actions Tools Show/Hide

A3 Sum of jumlah

	A	B	C	D	E	F
1						
2						
3	Sum of jumlah	Column Labels				
4	Row Labels	2010	2011	2012	Grand Total	
5	Bahan	1	8	8	17	
6	Batik			1	1	
7	Bolero		1		1	
8	Celana	17		17	34	
9	Hem	5	8	4	17	
10	Jam			44	44	
11	Jarik		2	4	6	
12	Kaos	1	14		15	
13	Rok		1		1	
14	Sarimbit	1		1		
15	Grand Total	23	21	93	137	
16						
17						
18						
19						
20						
21						
22						
23						
24						
25						

PivotTable Field List

Choose fields to add to report:

bulan  
 kuartal  
 tahun  
 nama\_produk  
 nama\_kategori  
 nama\_subkategori  
 nama\_pola  
 nama\_pelanggan  
 jenis\_kelamin  
 nama\_wilayah  
 jumlah  
 harga

Drag fields between areas below:

Report Filter Column Labels

tahun

Row Labels Row Labels

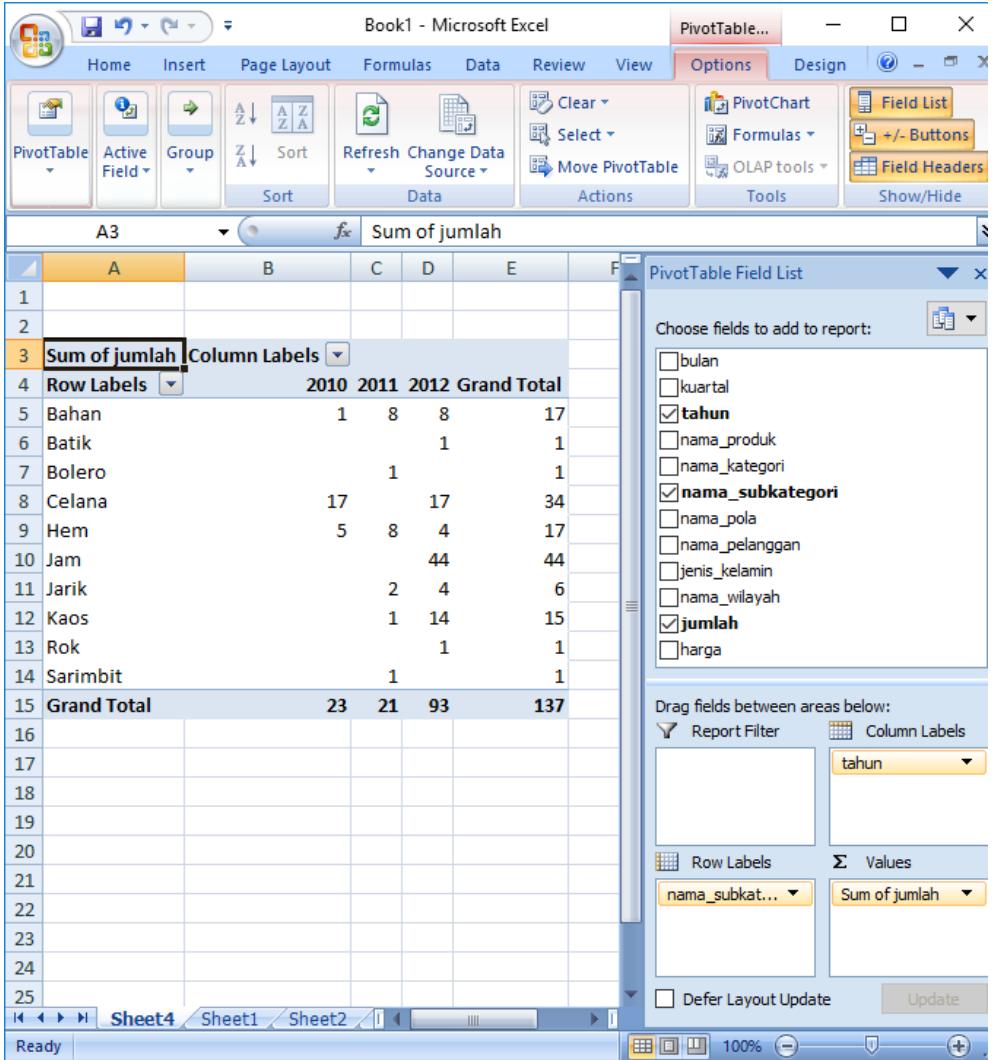
nama\_subkat...

Σ Values Sum of jumlah

Defer Layout Update Update

Sheet4 Sheet1 Sheet2

Ready



## Kegiatan 2

- a. tambahkan Field jumlah kembali ke kotak Value dengan cara drag. Sehingga akan muncul Sum of jumlah2

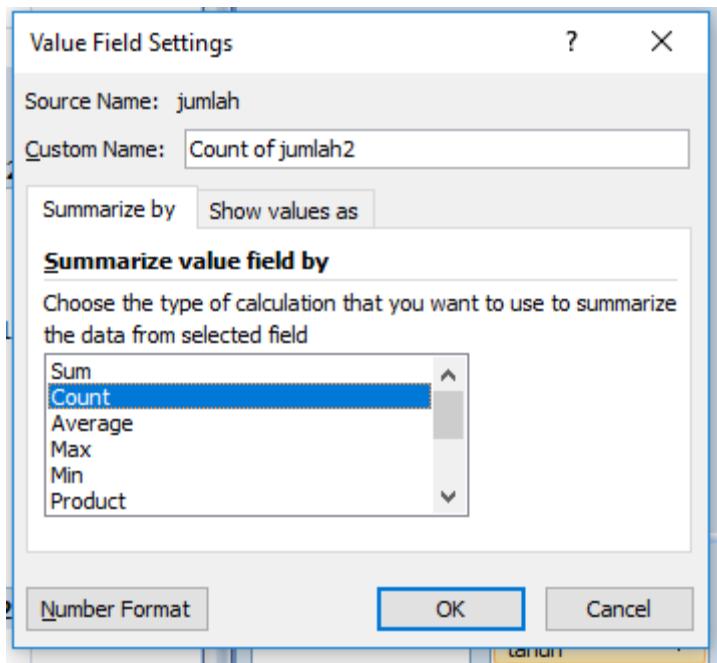
The screenshot shows a Microsoft Excel window titled "Book1 - Microsoft Excel". A PivotTable is being created in cells A3:D25. The PivotTable Field List dialog is open on the right, showing fields to add to the report. The "tahun" and "jumlah" fields are checked. In the "Values" section, "Sum of jumlah" is currently selected. An arrow points from the text "b. akan ada tambahan satu kolom perhitungan baru yang sama dengan hasil sebelumnya pada masing masing tahun tapi bukan ini yang diinginkan" towards this section.

	A	B	C	D
1				
2				
3	Column Labels			
4		2010		2011
5	Row Labels	Sum of jumlah	Sum of jumlah2	Sum of jumlah
6	Bahan	1	1	
7	Batik			
8	Bolero			
9	Celana	17	17	
10	Hem	5	5	
11	Jam			
12	Jarik			
13	Kaos			
14	Rok			
15	Sarimbit			
16	Grand Total	23	23	23
17				
18				
19				
20				
21				
22				
23				
24				
25				

- b. akan ada tambahan satu kolom perhitungan baru yang sama dengan hasil sebelumnya pada masing masing tahun tapi bukan ini yang diinginkan

c. kembali ke area values dan klik tombol panah kebawah pada field su, of jumlah2 Pilih item value field Setting

d. lalu pada dialog Value Field setting, ubah Sum menjadi Count. Lalu nama field akan berubah menjadi Count of Jumlah2

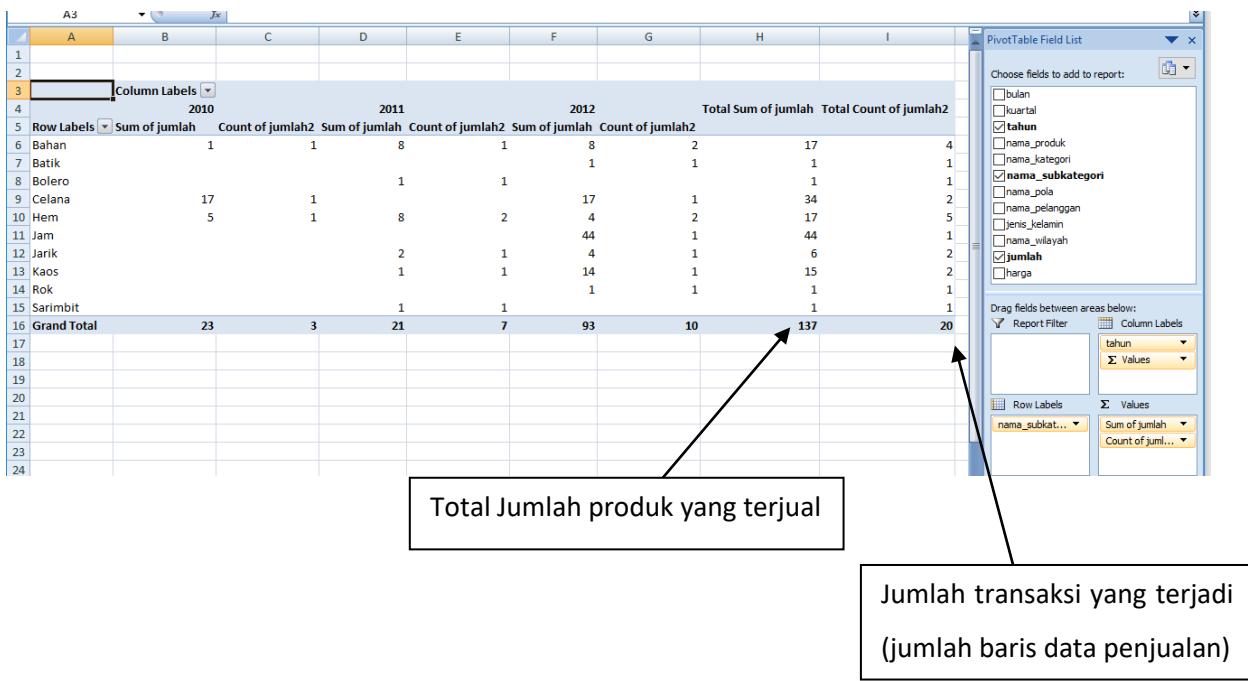


e. klik OK

f. pada area Pivot Table, didapatkan dua summary yaitu:

a) nilai jumlah unit penjualan yang terjadi (sum)

b) jumlah transaksi yang terjadi (count)



g. Simpan kembali dengan nama file yang sama

### Kegiatan 3

- a. buka sheet 1 fakta\_penjualan.xls, dan letakkan kursor kearea pivotTable
- b.Pada Menu Ribbon PivotTable Tools |Options, klik button formulas dan pilih Calculated Field

The screenshot shows a Microsoft Excel window with the ribbon set to 'PivotTable Tools | Options'. A black arrow points to the 'Formulas' button in the ribbon's dropdown menu. The main area displays a PivotTable with data for various items across different years. The PivotTable Field List pane on the right shows fields being assigned to the report, with 'tahun' selected for the Column Labels and 'Sum of jumlah' and 'Count of jumlah' selected for the Values.

- c. pada kontak dialog Insert Calculated Field yang muncul, masukan nilai kemudian klik OK

a)Name : Pendapatan

b)Formula = jumlah \* harga

The screenshot shows a Microsoft Excel window with a PivotTable named 'PivotTable1'. The PivotTable displays data for the year 2010, categorized by items like Bahau, Batik, Bolero, Celana, Hem, Jam, Jarik, Kaos, Rok, and Sarimbit, with a Grand Total of 23. A calculated field 'pendapatan' is being created with the formula  $=jumlah*harga$ . The PivotTable Field List on the right shows fields like tahun, nama\_subkategori, jumlah, and others.

d. Field baru “Sum of Pendapatan” akan muncul pada Pivot Table

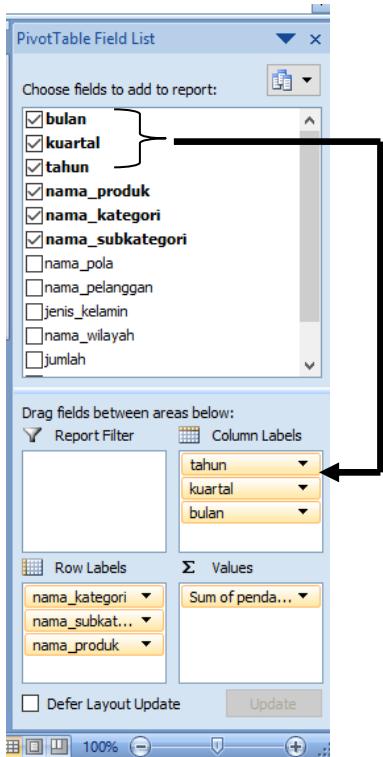
The screenshot shows the same Microsoft Excel window after the calculated field 'pendapatan' has been added. The PivotTable now includes columns for 'Sum of pendapatan' and 'Total Sum of pendapatan'. The PivotTable Field List on the right shows the newly added field 'pendapatan' in the Values area.

## Kegiatan 4

- a. buka sheet 1 letakkan kursor pada area pivot table
- b.pada pivot table list hilangkan tanda centang field jumlah.
- c. beri tanda centang dan letakkan field” berikut pada kotak row labels atau column Labels sesuai dengan kebutuhan tampilan cube
- d. misalkan pada row labels akan ditampilkan data erdasarkan urutan nama\_kategori, nama\_subkategori, dan nama produk. Lalu drag and drop dan letakkan pada kotak row labels.

The screenshot shows the Microsoft Excel ribbon with the 'PivotTable Tools' tab selected. The 'Options' tab is active. On the left, a PivotTable is displayed with the formula 'Sum of pendapatan' in cell A3. To the right is the 'PivotTable Field List' dialog box. Under 'Choose fields to add to the report', several fields are listed with checkboxes: bulan (unchecked), kuartal (unchecked), tahun (checked), nama\_produk (checked), nama\_kategori (checked), nama\_subkategori (checked), nama\_pola (unchecked), nama\_pelanggan (unchecked), jenis\_kelamin (unchecked), nama\_wilayah (unchecked), and jumlah (unchecked). Below this, under 'Drag fields between areas below', 'Report Filter' is set to 'tahun'. The 'Column Labels' section contains the field 'tahun'. The 'Row Labels' section contains the fields 'nama\_kategori' and 'nama\_subkategor...'. The 'Values' section contains the field 'Sum of pendapatan'. At the bottom of the dialog, there is a 'Defer Layout Update' checkbox and an 'Update' button. The status bar at the bottom of the Excel window shows 'Ready', 'Average: 5746622,952', 'Count: 212', 'Sum: 959686033', and '100%'. The tabs at the bottom of the screen show 'Sheet4', 'Sheet1', 'Sheet2', and 'Sheet3'.

- e. pada column labels akan ditampilkan data berdasarkan urutan **tahun,kuartan** dan **bulan**. Beri tanda centang pada field tersebut (drag and drop) dan letakkan pada kotak **Column Labels**



f. Lihat kembali pada cube setelah ditambahkan field-field untuk operasi roll up dan drill down.

g. Pada masing masing **Row Labels** dan **Column Labels** telah bertambah fiel field yang bisa diperinci dan diringkas sesuai urutan kategori data yang lebih spesifik.

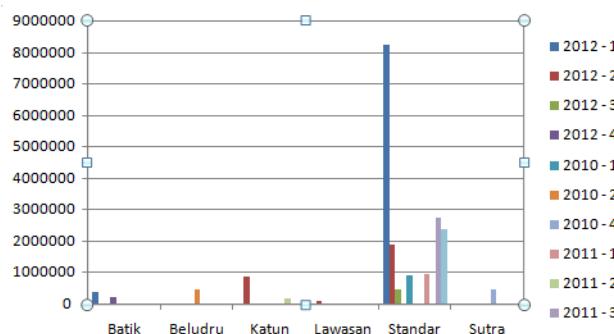
Screenshot of Microsoft Excel showing a PivotTable setup. The PivotTable is named "PivotTable1" and has "Sum of pendapatan" as the active field. The PivotTable Field List shows fields like bulan, kuartal, tahun, nama\_produk, nama\_kategori, and nama\_subkategor. The PivotTable itself displays data for various months and years, with a total sum of 847555,847.

Screenshot of Microsoft Excel showing a PivotTable setup. The PivotTable is named "PivotTable1" and has "kuartal" as the active field. The PivotTable Field List shows fields like bulan, kuartal, tahun, nama\_produk, nama\_kategori, and nama\_subkategor. The PivotTable displays data grouped by quarter, with a total sum of 1620000.

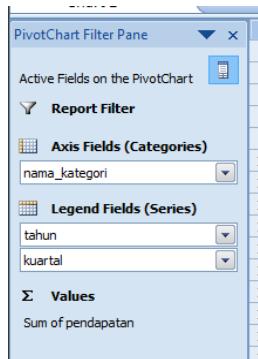
h. Klik tanda – untuk melakukan operasi **Roll Up** dan klik tanda untuk melakukan operasi **Drill Down**.

## Kegiatan 5

- a. buka sheet 1 fakta\_penjualan.xls, dan letakkan kursor ke area pivotTable
- b. pada menu Option klik menu PivotChart
- c. lalu pilih gambar grafik yang kamu inginkan. Lalu tekan Ok
- d. Grafik akan ditampilkan dengan sumbu X dan sumbu Y menyesuaikan dengan Row Labels dan Column Labels.
- e. Jika grafik terlalu rinci maka bisa dibuat secara lebih umum dengan menghilangkan kembalintanda centang pada field dalam **pivotTable Field**. centang saja **nama\_produk, nama\_subkategori, dan bulan.**



- f. dengan melihat grafi PivotChart diatas pola transaksi dari kuartal pertama hingga kuartal 4 dapat dilihat dengan mudah apakah terjadi kenaikan atau penurunann bahkan stabil untuk masing masing kategori produk
- g. Jendela PivotChart Filter Pane berfungsi untuk menyaring data data khusus yang akan ditampilkan saja.



## 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
  2. Buatlah pivot table dan PivotChart untuk melihat PPN dan otal Penghasilan tersebut selama tahun 2010-2012 . kategori produk apakah yang memberikan nilai penghasilan terbanyak selama 3 tahun tersebut.

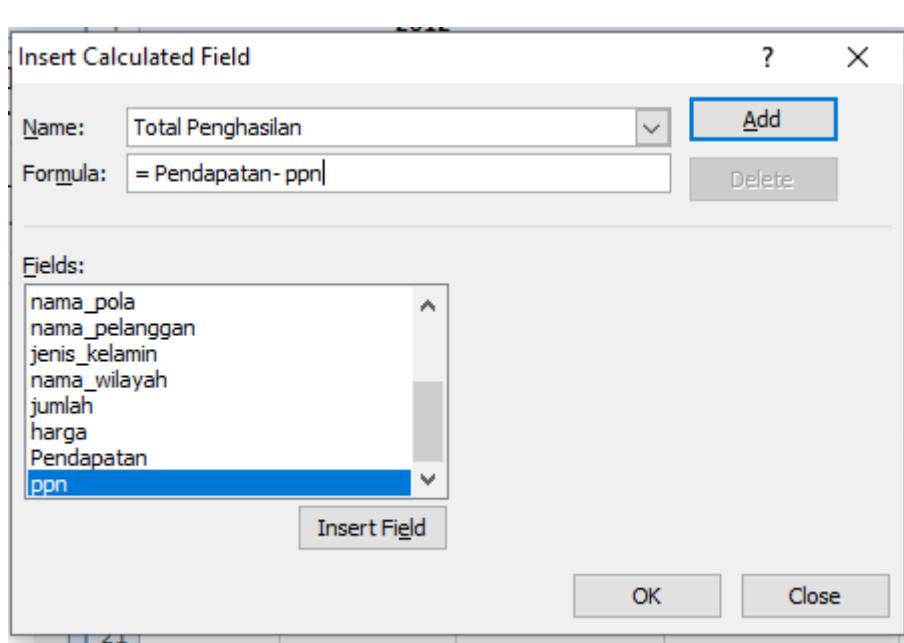
Jawaban :

1. A

- a. buka sheet 1 fakta\_penjualan.xls, dan letakkan kursor ke area pivotTable
  - b. Pada Menu Ribbon PivotTable Tools |Options, klik button formulas dan pilih Calculated Field.
  - c. pada kontak dialog Insert Calculated Field yang muncul, masukan nilai kemudian klik OK
    - a) Name : ppn
    - b) Formula =  $10/100 * \text{pendapatan}$
  - d. Field baru “ppn” akan muncul pada Pivot Table

**B.**

- a. buka sheet 1 fakta\_penjualan.xls, dan letakkan kursor ke area pivotTable
- b. Pada Menu Ribbon PivotTable Tools |Options, klik button formulas dan pilih Calculated Field.
- c. pada kontak dialog Insert Calculated Field yang muncul, masukan nilai kemudian klik OK
  - a) Name : Total Penghasilan
  - b) Formula = pendapatan - ppn



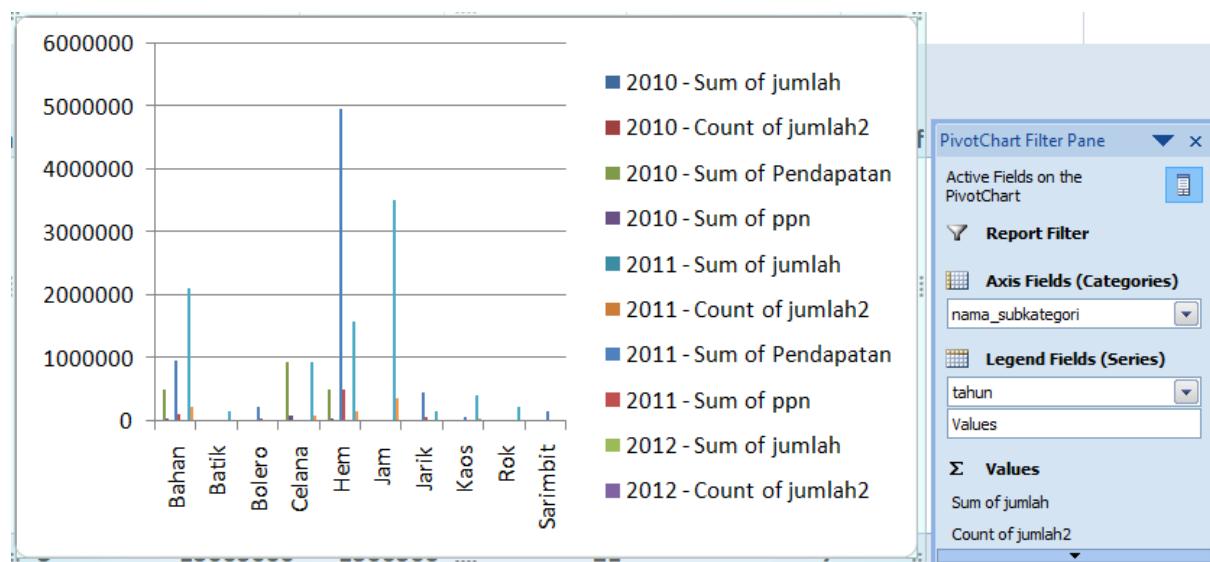
- d. Field baru "Total Penghasilan" akan muncul pada Pivot Table

2012	Sum of jumlah	Count of jumlah2	Sum of Pendapatan	Sum of ppn	Sum of Total Penghasilan	Total Sum of jumlah	Total Count of jumlah2	Total Su
	8	2	2120000	212000	1908000	17	4	
	1	1	150000	15000	135000	1	1	
			0	0	0	1	1	
	17	1	935000	93500	841500	34	2	
	4	2	1596000	159600	1436400	17	5	
	44	1	3520000	352000	3168000	44	1	
	4	1	160000	16000	144000	6	2	
	14	1	420000	42000	378000	15	2	
	1	1	225000	22500	202500	1	1	
			0	0	0	1	1	
	93	10	115692000	11569200	104122800	137	20	

			Total Sum of jumlah	Total Count of jumlah2	<b>Total Sum of Pendapatan</b>	Total Sum of ppn
Sum of Pendapatan	Sum of ppn	Sum of Total Penghasilan				
2120000	212000	1908000	17	4	15045000	150450
150000	15000	135000	1	1	150000	1500
0	0	0	1	1	225000	2250
935000	93500	841500	34	2	3740000	37400
1596000	159600	1436400	17	5	19023000	190230
3520000	352000	3168000	44	1	3520000	35200
160000	16000	144000	6	2	1590000	15900
420000	42000	378000	15	2	1350000	13500
225000	22500	202500	1	1	225000	2250
0	0	0	1	1	150000	1500
<b>115692000</b>	<b>11569200</b>	<b>104122800</b>	<b>137</b>	<b>20</b>	<b>451963000</b>	<b>4519630</b>

Total Sum of jumlah	Total Count of jumlah2	<b>Total Sum of Pendapatan</b>	Total Sum of ppn	Total Sum of Total Penghasilan
17	4	15045000	1504500	13540500
1	1	150000	15000	135000
1	1	225000	22500	202500
34	2	3740000	374000	3366000
17	5	19023000	1902300	17120700
44	1	3520000	352000	3168000
6	2	1590000	159000	1431000
15	2	1350000	135000	1215000
1	1	225000	22500	202500
1	1	150000	15000	135000
<b>137</b>	<b>20</b>	<b>451963000</b>	<b>45196300</b>	<b>406766700</b>

## 2. membuat pivot Chart.



Kategori produk yang memberikan nilai penghasilan terbanyak selama 3 tahun adalah **hem**.

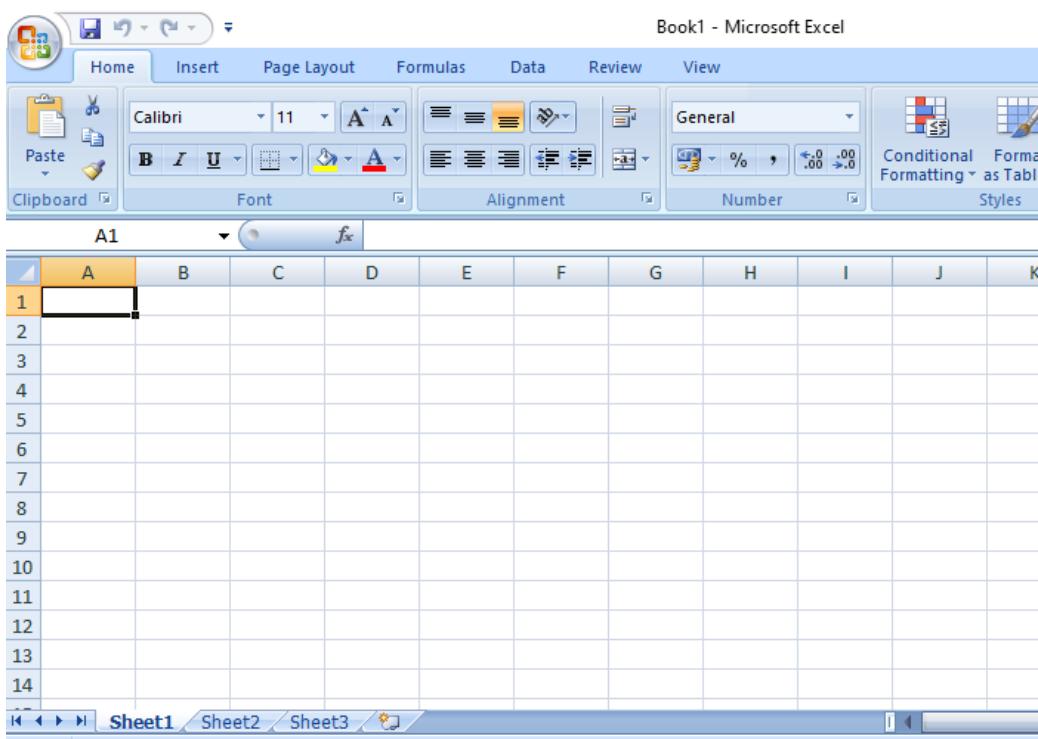
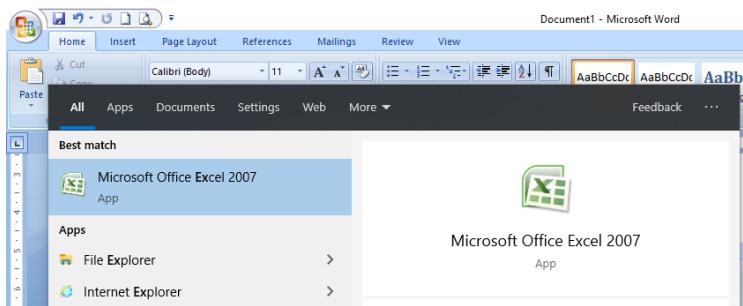
MODUL 6

## TUGAS

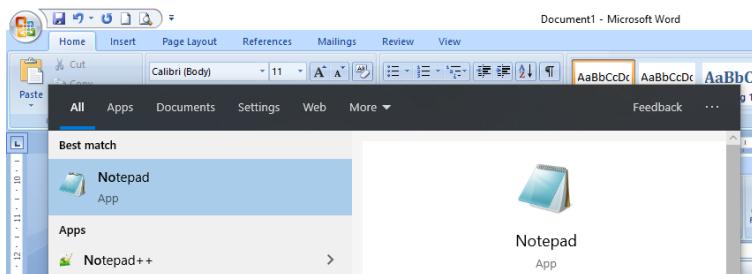
## TUGAS MODUL 6 PERTEMUAN KE 3

## TUGAS nomor 3

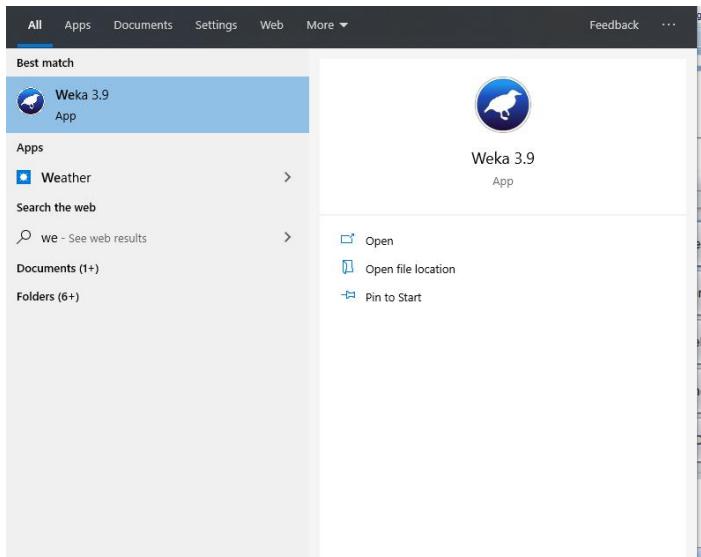
### 1. menginstal microsoft excel



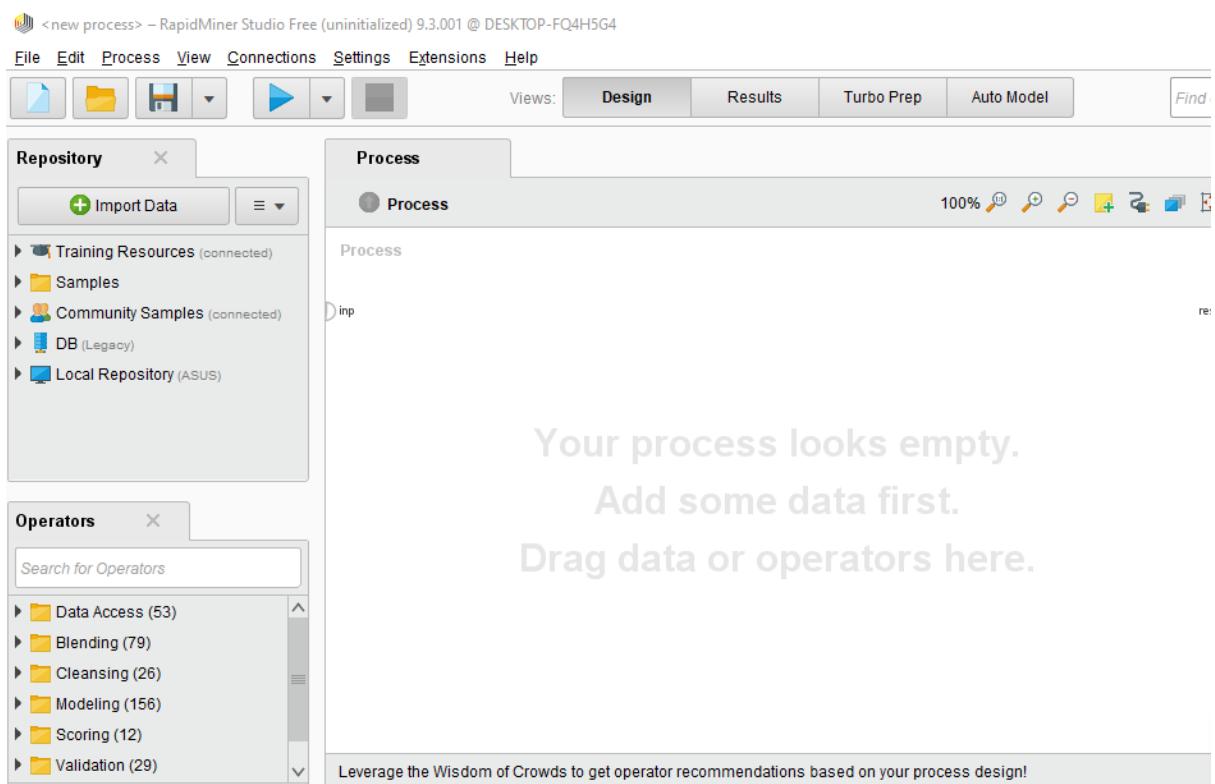
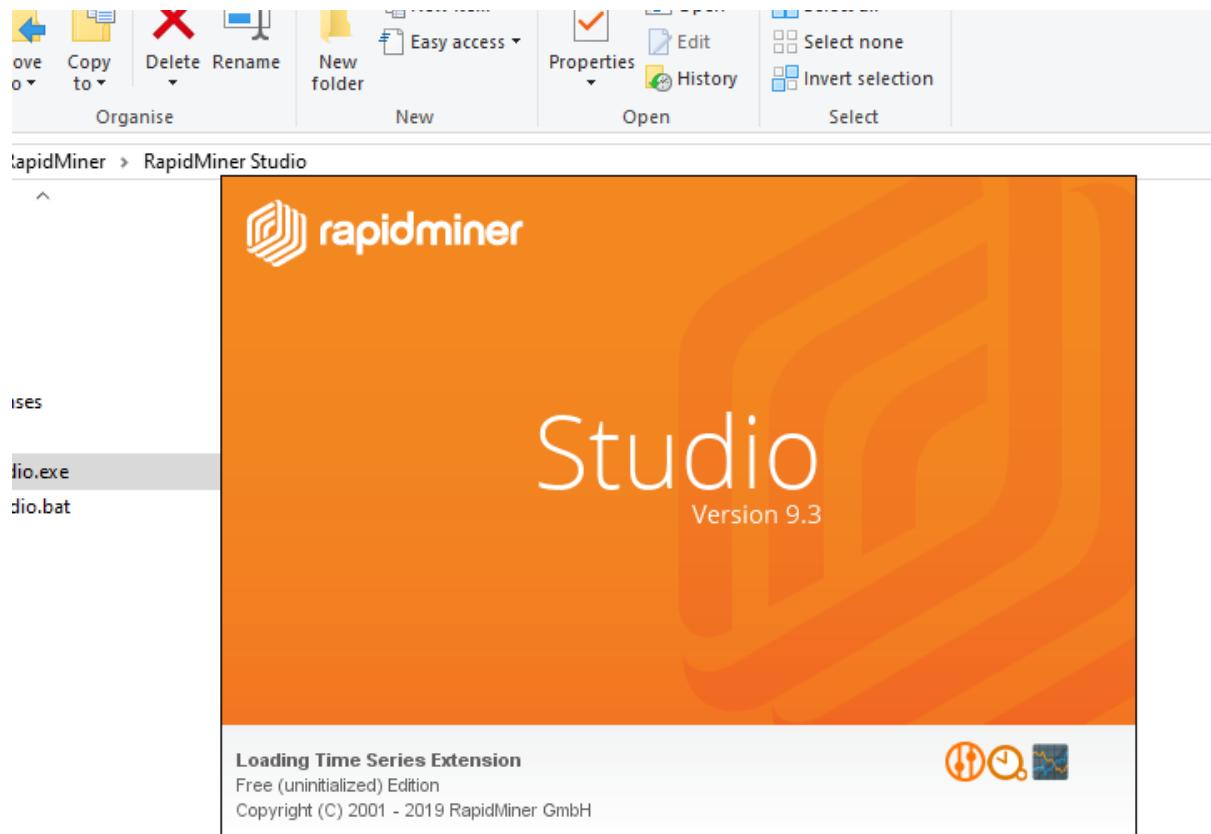
### 2. menginstal Notepad/notepad ++



### 3. menginstal Weka



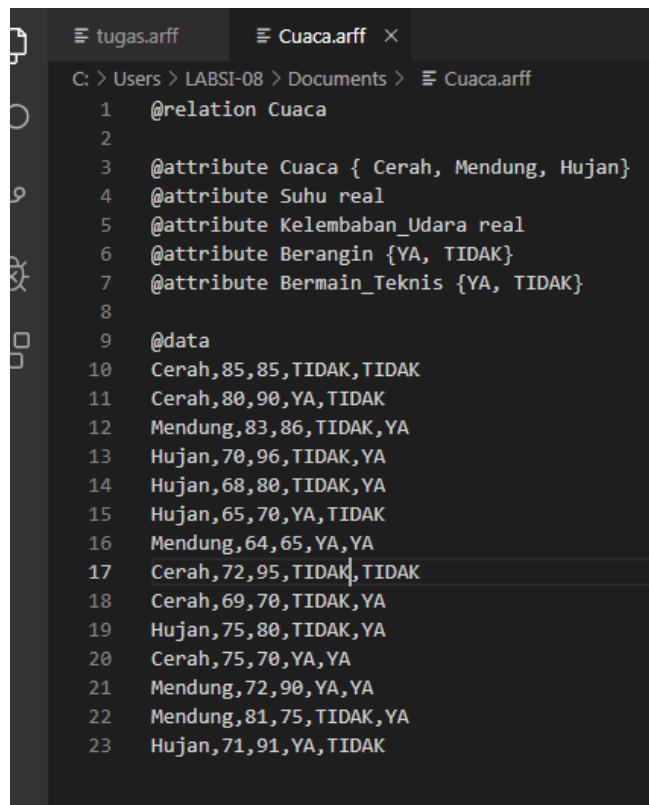
#### 4. Menginstal rapidminer



## MODUL 7

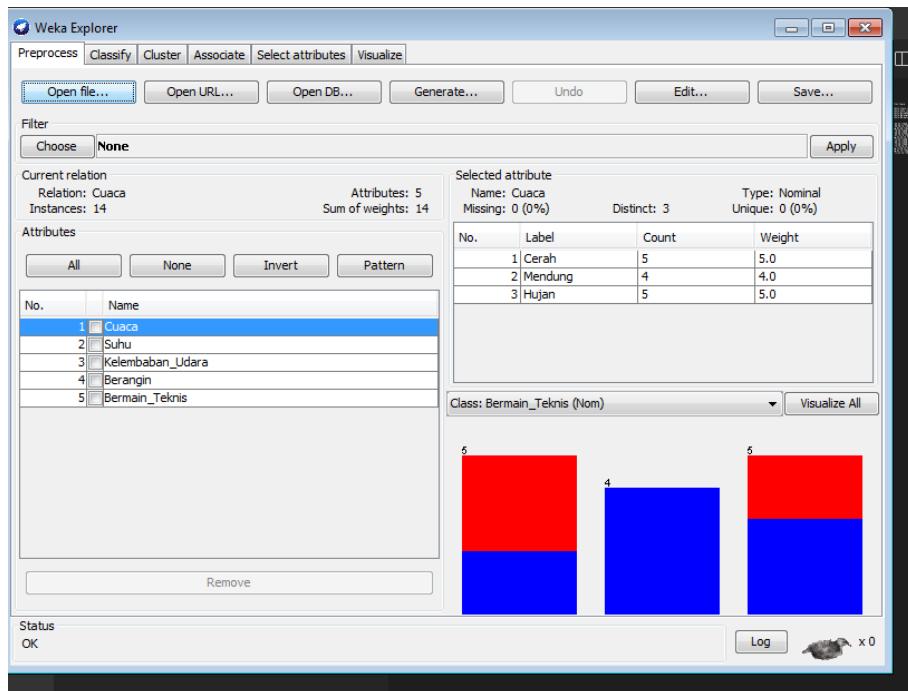
### Kegiatan

Hasil dari table cuaca

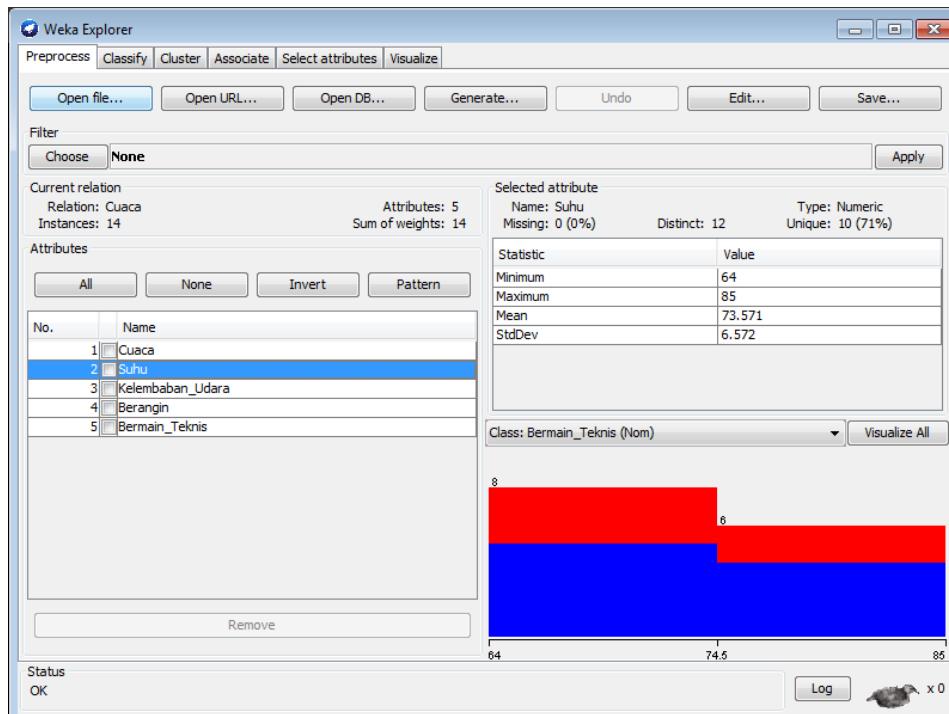


```
tugas.arff Cuaca.arff
C: > Users > LABSI-08 > Documents > Cuaca.arff
1 @relation Cuaca
2
3 @attribute Cuaca { Cerah, Mendung, Hujan}
4 @attribute Suhu real
5 @attribute Kelembaban_Udara real
6 @attribute Berangin {YA, TIDAK}
7 @attribute Bermain_Teknis {YA, TIDAK}
8
9 @data
10 Cerah,85,85,TIDAK,TIDAK
11 Cerah,80,90,YA,TIDAK
12 Mendung,83,86,TIDAK,YA
13 Hujan,70,96,TIDAK,YA
14 Hujan,68,80,TIDAK,YA
15 Hujan,65,70,YA,TIDAK
16 Mendung,64,65,YA,YA
17 Cerah,72,95,TIDAK,TIDAK
18 Cerah,69,70,TIDAK,YA
19 Hujan,75,80,TIDAK,YA
20 Cerah,75,70,YA,YA
21 Mendung,72,90,YA,YA
22 Mendung,81,75,TIDAK,YA
23 Hujan,71,91,YA,TIDAK
```

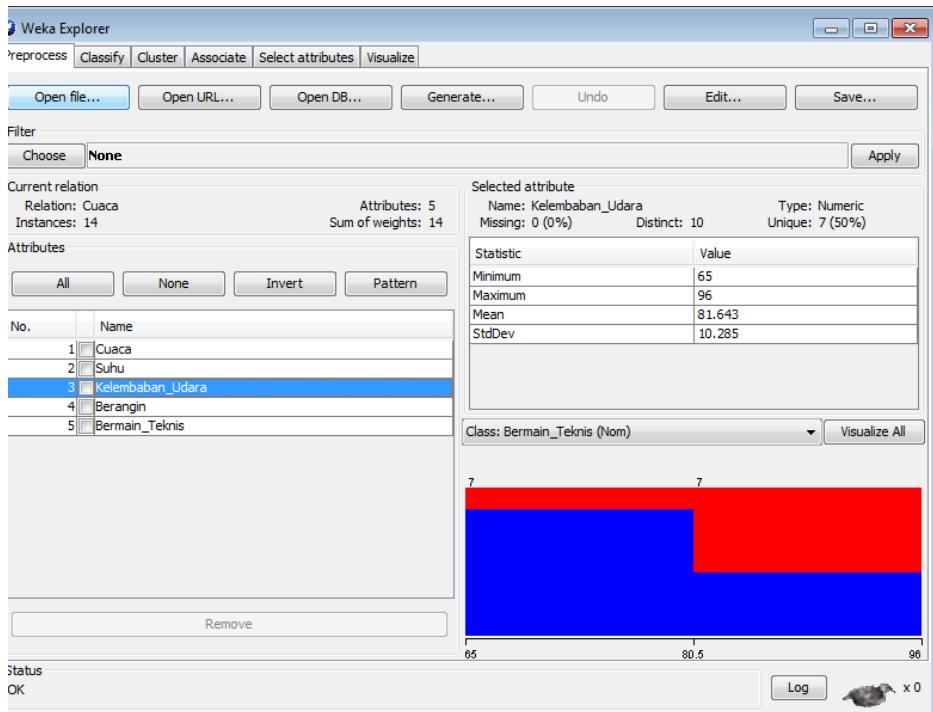
Grafik Cuaca :



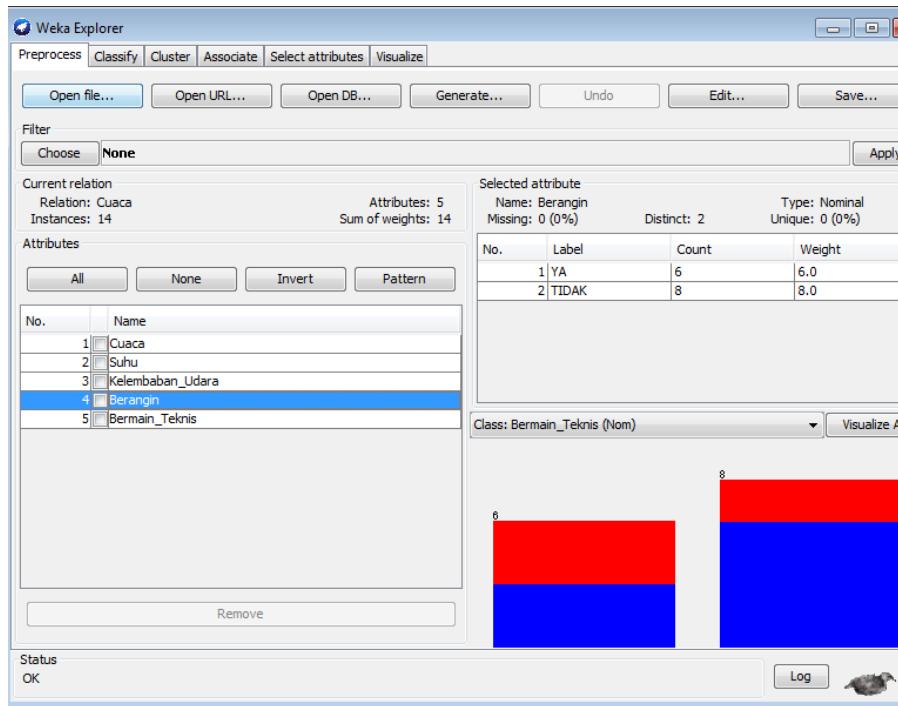
Grafik suhu :



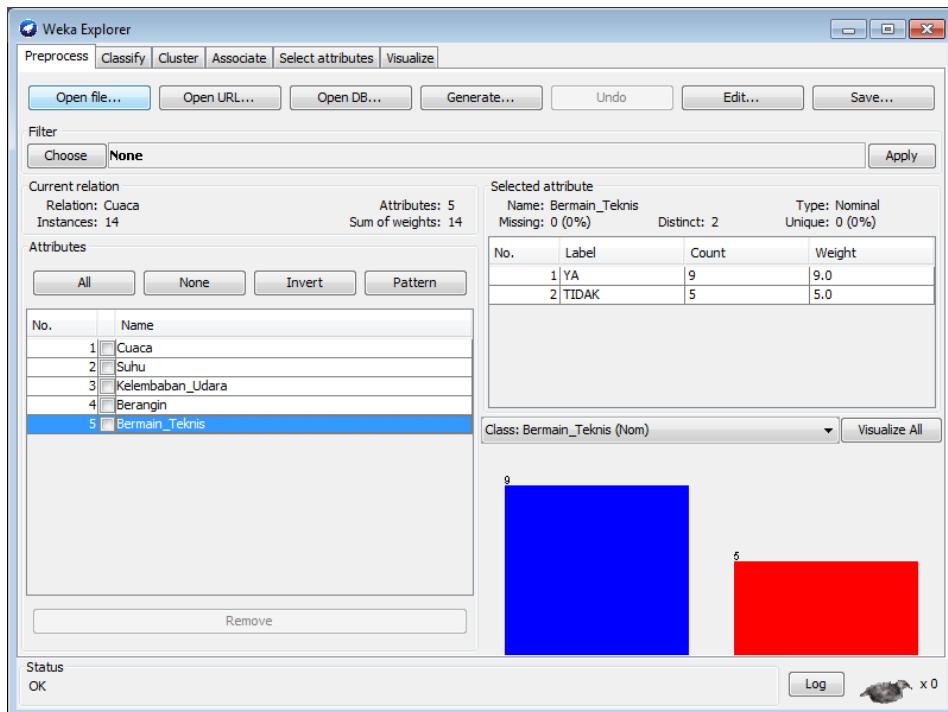
Grafik kelembaban udara :



Grafik berangin :



## Grafik Bermain tenis :



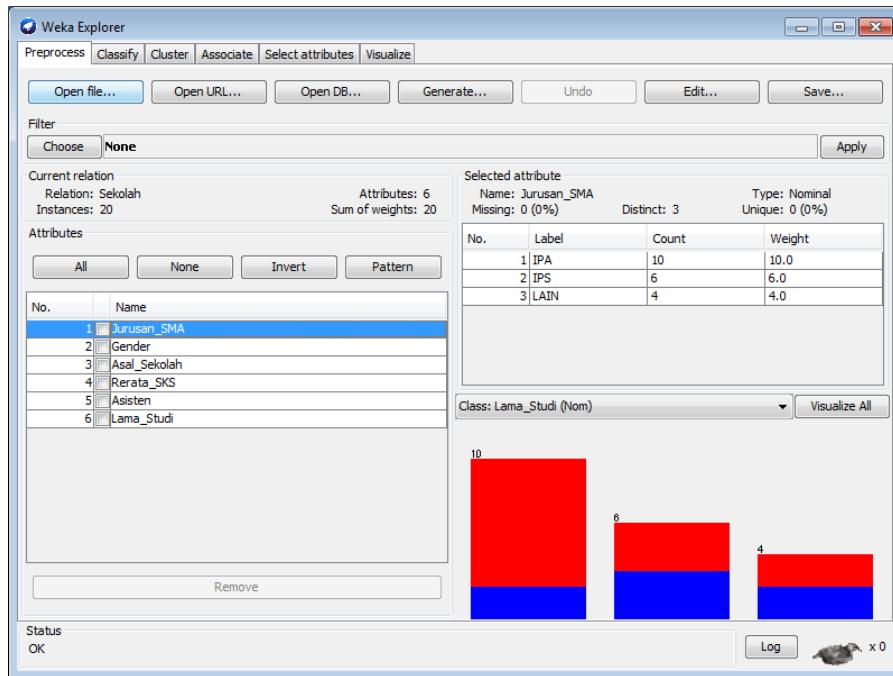
## TUGAS

1.

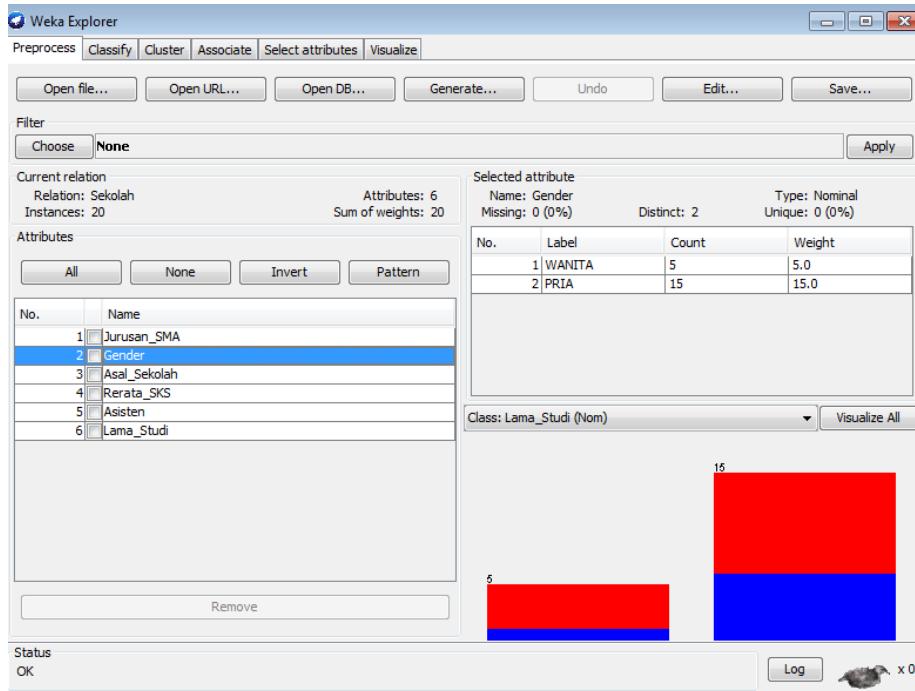
```
≡ tugas.arff ×
C: > Users > LABSI-08 > Documents > tugas.arff
1 @relation Sekolah
2
3 @attribute Jurusan_SMA{IPA, IPS, LAIN}
4 @attribute Gender{WANITA, PRIA}
5 @attribute Asal_Sekolah{SURAKARTA, LUAR}
6 @attribute Rerata_SKS real
7 @attribute Asisten{TIDAK, YA}
8 @attribute Lama_Studi{TERLAMBAT, TEPAT}
9
10 @data
11 IPA,WANITA,SURAKARTA,18,TIDAK,TERLAMBAT
12 IPA,PRIA,SURAKARTA,19,YA,TEPAT
13 LAIN,PRIA,SURAKARTA,19,TIDAK,TERLAMBAT
14 IPA,PRIA,LUAR,17,TIDAK,TERLAMBAT
15 IPA,WANITA,SURAKARTA,17,TIDAK,TEPAT
16 IPA,WANITA,LUAR,18,YA,TEPAT
17 IPA,PRIA,SURAKARTA,18,TIDAK,TERLAMBAT
18 IPA,PRIA,SURAKARTA,19,TIDAK,TEPAT
19 IPS,PRIA,LUAR,18,TIDAK,TERLAMBAT
20 LAIN,WANITA,SURAKARTA,18,TIDAK,TEPAT
21 IPA,WANITA,SURAKARTA,19,TIDAK,TEPAT
22 IPS,PRIA,SURAKARTA,20,TIDAK,TEPAT
23 IPS,PRIA,SURAKARTA,19,TIDAK,TEPAT
24 IPA,PRIA,SURAKARTA,19,TIDAK,TEPAT
25 IPA,PRIA,LUAR,22,YA,TEPAT
26 LAIN,PRIA,SURAKARTA,16,TIDAK,TERLAMBAT
27 IPS,PRIA,LUAR,20,TIDAK,TEPAT
28 LAIN,PRIA,LUAR,23,YA,TEPAT
29 IPA,PRIA,SURAKARTA,21,YA,TEPAT
30 IPS,PRIA,SURAKARTA,19,TIDAK,TERLAMBAT
31
```

2. Perlihatkan file ARFF dan juga gambar grafik untuk setiap data yang ditampilkan dalam weka berdasarkan file ARFF anda kepada dosen danasisten untuk dinilai.

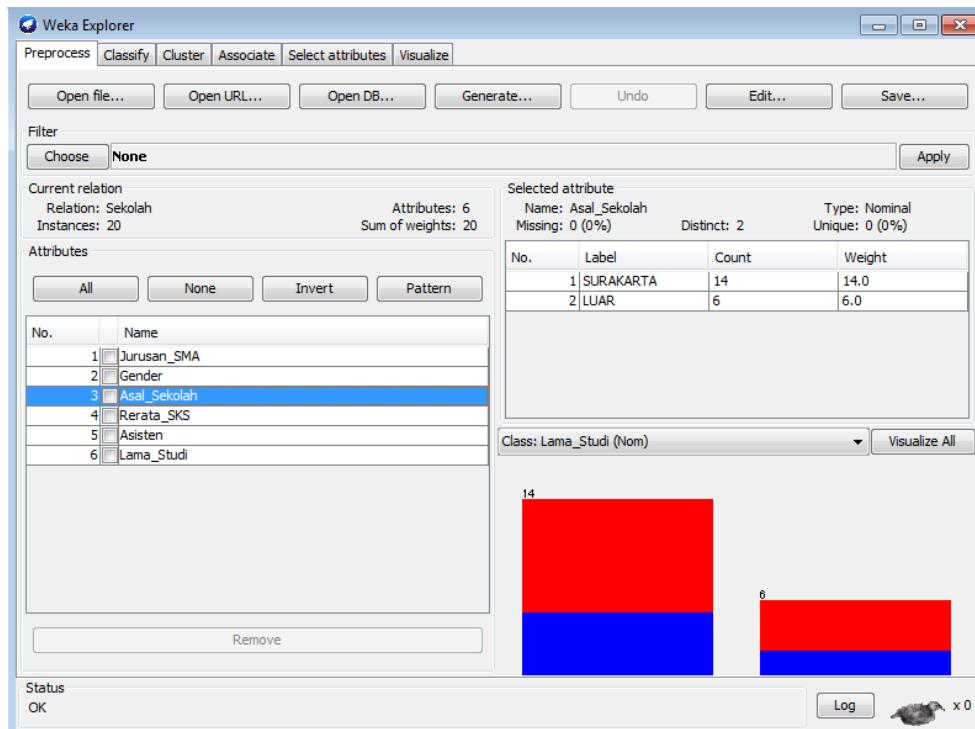
Grafik Jurusan\_SMA :



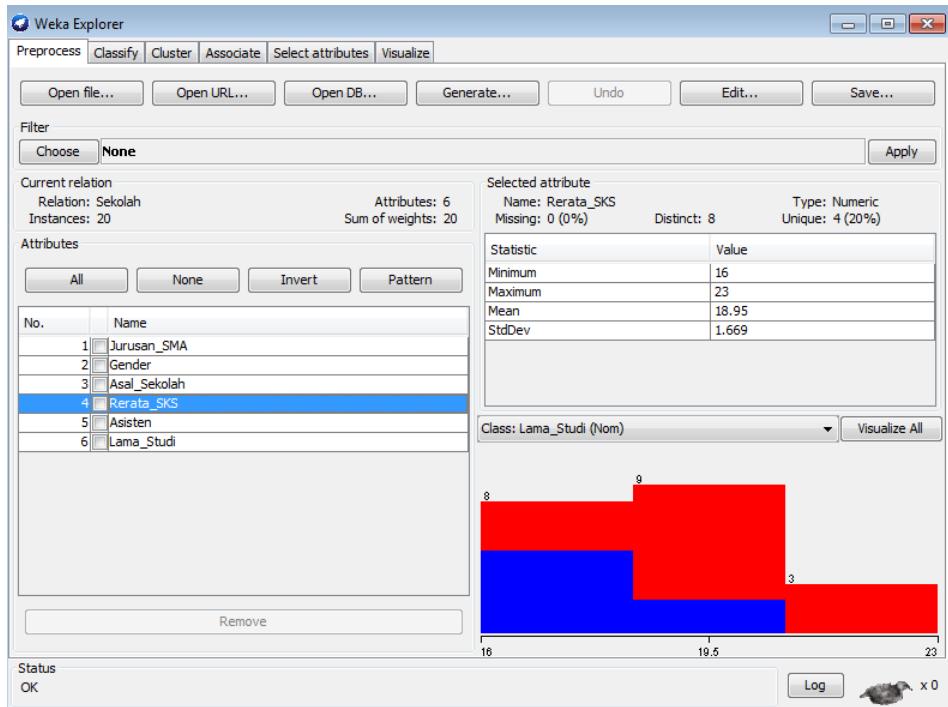
Grafik Gender :



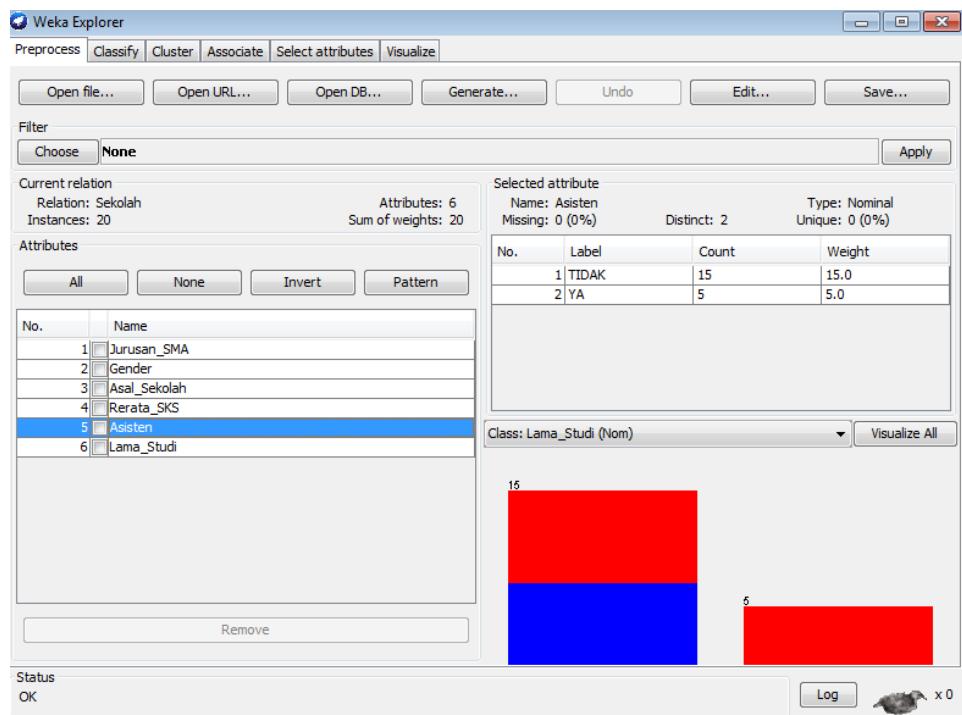
## Grafik Asal Sekolah :



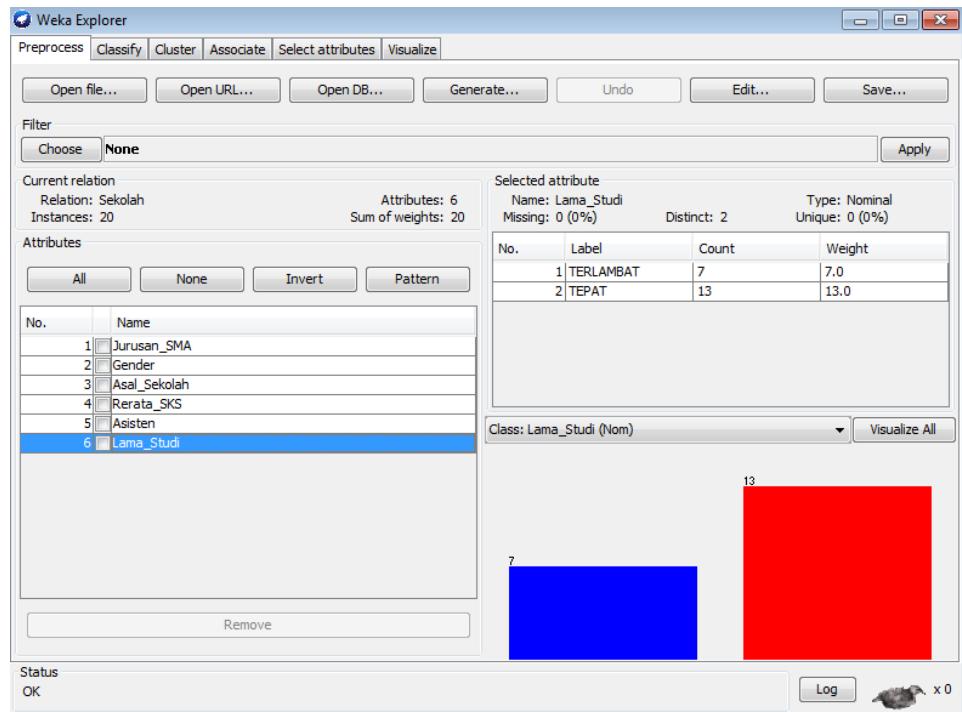
## Grafik Rerata\_SKS :



## Grafik Asisten :



## Grafik Lama\_Studi :



3. Berapa jumlah atribut yang bertipe binomial dan polynomial ?

- Binomial ada : 4 atribut yaitu Gender, Asal\_Sekolah, Asisten, Lama\_Studi
- Polynomial ada : 1 atribut yaitu Jurusan\_SMA

4. Brapa jumlah atribut yang bertipe real ?

- Real ada : 1 atribut yaitu Rerata\_SKS

5. Pada atribut Rerata\_SKS berapakah besarnya nilai Maximum, Minimum, Mean, dan StDev (Standard Deviation) ?

Maximum : 23 Value

Minimum : 16 Value

Mean : 18.95 value

StDev : 1.669 Value

Missing: 0 (0%)		Distinct: 8	Unique: 4 (20%)
Statistic	Value		
Minimum	16		
Maximum	23		
Mean	18.95		
StdDev	1.669		

Class: Lama\_Studi (Nom) ▾ Visualize All

## MODUL 8

Percobaan dengan weka

1. persiapkan file cuaca.arff

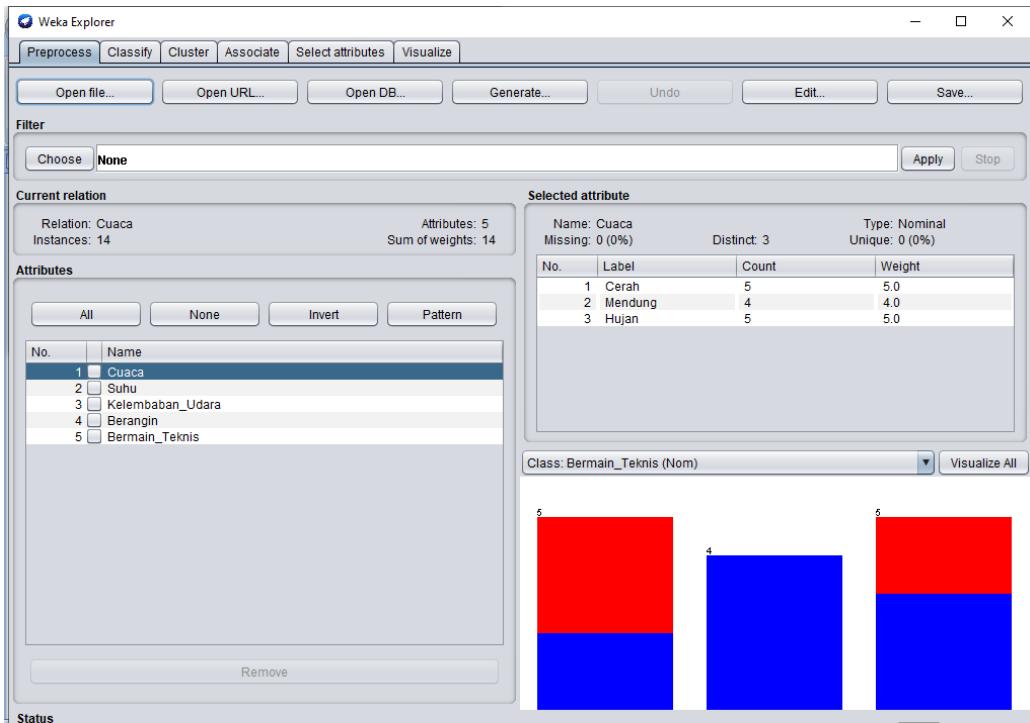
2. buatlah sebuah data testing dengan format arff dari table 8.1 sebagai data yang akan diprediksi dengan memiliki variabel variable independen dan variable dependen yang sama. Dengan ketentuan variable dipenden diisi dengan tanda ?.

3. simpan dengan nama Cuaca Testing.arff

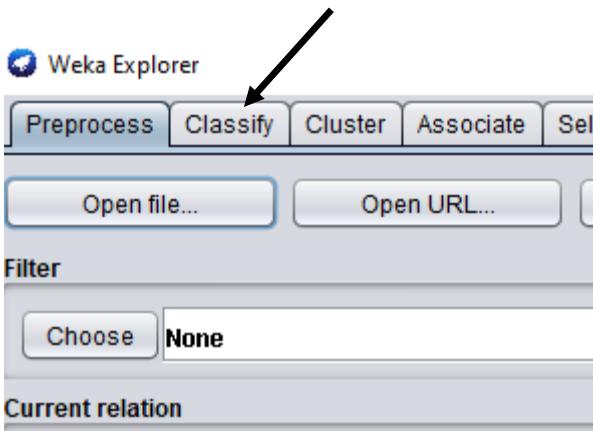
```
1 @relation Cuaca
2
3 @attribute Cuaca { Cerah, Mendung, Hujan}
4 @attribute Suhu real
5 @attribute Kelembaban_Udara real
6 @attribute Berangin {YA, TIDAK}
7 @attribute Bermain_Teknis {YA, TIDAK}
8
9 @data
10 Cerah,75,65,TIDAK,?
11 Cerah,80,68,YA,?
12 Cerah,83,87,YA,?
13 Mendung,70,96,TIDAK,?
14 Mendung,68,81,TIDAK,?
15 Hujan,65,75,YA,?
16 Hujan,64,85,YA,?_
```

4. jika telah selesai membuat buka aplikasi weka, masuk dalam menu Weka explorer

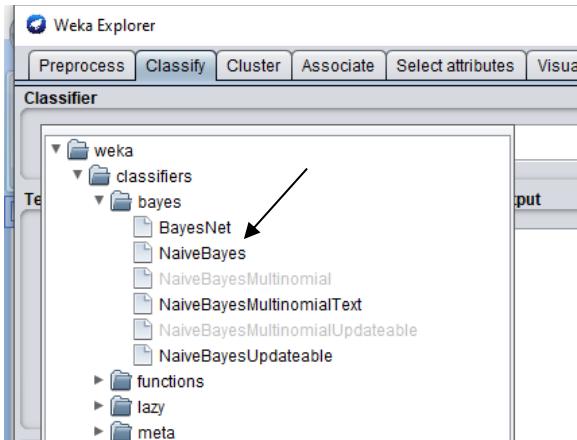
5. buka kembali file cuaca.arff



6. pilih tab classify pada jendela weka



7. klik, lalu akan muncul tombol choose dan klik untuk memilih metode / algoritma naive bayes

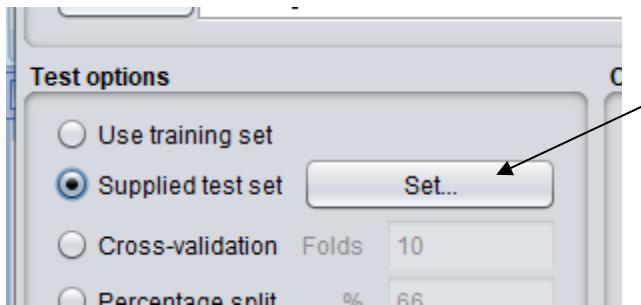


8. menentukan data testing sebagai data yang akan diprediksi variabel independennya. File CuacaTesting.

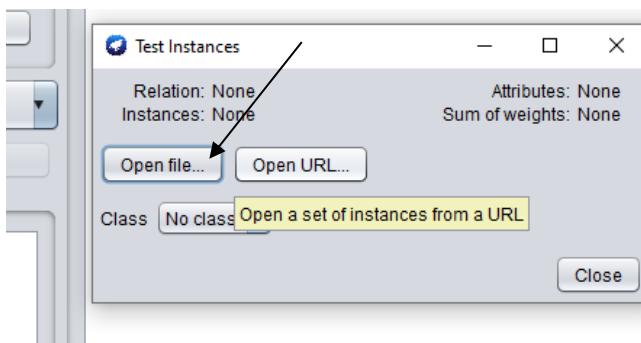
9. pada menu test options terdapat 4 pilihan pengujian yaitu :

- Use training set
- Supplied test st
- Cross-Validation
- Percentege split

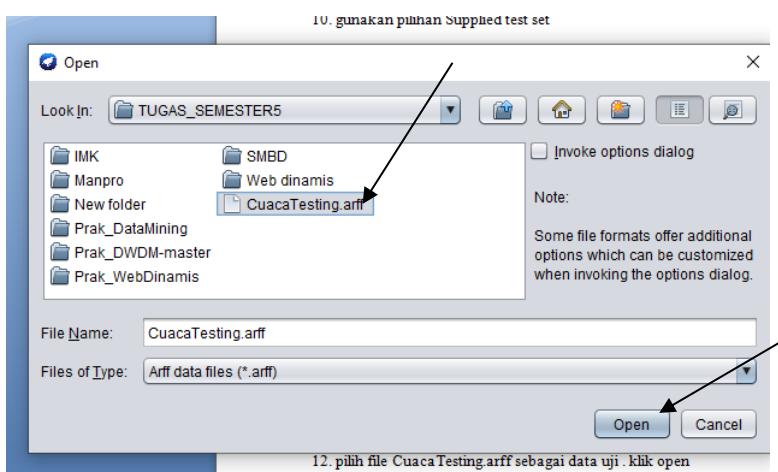
10. gunakan pilihan Supplied test set



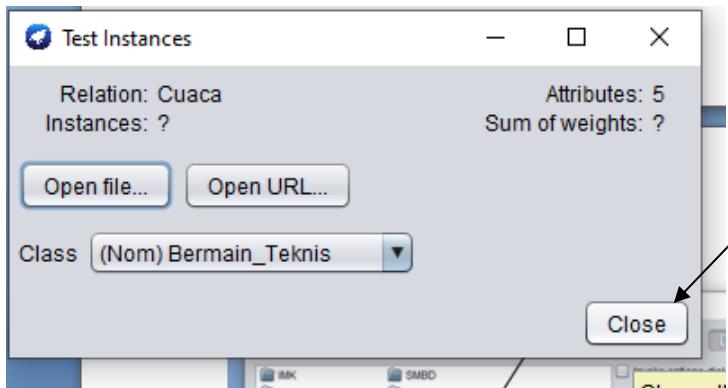
11. sehingga akan muncul jendela test instance. Klik open file



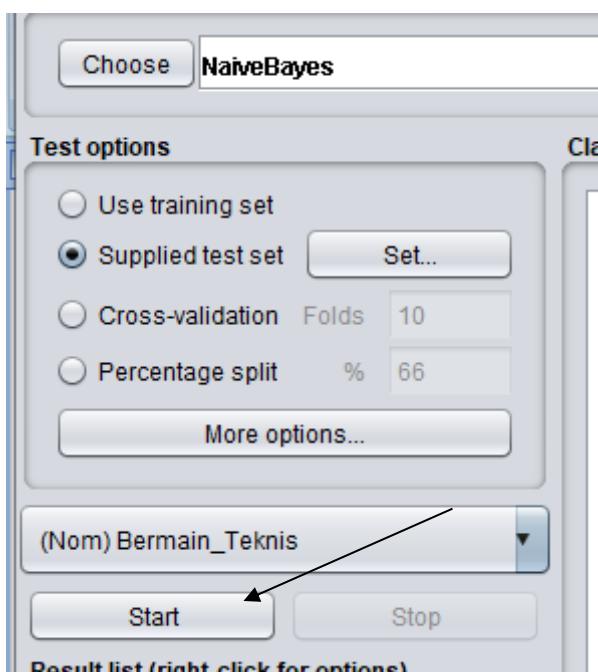
12. pilih file CuacaTesting.arff sebagai data uji . klik open



13. file CuacaTesting.arff akan diset sebagai data uji pada jendela test instance dengan variable predictor (class) adalah Bermain\_tenis. Lalu klik close



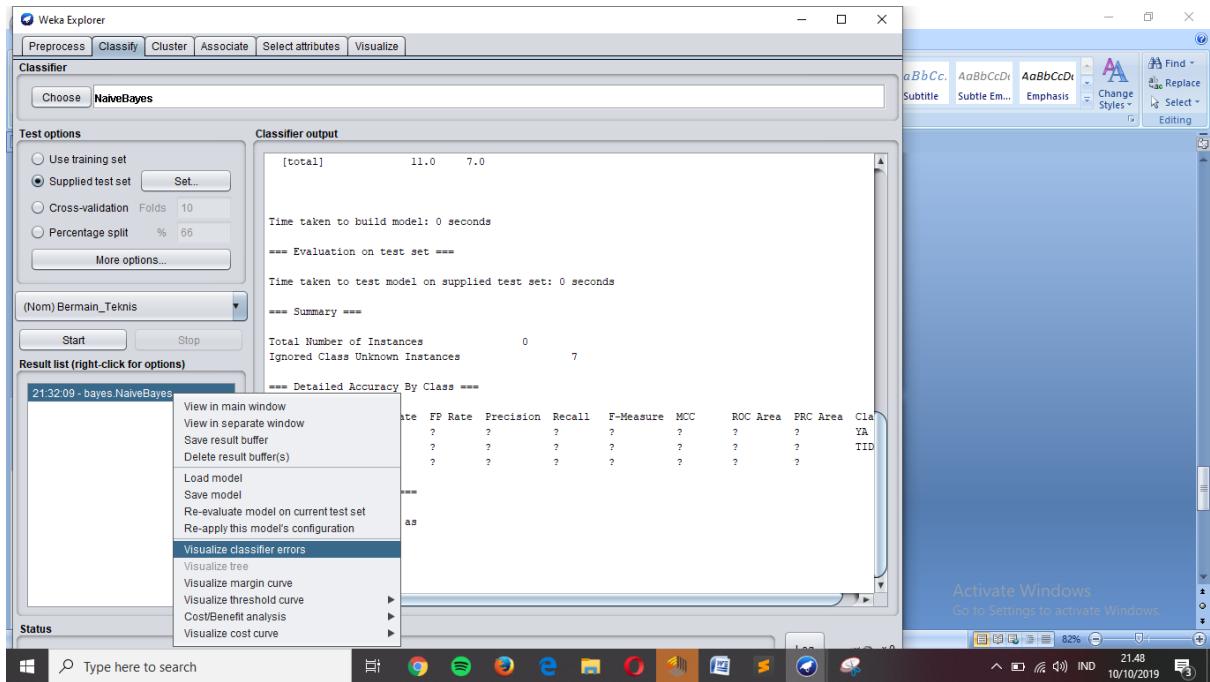
14. klik start untuk memulai proses naive bayes



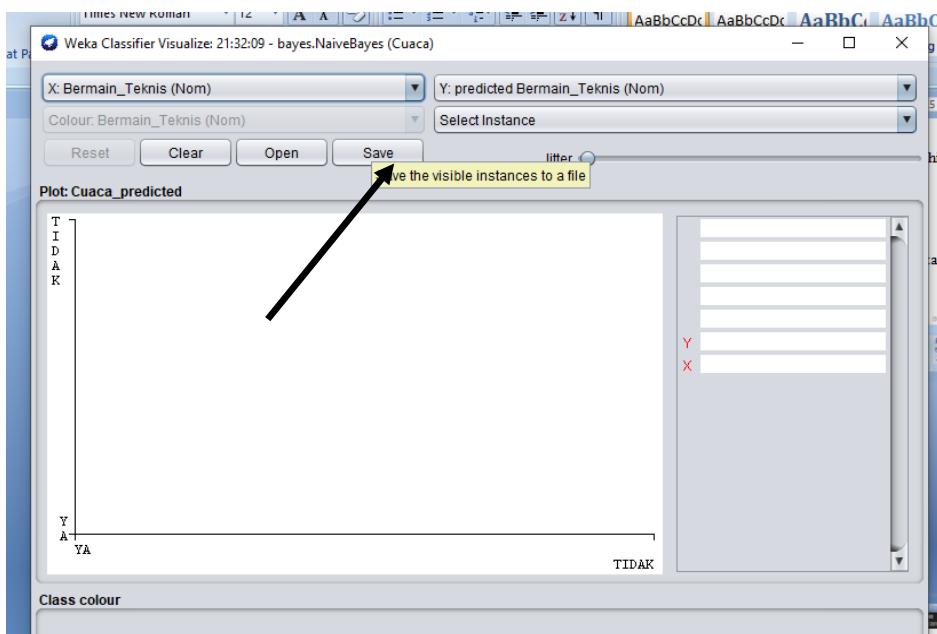
15. jika muncul jendela pesan classifier panel. Kita abaikan dengan mengeklik yes. Sehingga algoritma naive bayes akan diproses

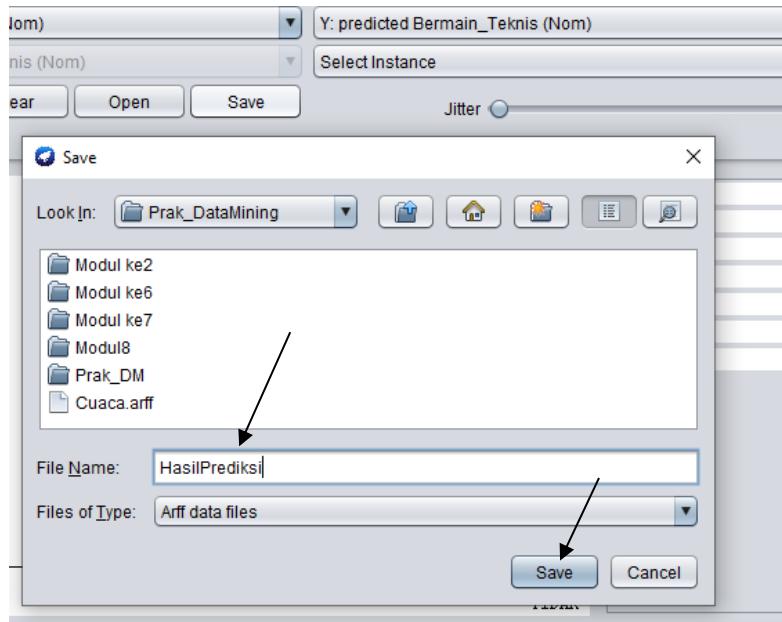
16. kita abaikan nilai nilai yang ditampilkan dalam jendela classifier output

17. untuk melihat hasil prediksi terhadap data uji, lihat nilai classifier errors. Kilik kanan pada hasil proses dalam kontak result list. Pilih menu visual classifier errors

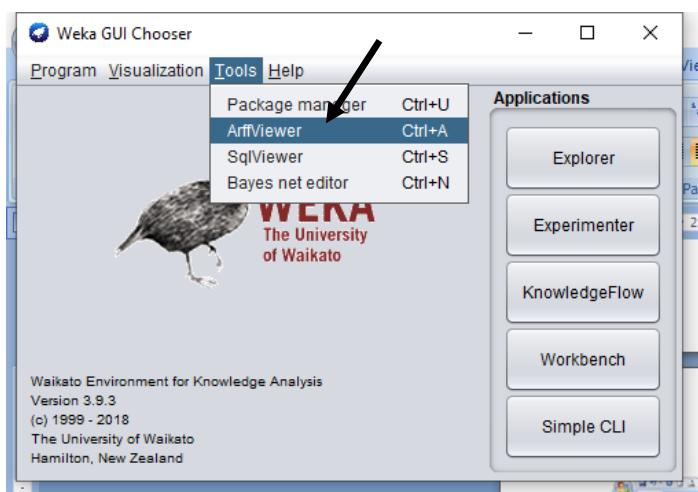


18. pada jendela weka classifier visualize, abaikan hasil apapun yang ditampilkan. Klik save . simpan dengan nama file HasilPrediksi.arff





19. tutup semua jendela. Lalu ke,bali ke weka GUI Chooser. Pilih menu Tools-ArffViewer



20. jendela arff-viewers akan ditampilkan. Buka menu file-open. Tunjukan pada file HasilPrediksi.arff jendela arff-viewers akan ditampilkan. Buka menu file-open. Tunjukan pada filenHasilPrediksi.arff ang telah anda simpan. Lihatlah hasil prediksi

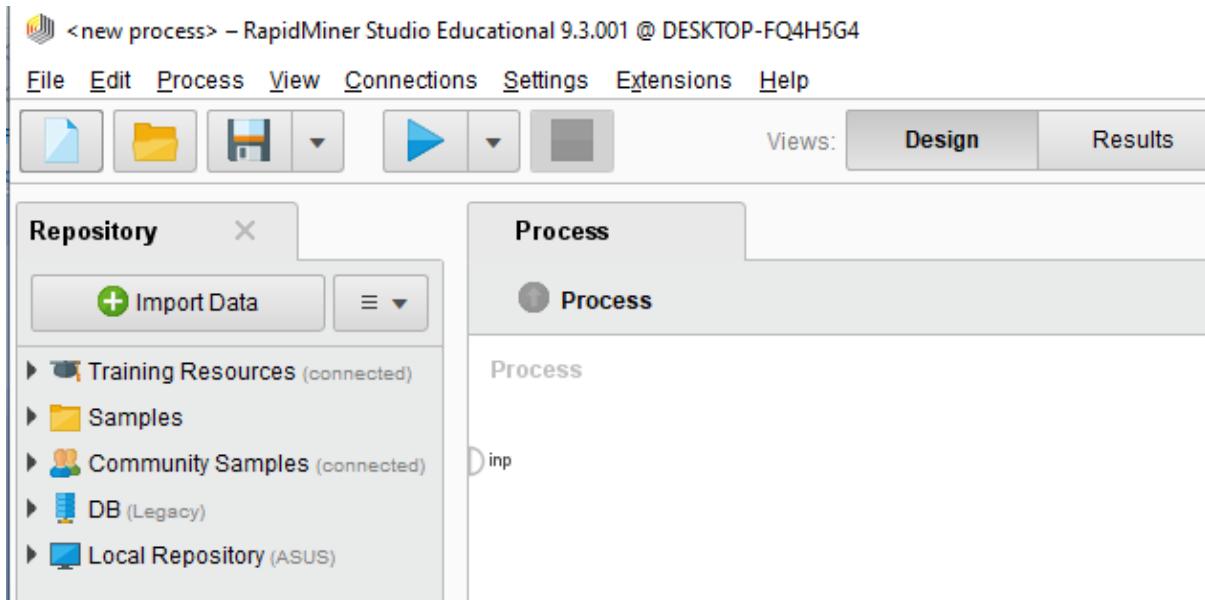
ARFF-Viewer - D:\UNIVERSITAS\TUGAS_SEMESTER5\Prak_DataMining\HasilPrediksi.arff						
File Edit View						
HasilPrediksi.arff						
Relation: Cuaca_predicted						
No. 1: Cuaca 2: Suhu 3: Kelembaban_Udara 4: Berangin 5: prediction margin 6: predicted Bermain_Teknis 7: Bermain_Teknis						
1	Cerah	75.0	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

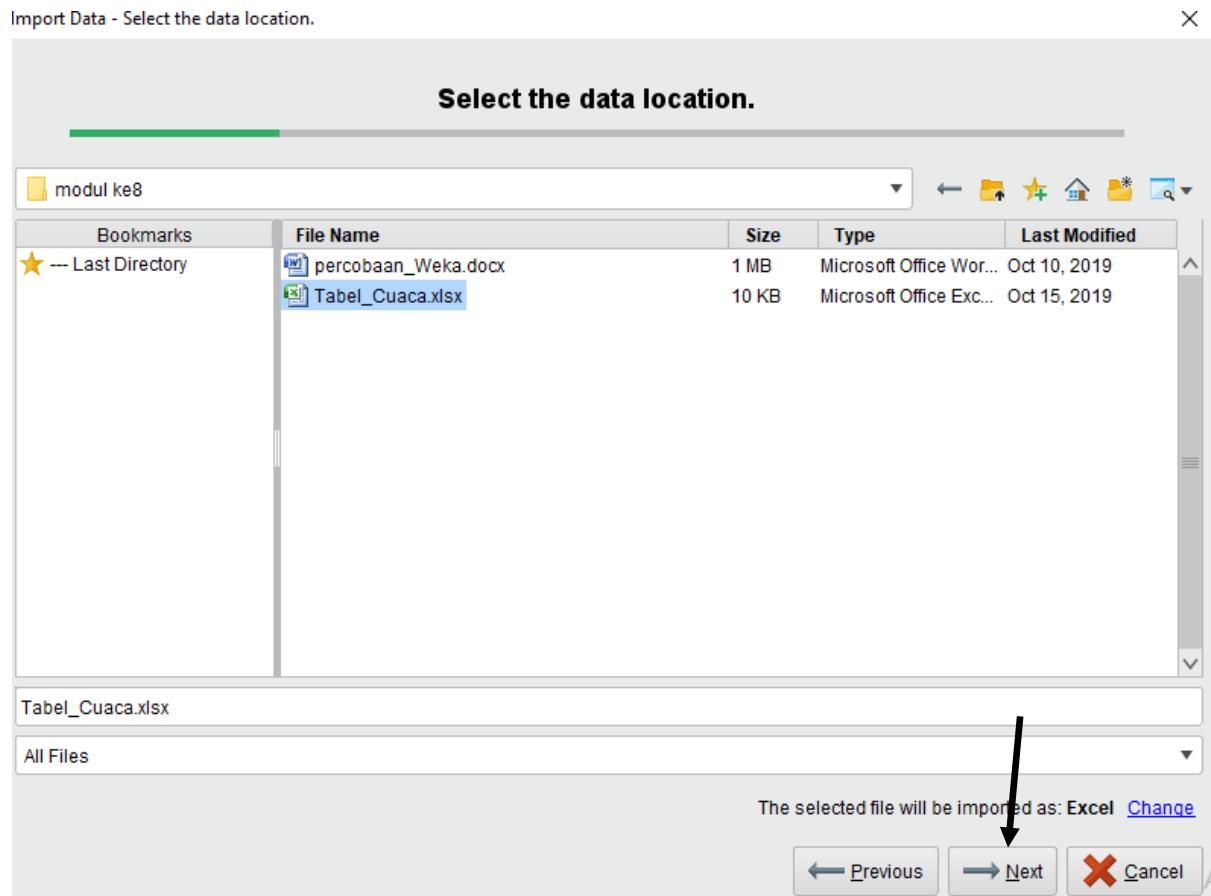
## Percobaan rapidMiner

siapkan table\_Cuaca.xlsx

	A	B	C	D
1	Cuaca	Suhu	Kelembaban_udara	Berangin
2	Cerah	75	65	TIDAK
3	Cerah	80	68	YA
4	Cerah	83	87	YA
5	Mendung	70	96	TIDAK
6	Mendung	68	81	TIDAK
7	Hujan	65	75	YA
8	Hujan	64	85	YA
9				
10				

2. buka aplikasi rapid miner, lalu klik import. Lalu arahkan direktori tempat penyimpanan file pada langkah select the data location. Lalu klik next





3. pastikan sel exel sesuai di langkah select the cells to import

Import Data - Select the cells to import.

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	...	...	...	...	...

← Previous    → Next    ✖ Cancel

4. pada langkah format your colums ubah kolom Bermain\_Tenis dengan tipe data binomial karena hanya ada dua keputusan YA/TIDAK

The screenshot shows the 'Format your columns' dialog in RapidMiner Studio. A context menu is open over the 'Bermain\_Tenis' column, specifically under the 'Change Type' submenu. The 'binominal' option is highlighted. The 'Parameters' panel on the right shows the 'Bermain\_Tenis' column has been set to 'binominal'. The 'no problems.' message at the bottom indicates the changes are successful.

	Cuaca	Suhu	Kelembaban_u...	Berangin	Bermain_Tenis
1	Cerah	85	85	TIDAK	TIDAK
2	Cerah	80	90	YA	TIDAK
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

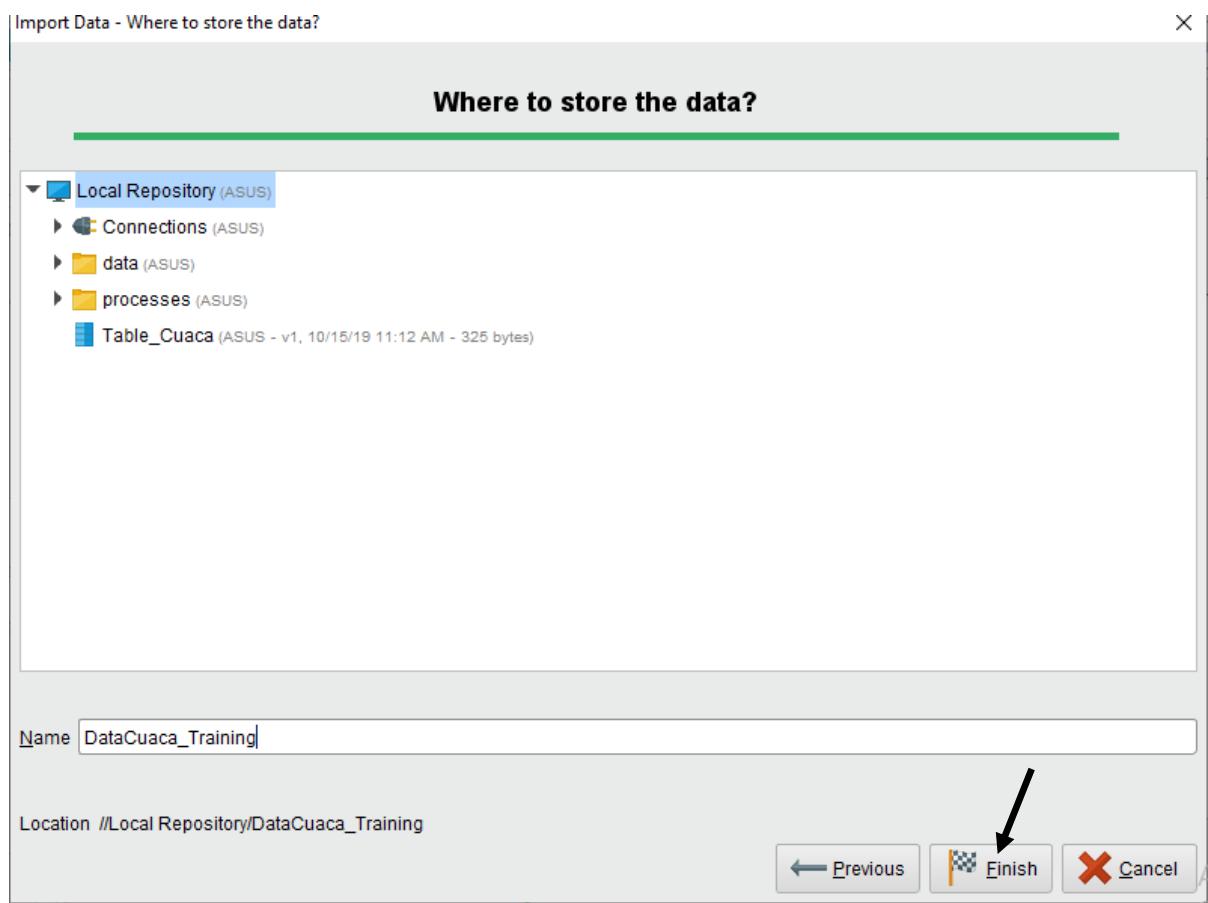
5. ubah juga label pada change Role

The screenshot shows the 'Format your columns' dialog in RapidMiner Studio. A context menu is open over the 'Bermain\_Tenis' column, specifically under the 'Change Role' submenu. A tooltip for 'Change Role' states 'Opens a dialog to change the role.' The 'Parameters' panel on the right shows the 'Bermain\_Tenis' column has been set to 'binominal'. The 'no problems.' message at the bottom indicates the changes are successful.

	Kelembaban_u...	Berangin	Bermain_Tenis
	integer	polynominal	binominal
85	TIDAK	TIDAK	
90	YA	TIDAK	
86	TIDAK	YA	
96	TIDAK	YA	
80	TIDAK	YA	
70	YA	TIDAK	
65	YA	YA	
95	TIDAK	TIDAK	
70	TIDAK	YA	
80	TIDAK	YA	
70	YA	YA	
90	YA	YA	
75	TIDAK	YA	

Ragam	Bermain_Tenis
binominal	binominal label
TIDAK	TIDAK
	TIDAK
TIDAK	YA
TIDAK	YA
TIDAK	YA
TIDAK	TIDAK
	YA
TIDAK	TIDAK

6. simpan dengan nama DataCuaca\_Training dilanjutkan klik tombol finish



7. hasil import file Tabel\_Cuaca.xls akan ditampilkan

ExampleSet (/Local Repository/DataCuaca\_Training)

Open in [Turbo Prep](#) [Auto Model](#)

Filter (14 / 14 examples): all

Row No.	Bermain_Te...	Cuaca	Suhu	Kelembaban...	Berangin
1	TIDAK	Cerah	85	85	TIDAK
2	TIDAK	Cerah	80	90	YA
3	YA	Mendung	83	86	TIDAK
4	YA	Hujan	70	96	TIDAK
5	YA	Hujan	68	80	TIDAK
6	TIDAK	Hujan	65	70	YA
7	YA	Mendung	64	65	YA
8	TIDAK	Cerah	72	95	TIDAK
9	YA	Cerah	69	70	TIDAK
10	YA	Hujan	75	80	TIDAK
11	YA	Cerah	75	70	YA
12	YA	Mendung	72	90	YA
13	YA	Mendung	81	75	TIDAK
14	TIDAK	Hujan	71	91	YA

8. kembali ke jendela design perspective dengan shorcuit tombol F8

Process

Process

100%

Process

inp

Your process looks empty.  
Add some data first.  
Drag data or operators here.

Leverage the Wisdom of Crowds to get operator recommendations based on your process design!

Activate Wisdom of Crowds

9. lakukan hal yg sama untuk data testing yg diambil dari Table\_Cuaca.xls pda sheet 2. Dengan mengulang langkah 2. Pastikan semua variable data testing terpilih 4 variabel, beda pada langkah ini tidak ada variable yg diubah tipe label

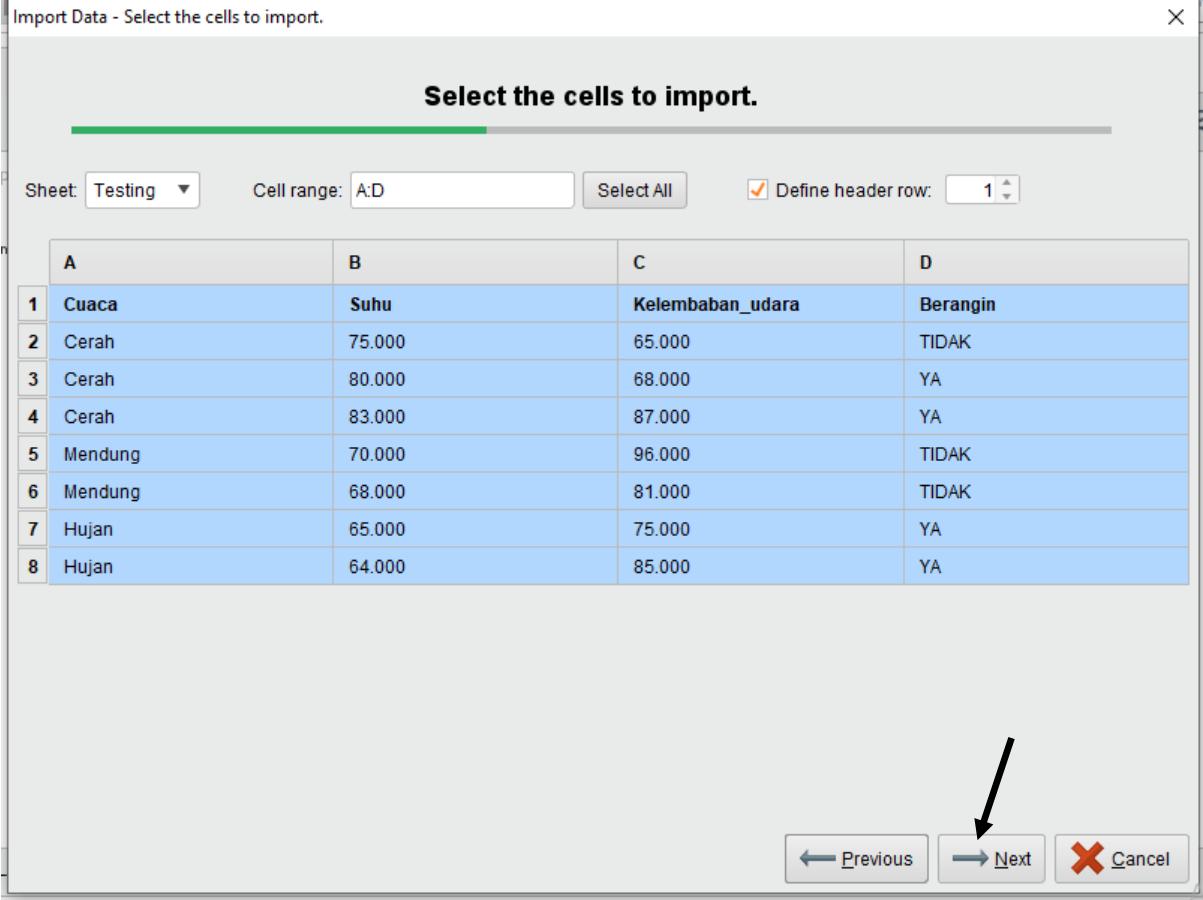
Import Data - Select the cells to import.

Select the cells to import.

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

	A	B	C	D
1	Cuaca	Suhu	Kelembaban_udara	Berangin
2	Cerah	75.000	65.000	TIDAK
3	Cerah	80.000	68.000	YA
4	Cerah	83.000	87.000	YA
5	Mendung	70.000	96.000	TIDAK
6	Mendung	68.000	81.000	TIDAK
7	Hujan	65.000	75.000	YA
8	Hujan	64.000	85.000	YA

← Previous → Next ✖ Cancel



Import Data - Format your columns.

### Format your columns.

Replace errors with missing values (i)

	Cuaca polynominal	Suhu integer	Kelembaban_udara integer	Berangin polynominal
1	Cerah	75	65	TIDAK
2	Cerah	80	68	YA
3	Cerah	83	87	YA
4	Mendung	70	96	TIDAK
5	Mendung	68	81	TIDAK
6	Hujan	65	75	YA
7	Hujan	64	85	YA

 no problems.

 Previous  Next  Cancel

Result History

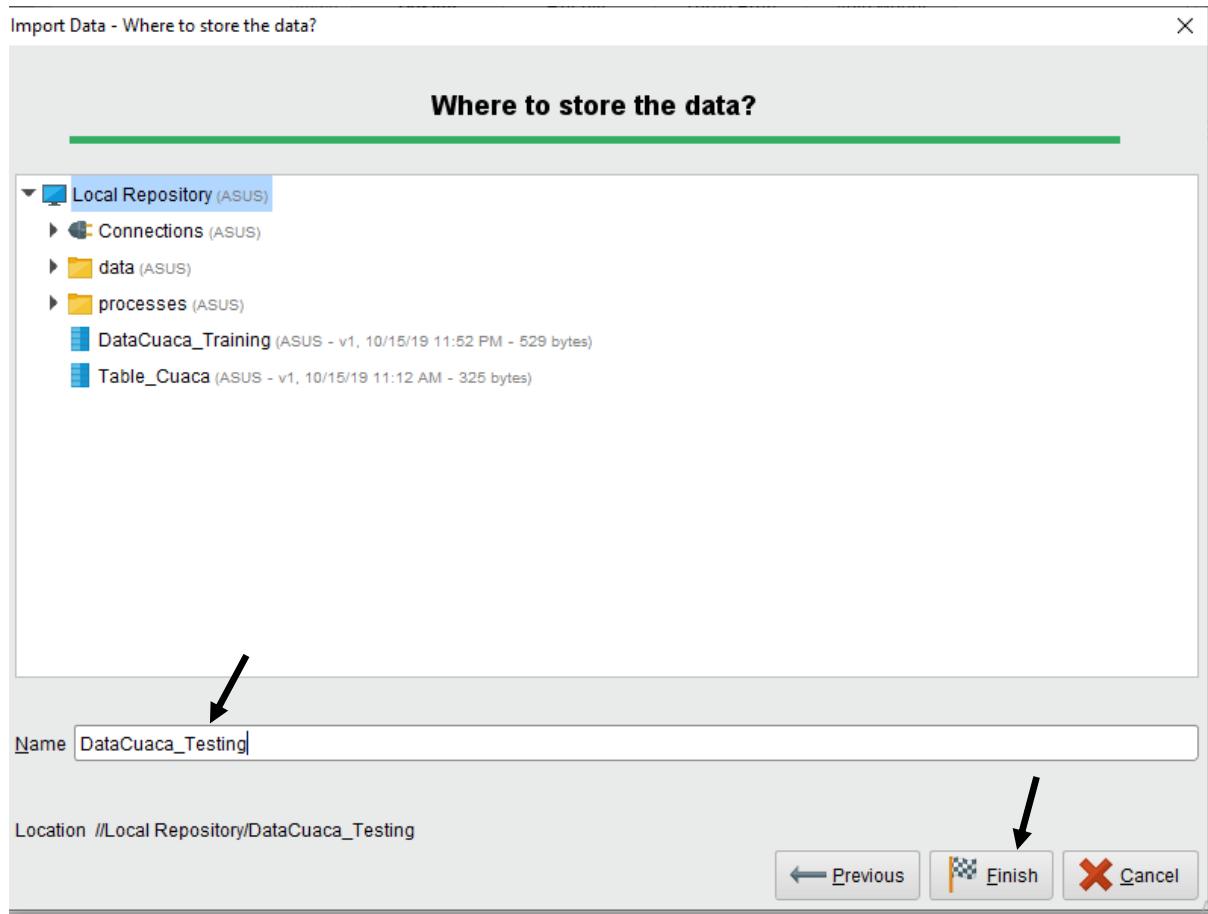
ExampleSet (/Local Repository/DataCuaca\_Testing)  ExampleSet (/Local

Open in  Turbo Prep  Auto Model

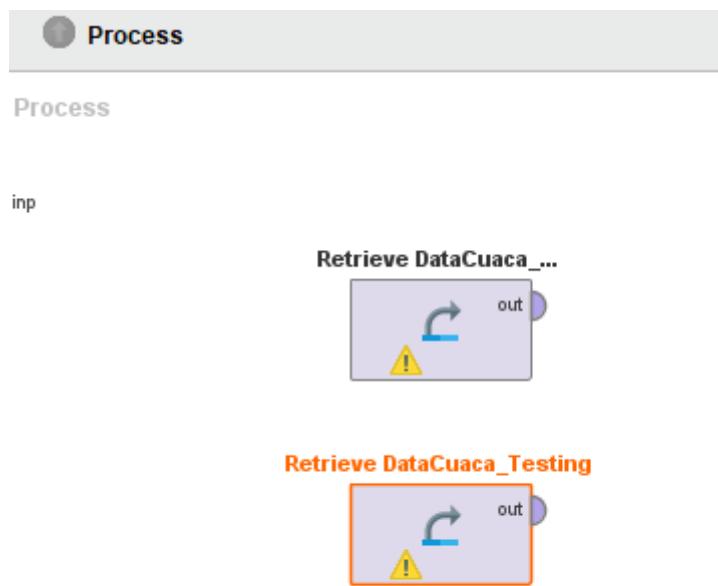
Row No.	Cuaca	Suhu	Kelembaban...	Berangin
1	Cerah	75	65	TIDAK
2	Cerah	80	68	YA
3	Cerah	83	87	YA
4	Mendung	70	96	TIDAK
5	Mendung	68	81	TIDAK
6	Hujan	65	75	YA
7	Hujan	64	85	YA

Data   
 Statistics   
 Visualizations   
 Annotations 

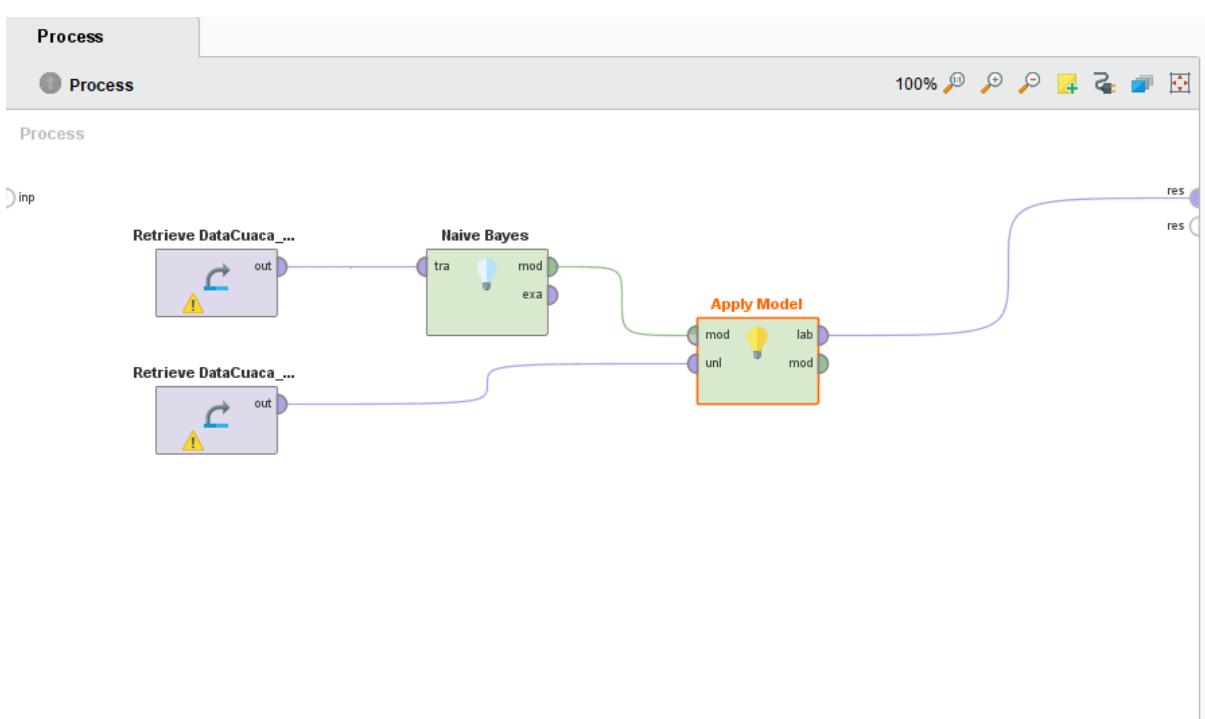
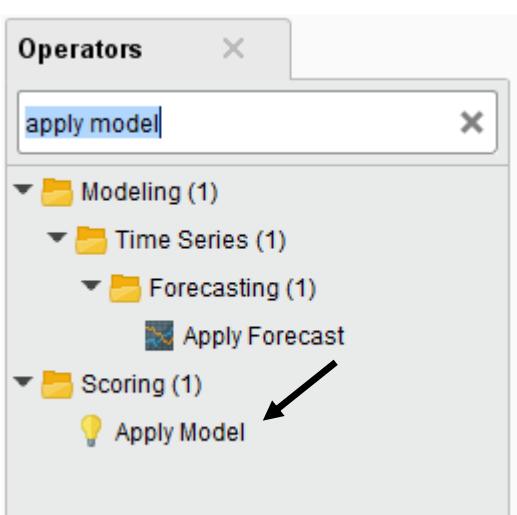
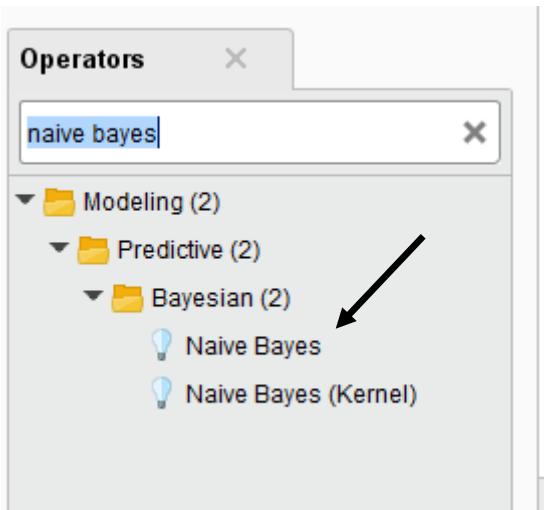
10. simpan dengan nama DataCuaca\_Testing



11. lalu buat design naive bayes. Drag DataCuaca\_Training dan DataCuaca\_Testing ke dalam jendela proces view



12. Masukan juga operator naive bayes dan apply model ke dalam process view. Hubungkan konektor masing masing data terhadap operator seperti gambara



13. jalankan proses naive bayes dengan menekan tombol Run

14. perhatikan hasil proses klasifikasi naive bayes. Pada tab Data, dapat kita lihat hasil prediksi terhadap data testing serta tingkat confidence nilai kelas pada masing masing data

Result History		ExampleSet (Apply Model)					
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
6	YA	0.371	0.629	Hujan	65	75	YA
7	TIDAK	0.568	0.432	Hujan	64	85	YA

Pada tab statistic, dapat dilihat bahwa distribusi nilai kelas pada variable Y (Bermain\_Tenis) rerata nilai confidence sebesar 0,353 untuk nilai TIDAK, dan 0,647 untuk nilai YA

Result History		ExampleSet (//Local Repository/DataCuaca_Testing)				ExampleSet (//Local Repository/DataCuaca_Training)			
Data	Name	Type	Missing	Filter (7 / 7 attributes):		Search for Attributes	▼		
	Binomial	0	Least TIDAK (2)	Most YA (5)	Values YA (5), TIDAK (2)				
Statistics	Real	0	Min 0.007	Max 0.856	Average 0.353				
Visualizations	Real	0	Min 0.144	Max 0.993	Average 0.647				
Annotations	Polynomial	0	Least Mendung (2)	Most Cerah (3)	Values Cerah (3), Hujan (2), ...[ more]				
	Integer	0	Min 64	Max 83	Average 72.143				
	Integer	0	Min 65	Max 96	Average 79.571				
	Polynomial	0	Least TIDAK (3)	Most YA (4)	Values YA (4), TIDAK (3)				

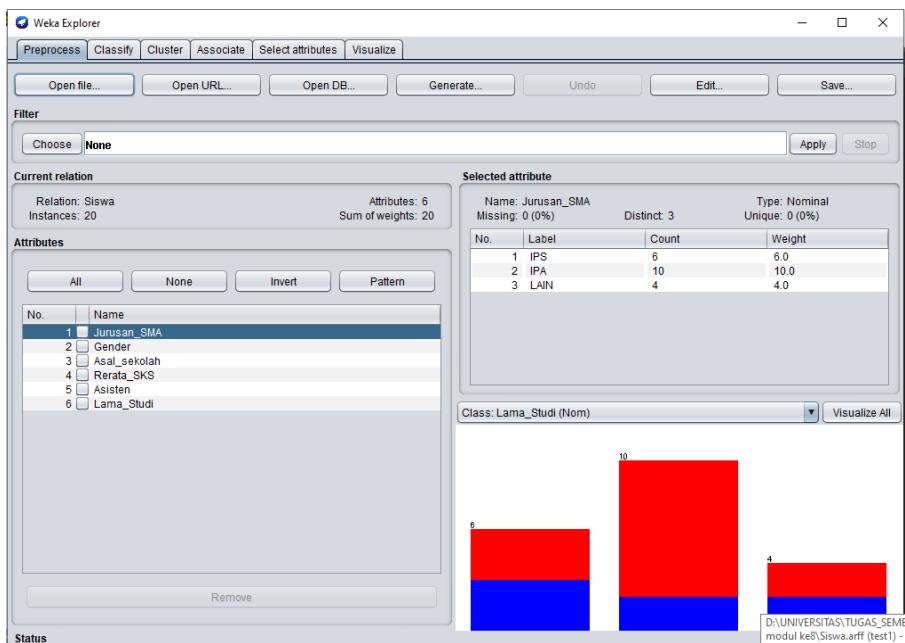
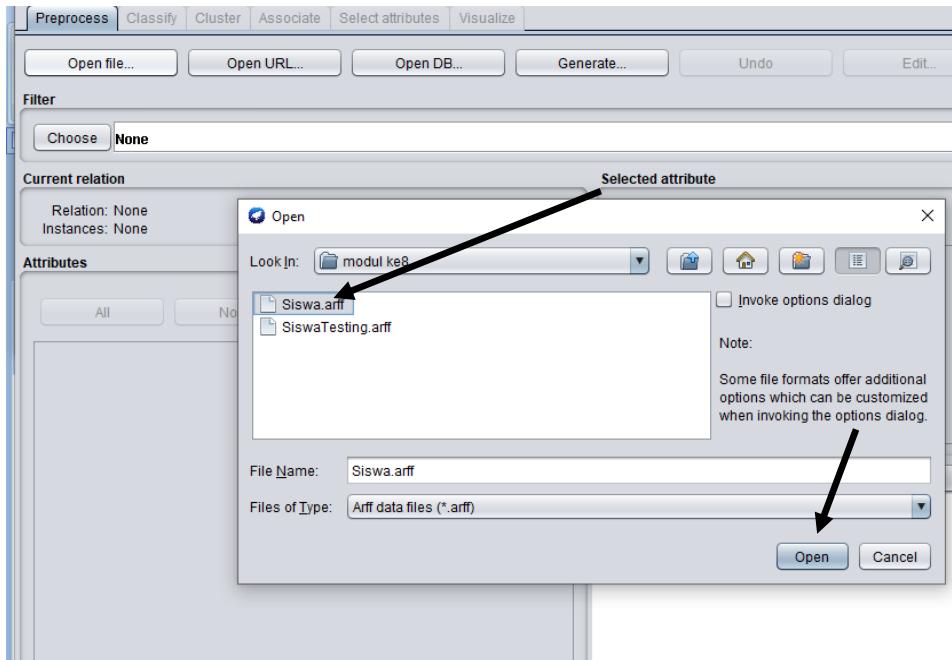
## TUGAS

- persiapkan file Siswa.arff
- buatlah sebuah data testing dengan format arff sebagai data yang akan diprediksi dengan memiliki variabel variable independen dan variable dependen yang sama.
- simpan dengan nama SiswaTesting.arff

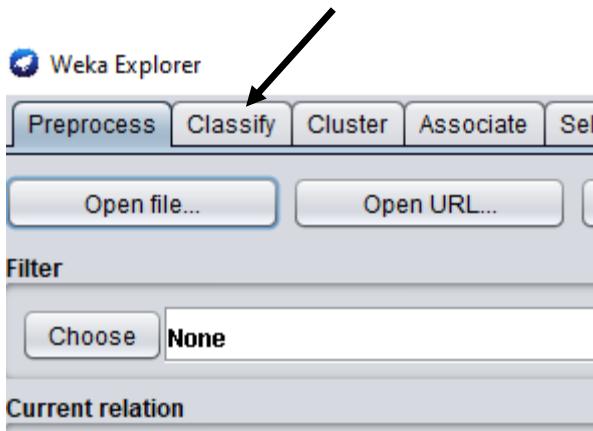
```
1 @relation Siswa
2 @attribute Jurusan_SMA {IPS, IPA, LAIN}
3 @attribute Gender {PRIA, WANITA}
4 @attribute Asal_sekolah {SURAKARTA, LUAR}
5 @attribute Rerata_SKS real
6 @attribute Asisten {YA, TIDAK}
7 @attribute Lama_Studi {TERLAMBAT, TEPAT}
8
9
10 @data
11
12 LAIN, WANITA, SURAKARTA, 18, TIDAK, ?
13 IPA, PRIA, SURAKARTA, 19, YA, ?
14 LAIN, PRIA, SURAKARTA, 19, TIDAK, ?
15 IPS, PRIA, LUAR, 17, TIDAK, ?
16 LAIN, WANITA, SURAKARTA, 17, TIDAK, ?
17 IPA, WANITA, LUAR, 18, YA, ?
18 IPA, PRIA, SURAKARTA, 18, TIDAK, ?
19 IPA, PRIA, SURAKARTA, 19, TIDAK, ?
20 IPS, PRIA, LUAR, 18, TIDAK, ?
21 LAIN, WANITA, SURAKARTA, 18, TIDAK, ?|
```

4. jika telah selesai membuat buka aplikasi weka, masuk dalam menu Weka explorer
5. buka kembali file Siswa.arff

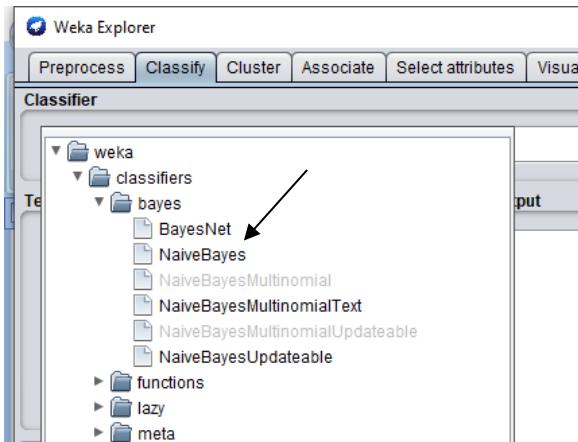
```
◀ ▶ Siswa1.arff × siswa.arff × Siswa.arff ×
1 @relation Siswa
2 @attribute Jurusan_SMA {IPS, IPA, LAIN}
3 @attribute Gender {PRIA, WANITA}
4 @attribute Asal_sekolah {SURAKARTA, LUAR}
5 @attribute Rerata_SKS real
6 @attribute Asisten {YA, TIDAK}
7 @attribute Lama_Studi {TERLAMBAT, TEPAT}
8
9 @data
10 IPS, WANITA, SURAKARTA, 18, TIDAK, TERLAMBAT
11 IPA, PRIA, SURAKARTA, 19, YA, TEPAT
12 LAIN, PRIA, SURAKARTA, 19, TIDAK, TERLAMBAT
13 IPA, PRIA, LUAR, 17, TIDAK, TERLAMBAT
14 IPA, WANITA, SURAKARTA, 17, TIDAK, TEPAT
15 IPA, WANITA, LUAR, 18, YA, TEPAT
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, 20, TIDAK, TEPAT
27 LAIN, PRIA, LUAR, 23, YA, TEPAT
28 IPA, PRIA, SURAKARTA, 21, YA, TEPAT
29 IPS, PRIA, SURAKARTA, 19, TIDAK, TERLAMBAT
```



6. pilih tab classify pada jendela weka



7. klik, lalu akan muncul tombol choose dan klik untuk memilih metode / algoritma naive bayes

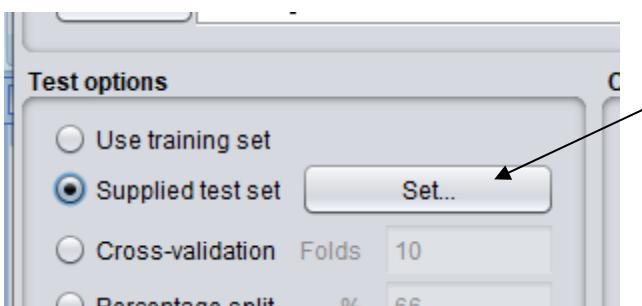


8. menentukan data testing sebagai data yang akan diprediksi variabel independennya. File SiswaTesting.

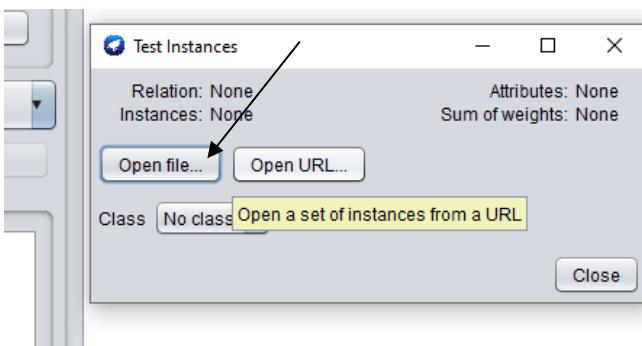
9. pada menu test options terdapat 4 pilihan pengujian yaitu :

- Use training set
- Supplied test set
- Cross-Validation
- Percentage split

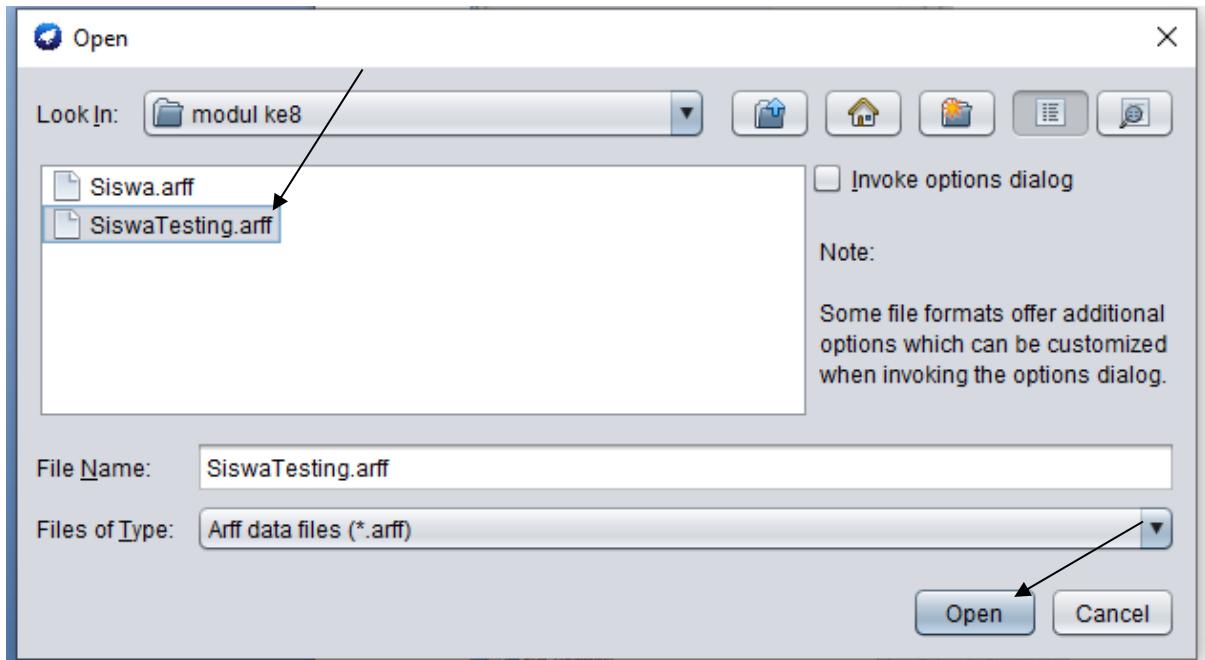
10. gunakan pilihan Supplied test set



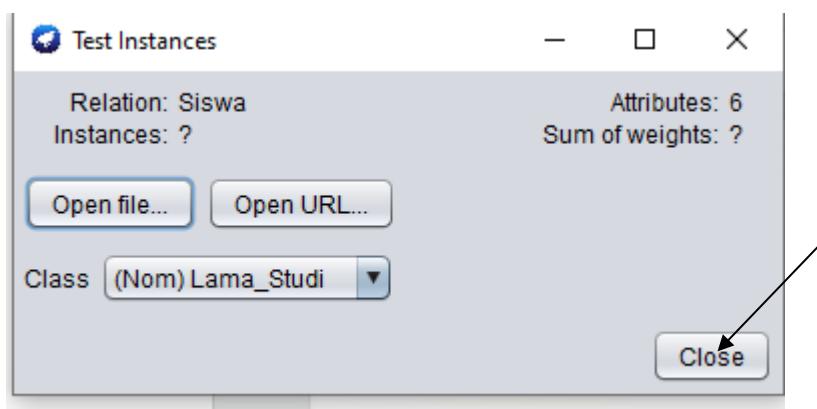
11. sehingga akan muncul jendela test instance Klik open file



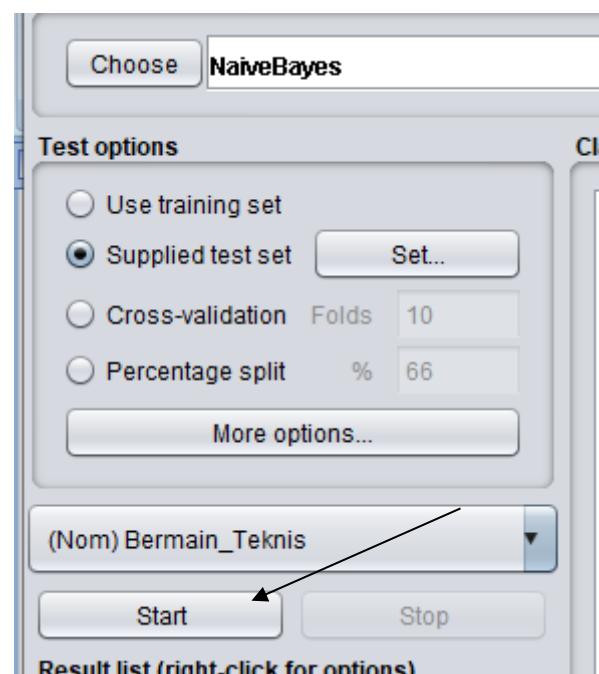
12. pilih file SiswaTesting.arff sebagai data uji. klik open



13. Lalu klik close



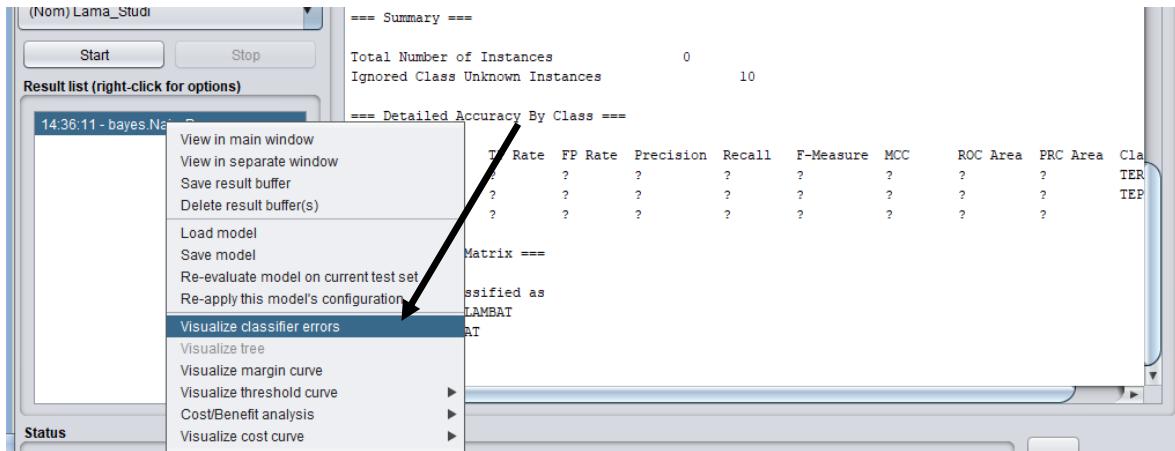
14. klik start untuk memulai proses naive bayes



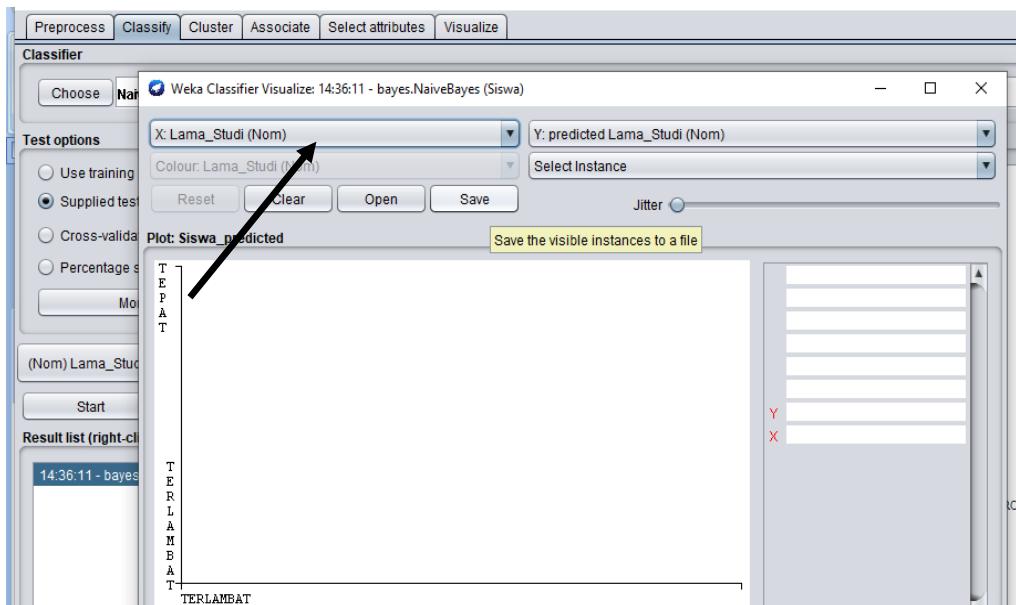
15. jika muncul jendela pesan classifier panel. Kita abaikan dengan mengeklik yes. Sehingga algoritma naive bayes akan diproses

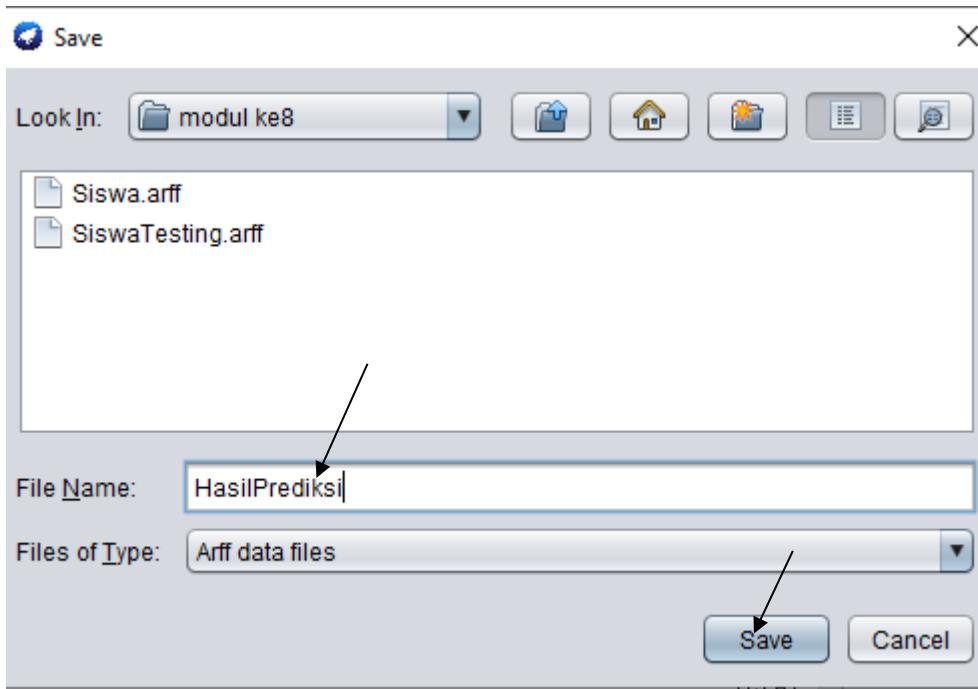
16. kita abaikan nilai nilai yang ditampilkan dalam jendela classifier output

17. untuk melihat hasil prediksi terhadap data uji, lihat nilai classifier errors. Kilik kanan pada hasil proses dalam kontak result list. Pilih menu visual classifier errors

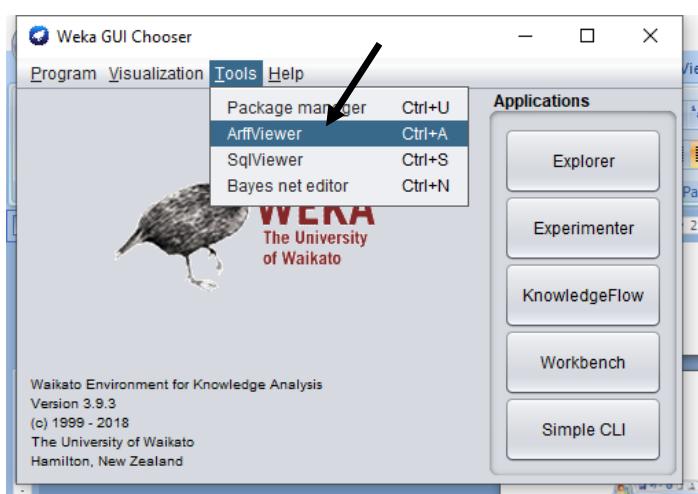


18. pada jendela weka classifier visualize, abaikan hasil apapun yang ditampilkan. Klik save . simpan dengan nama file HasilPrediksi.arff





19. tutup semua jendela. Lalu ke,bali ke weka GUI Chooser. Pilih menu Tools-ArffViewer



20. jendela arff-viewers akan ditampilkan. Buka menu file-open. Tunjukan pada file HasilPrediksi.arff jendela arff-viewers akan ditampilkan. Buka menu file-open. Tunjukan pada file HasilPrediksi.arff yang telah anda simpan. Lihatlah hasil prediksi

ARFF-Viewer - D:\UNIVERSITAS\TUGAS_SEMESTERS\Prak_DataMining\modul ke8\HasilPrediksi.arff							
Relation: Siswa_predicted							
No.	1: Jurusan_SMA	2: Gender	3: Asal_sekolah	4: Rerata_SKS	5: Asisten	6: prediction margin	7: predicted Lama_Studi
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

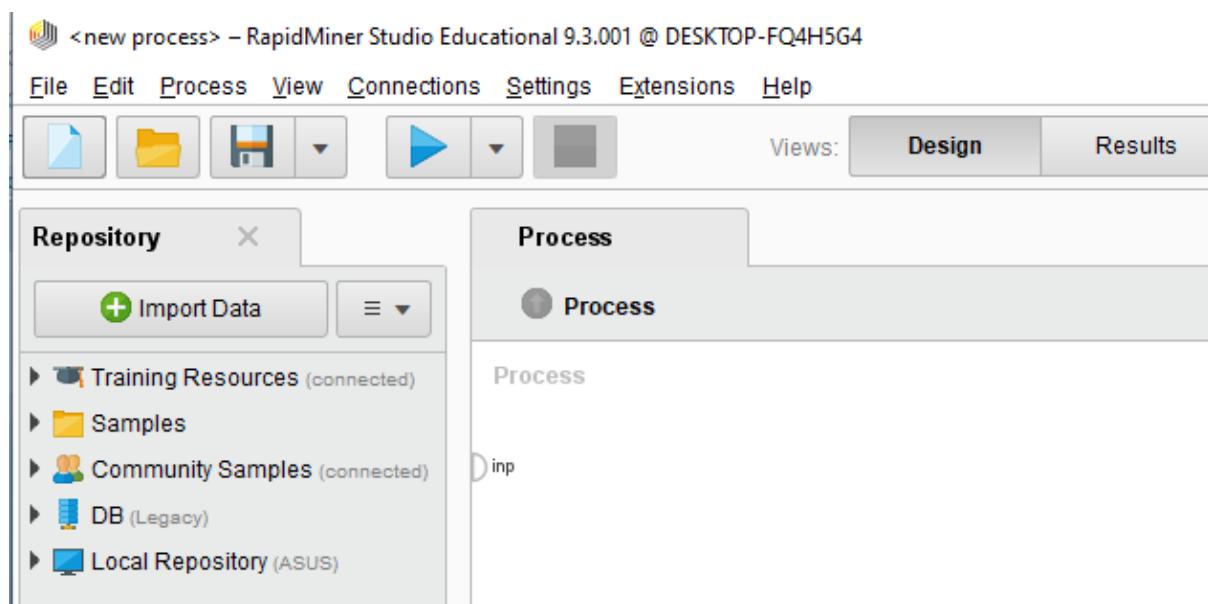


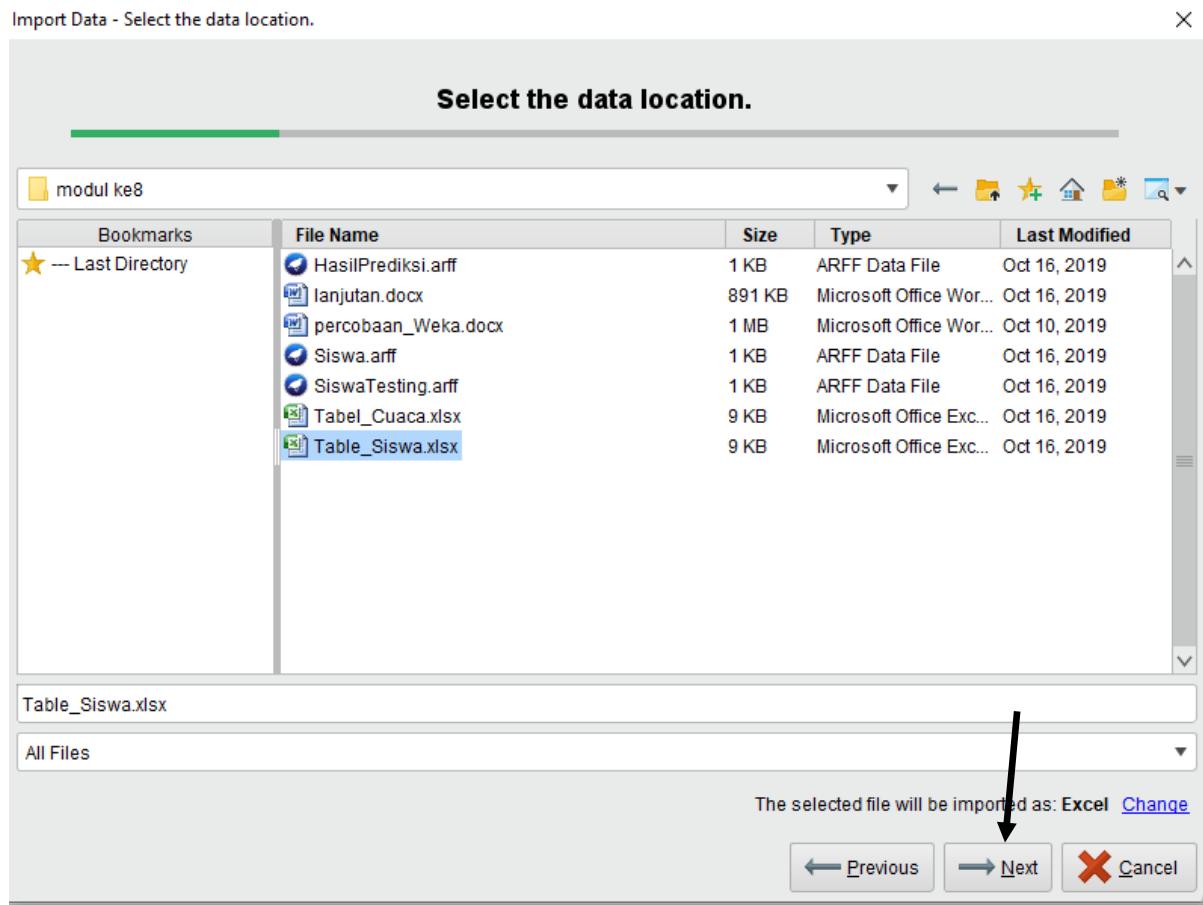
## Menggunakan Naive Bayes dengan RapidMiner

1. siapkan Table\_Siswa.xlsx

	A	B	C	D	E	
1	Jurusan_S	Gender	Asal_Sekolah	Rerata_SK	Asisten	
2	LAIN	WANITA	SURAKARTA	18	TIDAK	
3	IPA	PRIA	SURAKARTA	19	YA	
4	LAIN	PRIA	SURAKARTA	19	TIDAK	
5	IPS	PRIA	LUAR	17	TIDAK	
6	LAIN	WANITA	SURAKARTA	17	TIDAK	
7	IPA	WANITA	LUAR	18	YA	
8	IPA	PRIA	SURAKARTA	18	TIDAK	
9	IPA	PRIA	SURAKARTA	19	TIDAK	
10	IPS	PRIA	LUAR	18	TIDAK	
11	LAIN	WANITA	SURAKARTA	18	TIDAK	
12						

2. buka aplikasi rapid miner, lalu klik import. Lalu arahkan direktori tempat penyimpanan file pada langkah select the data location. Lalu klik next





3. pastikan sel exel sesuai di langkah select the cells to import

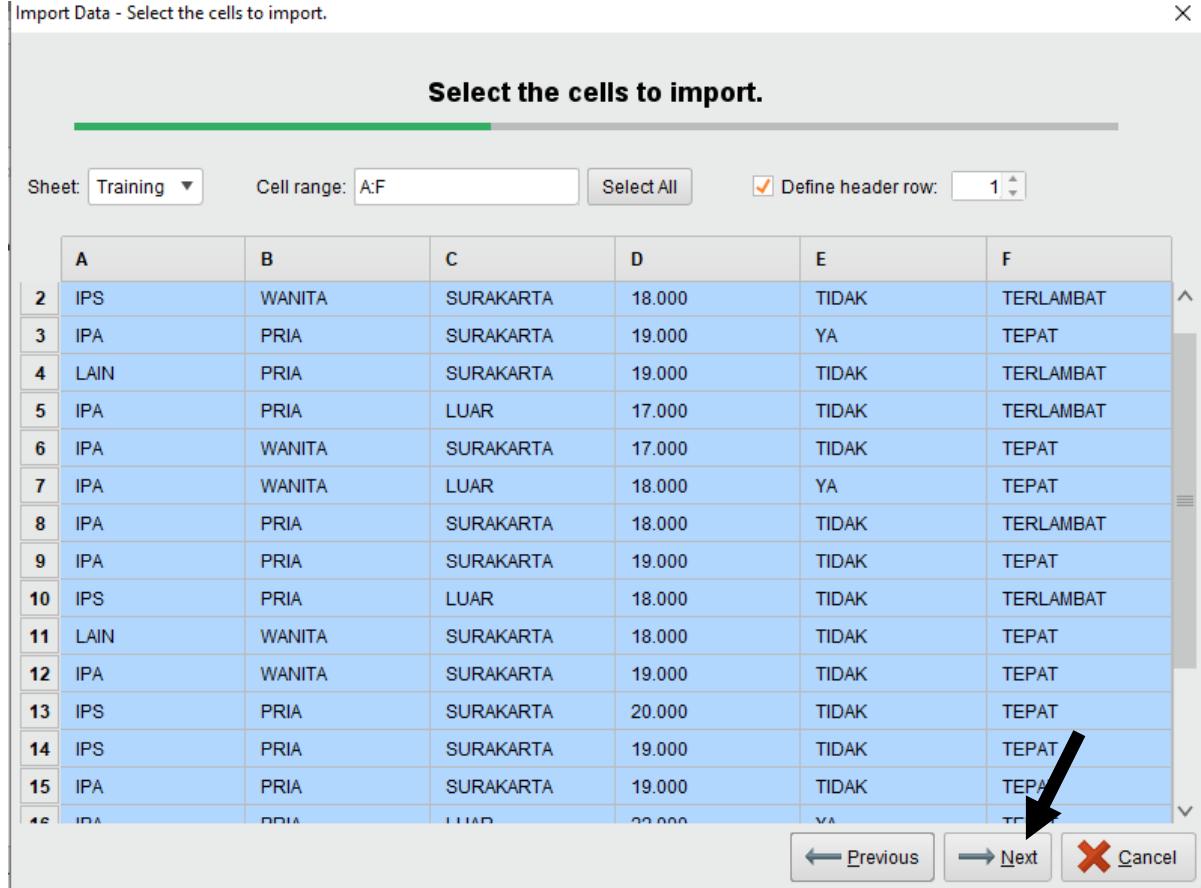
Import Data - Select the cells to import.

Select the cells to import.

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

	A	B	C	D	E	F
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	19.000	TIDAK	TEPAT
16	IPA	PRIA	LUAR	20.000	YA	TEPAT

← Previous    Next →    X Cancel



4. pada langkah format your colums ubah kolom Lama\_Studi dan Asisten dengan tipe data binomial karena hanya ada dua keputusan YA/TIDAK

	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten	Lama_Studi
1	IPS	WANITA	SURAKARTA	18	TIDAK	
2	IPA	PRIA	SURAKARTA	19	YA	
3	LAIN	PRIA	SURAKARTA	19	TIDAK	
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	LUAR	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	SIRAKARTA	19	TIDAK	TEPAT

5. ubah juga label pada change Role

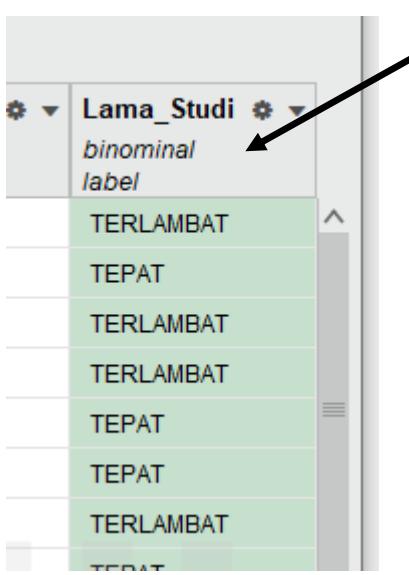
Screenshot of RapidMiner Studio Educational 9.3.001 showing the 'Format your columns' dialog.

The dialog displays a table with 11 rows and 6 columns. The columns are: Jurusan\_SMA (polynominal), Gender (polynominal), Asal\_Sekolah (polynominal), Rerata\_SKS (integer), Asisten (binominal), and Lama\_Studi (binominal label).

The 'Lama\_Studi' column contains values: TERLAMBAT, TEPAT, TERLAMBAT, TERLAMBAT, TEPAT, TEPAT, TERLAMBAT, TERLAMBAT, TEPAT, TEPAT, TEPAT.

A context menu is open over the 'Lama\_Studi' column, with the 'Change Role' option highlighted. A tooltip for 'Change Role' states: 'Opens a dialog to change the role.' Other options in the menu include 'Change Type' and 'Exclude column.'

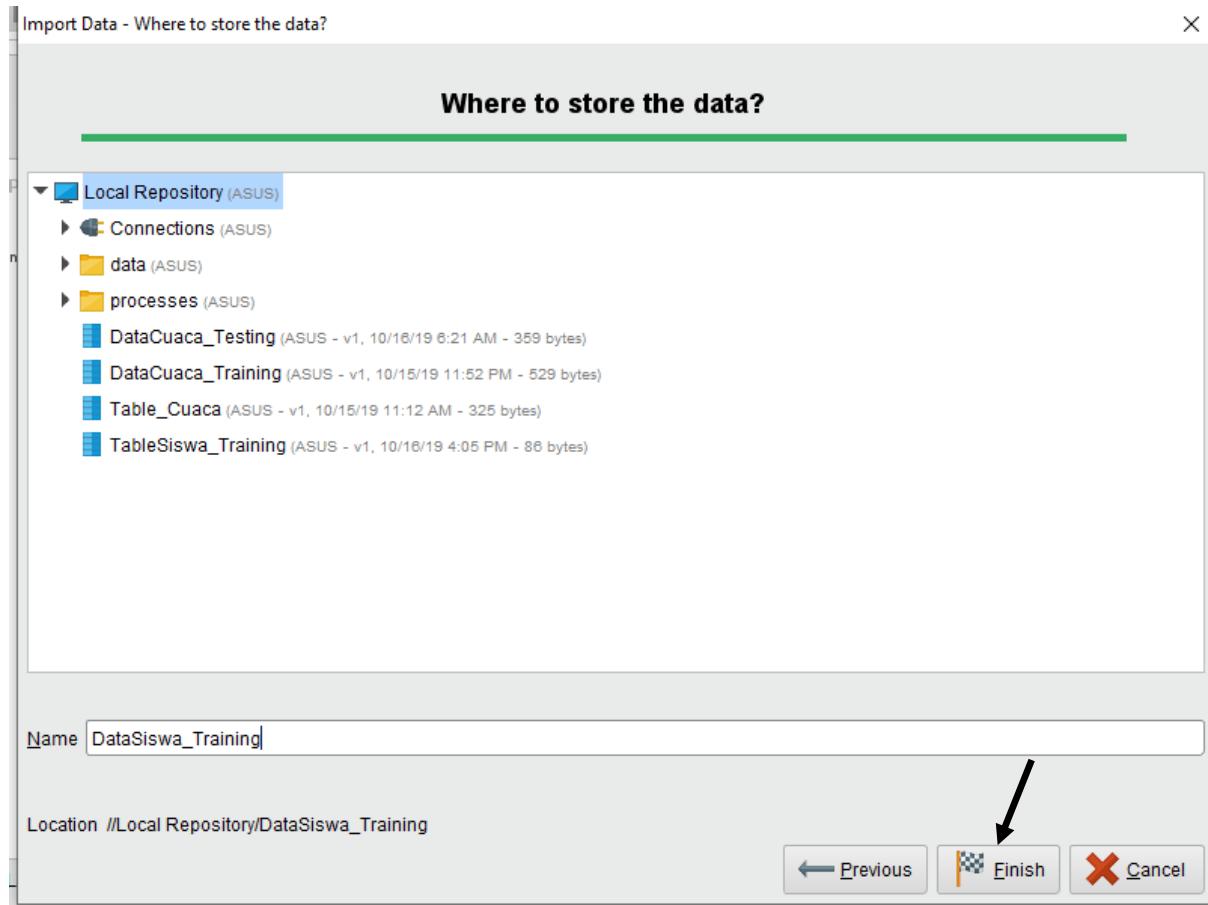
The 'Parameters' panel shows 'Verbosity' set to 'init' and 'logfile' set to 'None'. The 'Help' panel provides information about the 'Process' and 'Synopsis'.



The screenshot shows a list view of the 'Lama\_Studi' column. The column header is 'Lama\_Studi binominal label'. The list contains the following values:

- TERLAMBAT
- TEPAT
- TERLAMBAT
- TERLAMBAT
- TEPAT
- TEPAT
- TERLAMBAT
- TERLAMBAT
- TEPAT
- TEPAT
- TEPAT

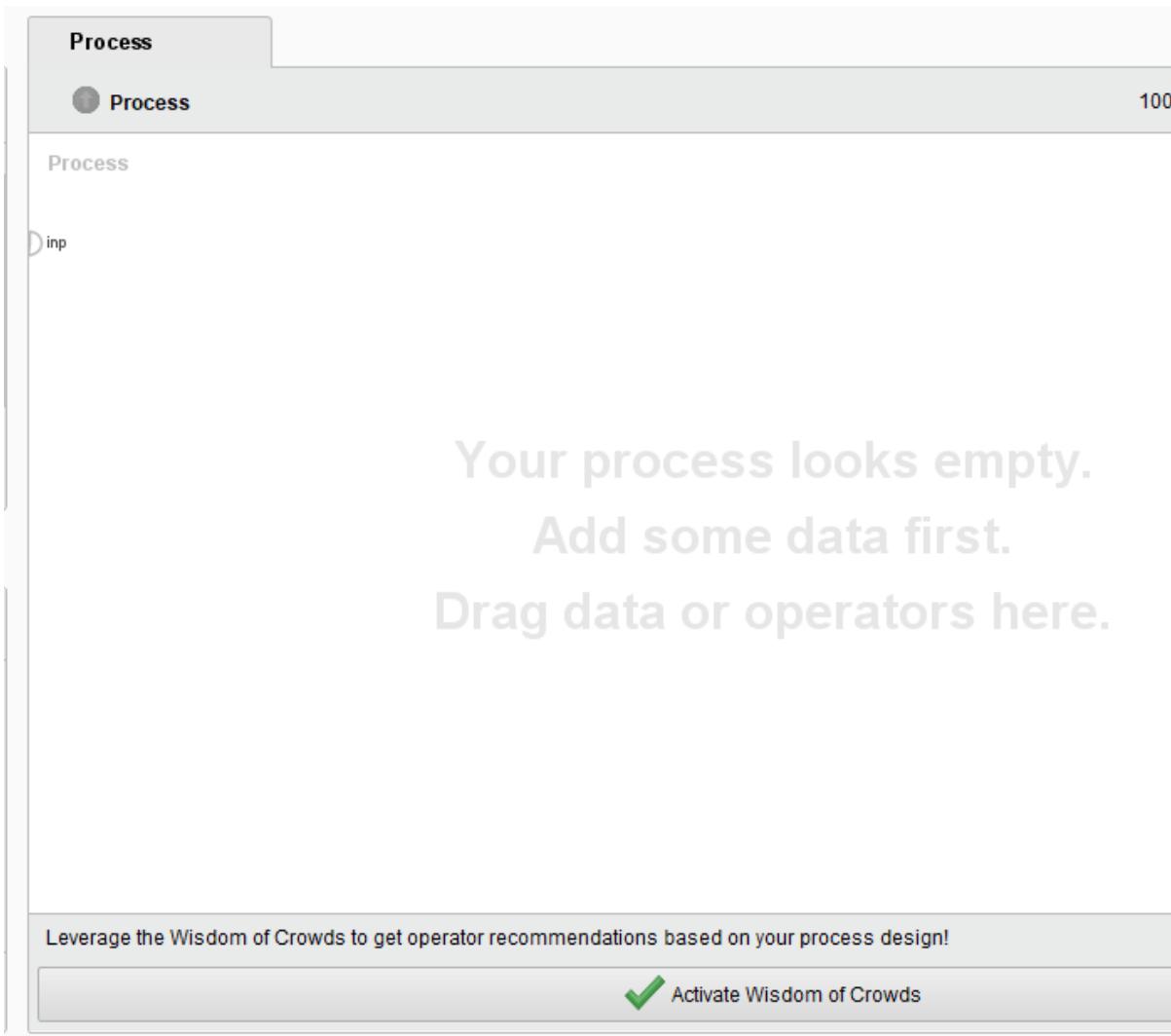
6. simpan dengan nama DataSiswa\_Training dilanjutkan klik tombol finish



7. hasil import file Tabel\_Siswa.xls akan ditampilkan

Row No.	Lama_Studi	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	TERLAMBAT	IPS	WANITA	SURAKARTA	18	TIDAK
2	TEPAT	IPA	PRIA	SURAKARTA	19	YA
3	TERLAMBAT	LAIN	PRIA	SURAKARTA	19	TIDAK
4	TERLAMBAT	IPA	PRIA	LUAR	17	TIDAK
5	TEPAT	IPA	WANITA	SURAKARTA	17	TIDAK
6	TEPAT	IPA	WANITA	LUAR	18	YA
7	TERLAMBAT	IPA	PRIA	SURAKARTA	18	TIDAK
8	TEPAT	IPA	PRIA	SURAKARTA	19	TIDAK
9	TERLAMBAT	IPS	PRIA	LUAR	18	TIDAK
10	TEPAT	LAIN	WANITA	SURAKARTA	18	TIDAK
11	TEPAT	IPA	WANITA	SURAKARTA	19	TIDAK
12	TEPAT	IPS	PRIA	SURAKARTA	20	TIDAK
13	TEPAT	IPS	PRIA	SURAKARTA	19	TIDAK
14	TEPAT	IPA	PRIA	SURAKARTA	19	TIDAK

8. kembali ke jendela design perspective dengan shorcut tombol F8



9. lakukan hal yg sama untuk data testing yg diambil dari Table\_Cuaca.xls pda sheet 2. Dengan mengulang langkah 2. Pastikan semua variable data testing terpilih 4 variabel, beda pada langkah ini tidak ada variable yg diubah tipe label

Import Data - Select the cells to import.

Select the cells to import.

Sheet: Testing ▾ Cell range: A1:A1 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 ✖ Cancel

Import Data - Format your columns.

Format your columns.

Replace errors with missing values ⓘ

	Jurusan_SMA polynomial	Gender polynomial	Asal_Sekolah polynomial	Rerata_SKS integer	Asisten binomial
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 ✖ Cancel

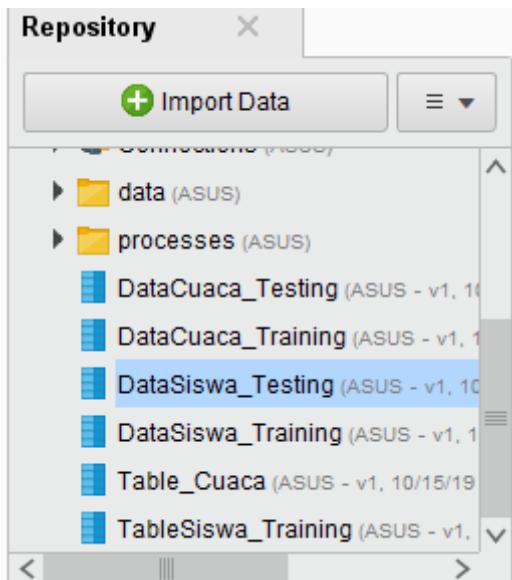
It History

ExampleSet //Local Repository/DataSiswa\_Testing) X

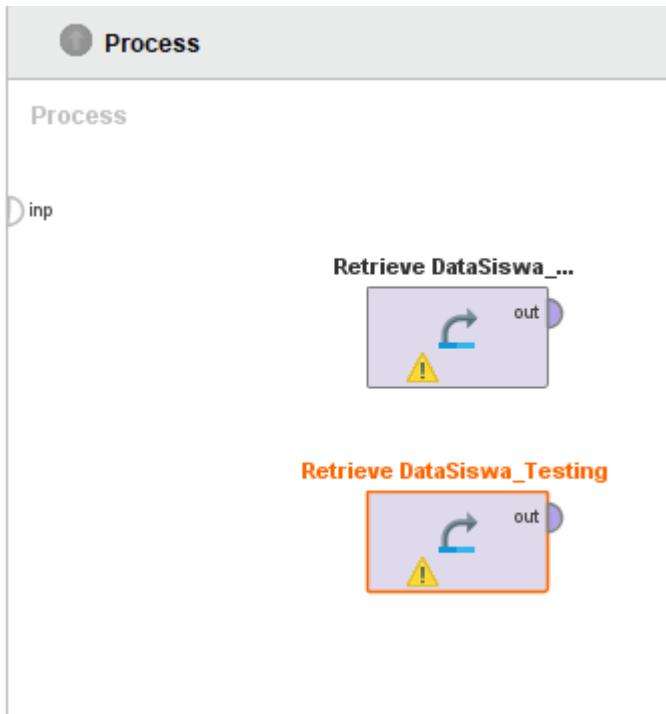
Open in Turbo Prep Auto Model Filter (10 / 10 examples): all

Row No.	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
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

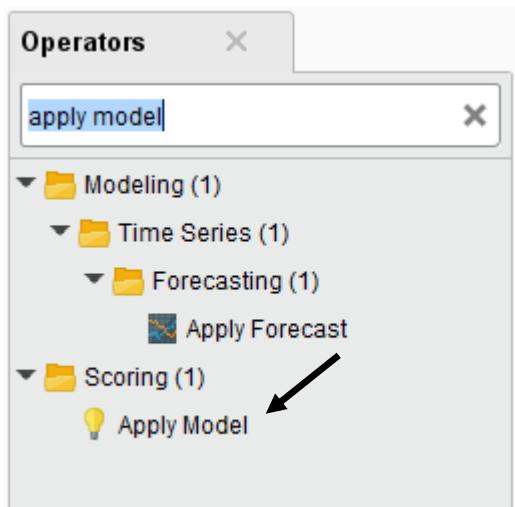
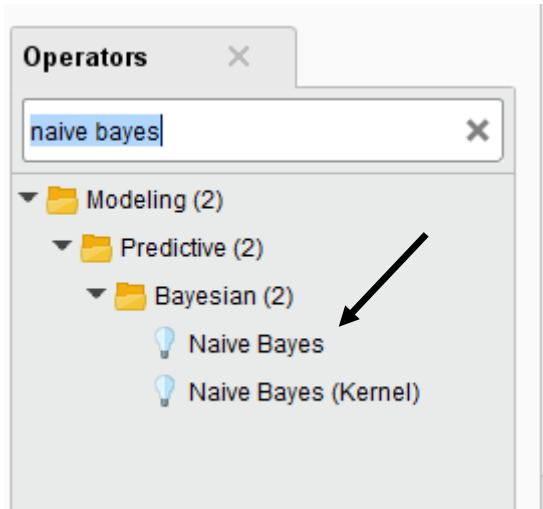
10. simpan dengan nama DataSiswa\_Testing

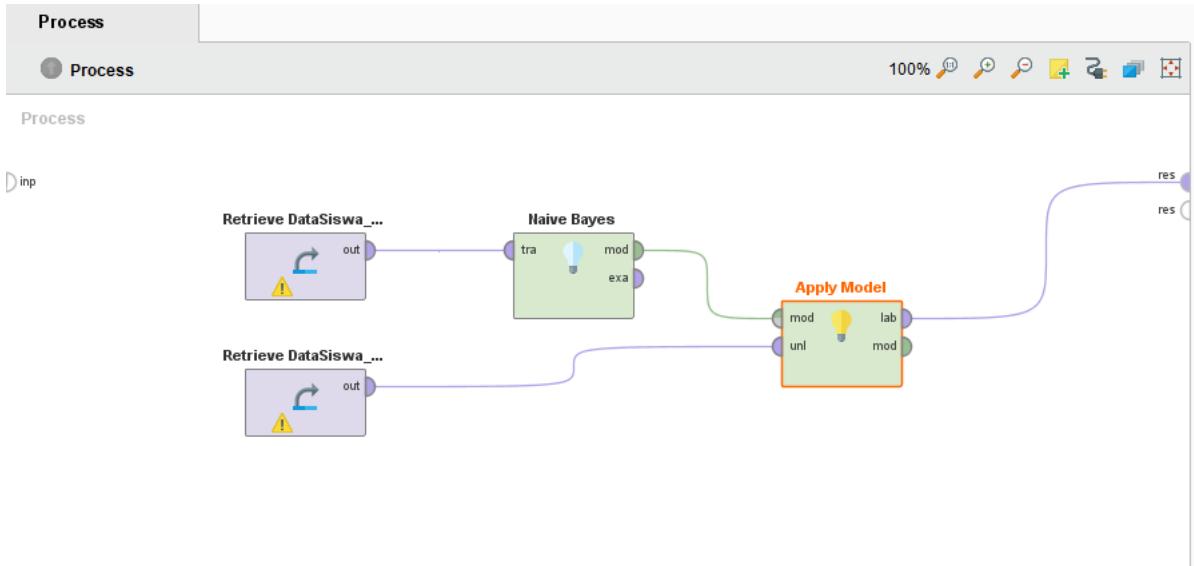


11. lalu buat design naive bayes. Drag DataSiswa\_Training dan DataSiswa\_Testing ke dalam jendela proces view



12. Masukan juga operator naive bayes dan apply model ke dalam process view. Hubungkan konektor masing masing data terhadap operator seperti gambara





13. jalankan proses naive bayes dengan menekan tombol Run

14. perhatikan hasil proses klasifikasi naive bayes. Pada tab Data, dapat kita lihat hasil prediksi terhadap data testing serta tingkat confidence nilai kelas pada masing masing data

ExampleSet (Apply Model)									ExampleSet (/Local Repository/DataSiswa_Testing)	
		Turbo Prep		Auto Model		Filter (10 / 10 examples):			all	
No.	prediction(L...)	confidence(...)	confidence(...)	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten		
	TERLAMBAT	0.648	0.352	LAIN	WANITA	SURAKARTA	18	TIDAK		
	TEPAT	0.005	0.995	IPA	PRIA	SURAKARTA	19	YA		
	TERLAMBAT	0.650	0.350	LAIN	PRIA	SURAKARTA	19	TIDAK		
	TERLAMBAT	0.868	0.132	IPS	PRIA	LUAR	17	TIDAK		
	TERLAMBAT	0.738	0.262	LAIN	WANITA	SURAKARTA	17	TIDAK		
	TEPAT	0.005	0.995	IPA	WANITA	LUAR	18	YA		
	TERLAMBAT	0.547	0.453	IPA	PRIA	SURAKARTA	18	TIDAK		
	TEPAT	0.321	0.679	IPA	PRIA	SURAKARTA	19	TIDAK		
	TERLAMBAT	0.811	0.189	IPS	PRIA	LUAR	18	TIDAK		
	TERLAMBAT	0.648	0.352	LAIN	WANITA	SURAKARTA	18	TIDAK		

Pada tab statistic, dapat dilihat bahwa rerata nilai confidence sebesar 0,524 untuk nilai TERLAMBAT, dan 0,476 untuk nilai TEPAT

ExampleSet (/Local Repository/DataSiswa\_Training)      ExampleSet (/Local Repository/TableSiswa\_Training)      ExampleSet (/Local Repository/DataSiswa\_Testing)

Attribute      ExampleSet (Apply Model)      ExampleSet (/Local Repository/DataSiswa\_Testing)

Name	Type	Missing	Filter (8 / 8 attributes):	Search for Attributes	Filter
Binomial	0	Least TEPAT (3)	Most TERLAMBAT (7)	Values TERLAMBAT (7), TEPAT (3)	
Real	0	Min 0.005	Max 0.868	Average 0.524	
Real	0	Min 0.132	Max 0.995	Average 0.476	
Polynominal	0	Least IPS (2)	Most IPA (4)	Values IPA (4), LAIN (4), ...[1 more]	
Polynominal	0	Least WANITA (4)	Most PRIA (6)	Values PRIA (6), WANITA (4)	
Polynominal	0	Least LUAR (3)	Most SURAKARTA (7)	Values SURAKARTA (7), LUAR (3)	
Integer	0	Min 17	Max 19	Average 18.100	

<      >

Showing attributes 1 - 8      Examples: 10   Special Attributes: 3   Regular Attributes: 5

15. tambahkan data testing berupa nama dewi dan jono

Table\_Siswa.xlsx

	A	B	C	D	E
1	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
2	IPA	WANITA	LUAR	18	TIDAK
3	LAIN	PRIA	SURAKARTA	17	YA
4					
5					

16. lakukan import data

Import Data - Select the cells to import.

Select the cells to import.

Sheet: Sheet3   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 IPA	WANITA	LUAR	18.000	TIDAK
3 LAIN	PRIA	SURAKARTA	17.000	YA

## 17. ubah Asisten menjadi Binomial

Import Data - Format your columns.

**Format your columns.**

Replace errors with missing values (i)

	Jurusan_SMA <i>polynominal</i>	Gender <i>polynominal</i>	Asal_Sekolah <i>polynominal</i>	Rerata_SKS <i>integer</i>	Asisten <i>binominal</i>
1	IPA	WANITA	LUAR	18	TIDAK
2	LAIN	PRIA	SURAKARTA	17	YA

## 18. simpan dengan nama Tugas\_Testing

Import Data - Where to store the data?

**Where to store the data?**

▼ Local Repository (ASUS)

- ▶ Connections (ASUS)
- ▶ data (ASUS)
- ▶ processes (ASUS)
  - DataCuaca\_Testing (ASUS - v1, 10/16/19 6:21 AM - 359 bytes)
  - DataCuaca\_Training (ASUS - v1, 10/15/19 11:52 PM - 529 bytes)
  - DataSiswa\_Testing (ASUS - v1, 10/16/19 4:16 PM - 483 bytes)
  - DataSiswa\_Training (ASUS - v1, 10/16/19 4:10 PM - 673 bytes)
  - Table\_Cuaca (ASUS - v1, 10/15/19 11:12 AM - 325 bytes)
  - TableSiswa\_Training (ASUS - v1, 10/16/19 4:05 PM - 86 bytes)

Name

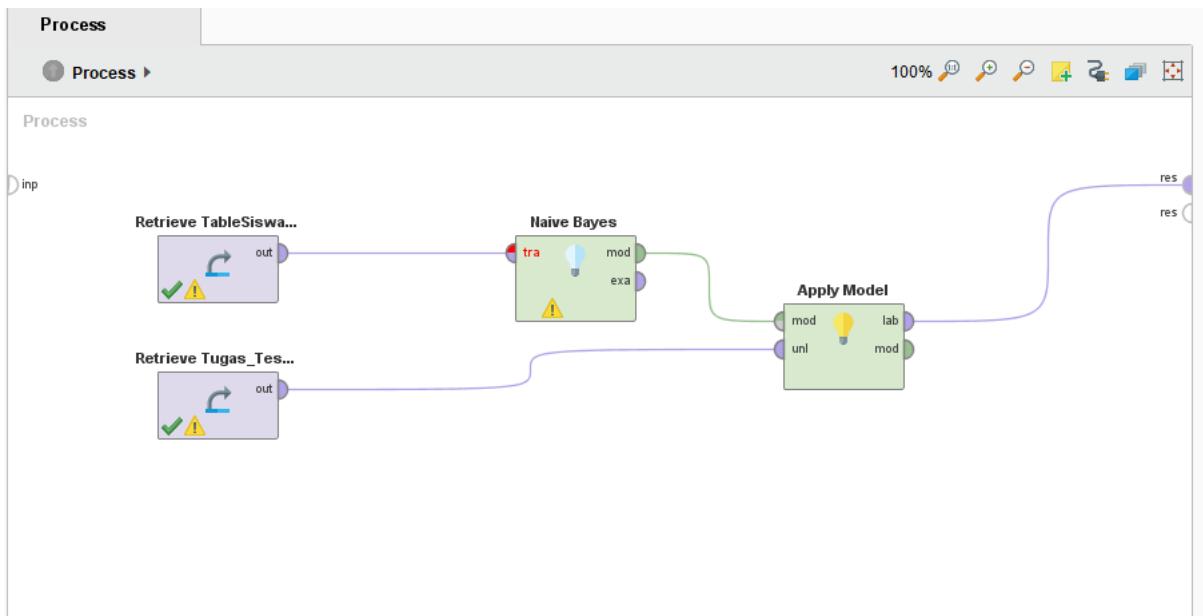
Location //Local Repository/Tugas\_Testing

19. lihat hasil dari penyimpanan

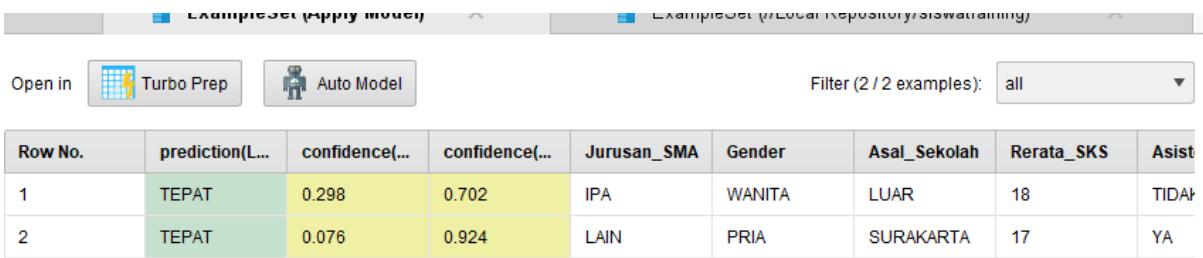


Row No.	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	IPA	WANITA	LUAR	18	TIDAK
2	LAIN	PRIA	SURAKARTA	17	YA

20. lalu drag TableSiswa\_Training dan Tugas\_Testing



21. lalu run, dan liat hasilnya



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

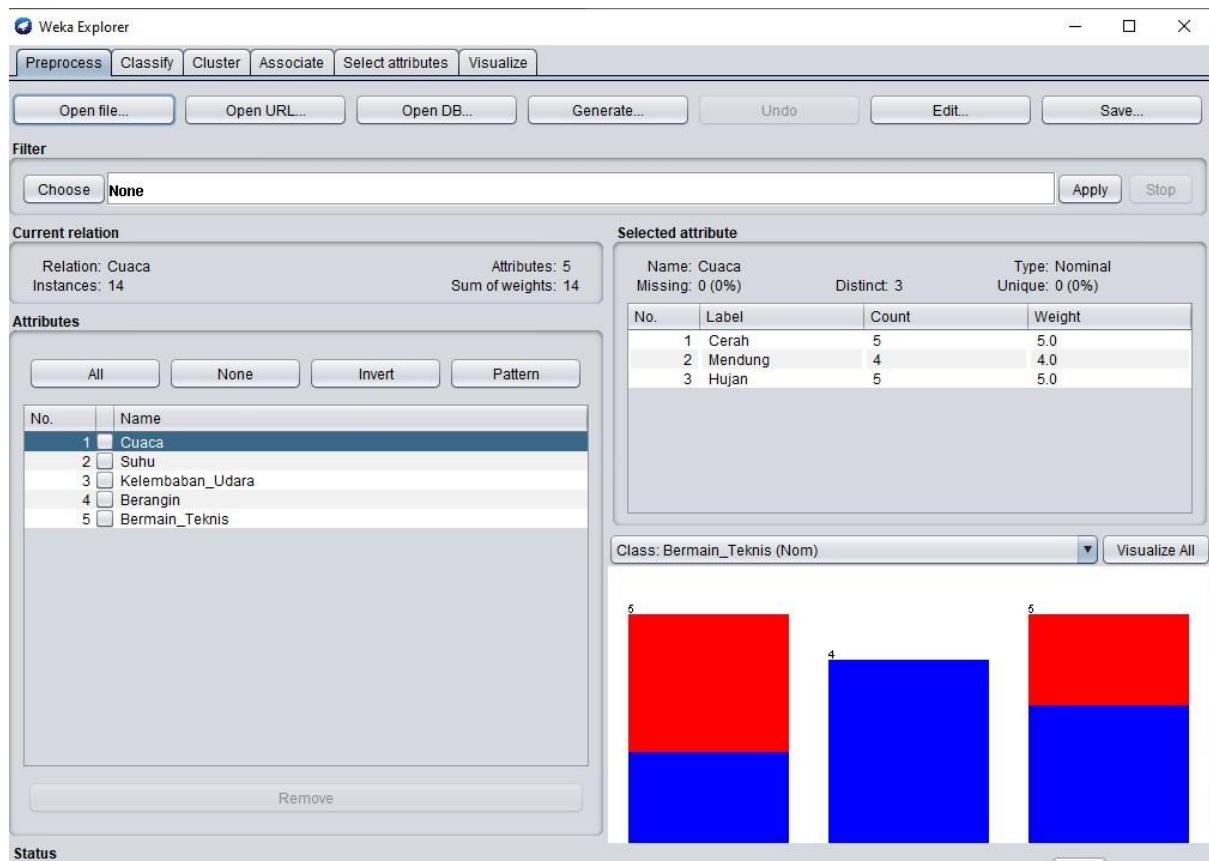
22. Pada tab statistic, dapat dilihat bahwa rerata nilai confidence sebesar 0,187 untuk nilai TERLAMBAT dan 0,813 untuk nilai TEPAT

ExampleSet (/Local Repository/DataSiswa_Training)		ExampleSet (/Local Repository/TableSiswa_Training)	
ExampleSet (/Local Repository/Tugas_Testing)		ExampleSet (/Local Repository/DataSiswa_Testing)	
Result History		ExampleSet (Apply Model)	
	Name	Type	Missing .. Filter (8 / 8 attributes): Search for Attributes
Data	Binomial	0 Least TERLAMBAT (0)	Most TEPAT (2) Values TEPAT (2), TERLAMBAT (0)
Statistics	Real	0 Min 0.076	Max 0.298 Average 0.187
Visualizations	Real	0 Min 0.702	Max 0.924 Average 0.813
Annotations	Polynomial	0 Least LAIN (1)	Most IPA (1) Values IPA (1), LAIN (1)
	Polynomial	0 Least WANITA (1)	Most PRIA (1) Values PRIA (1), WANITA (1)
	Polynomial	0 Least SURAKARTA (1)	Most LUAR (1) Values LUAR (1), SURAKARTA (1)
	Integer	0 Min 17	Max 18 Average 17.500
< Showing attributes 1 - 8		Examples: 2 Special Attributes: 3 Regular Attributes: 5	

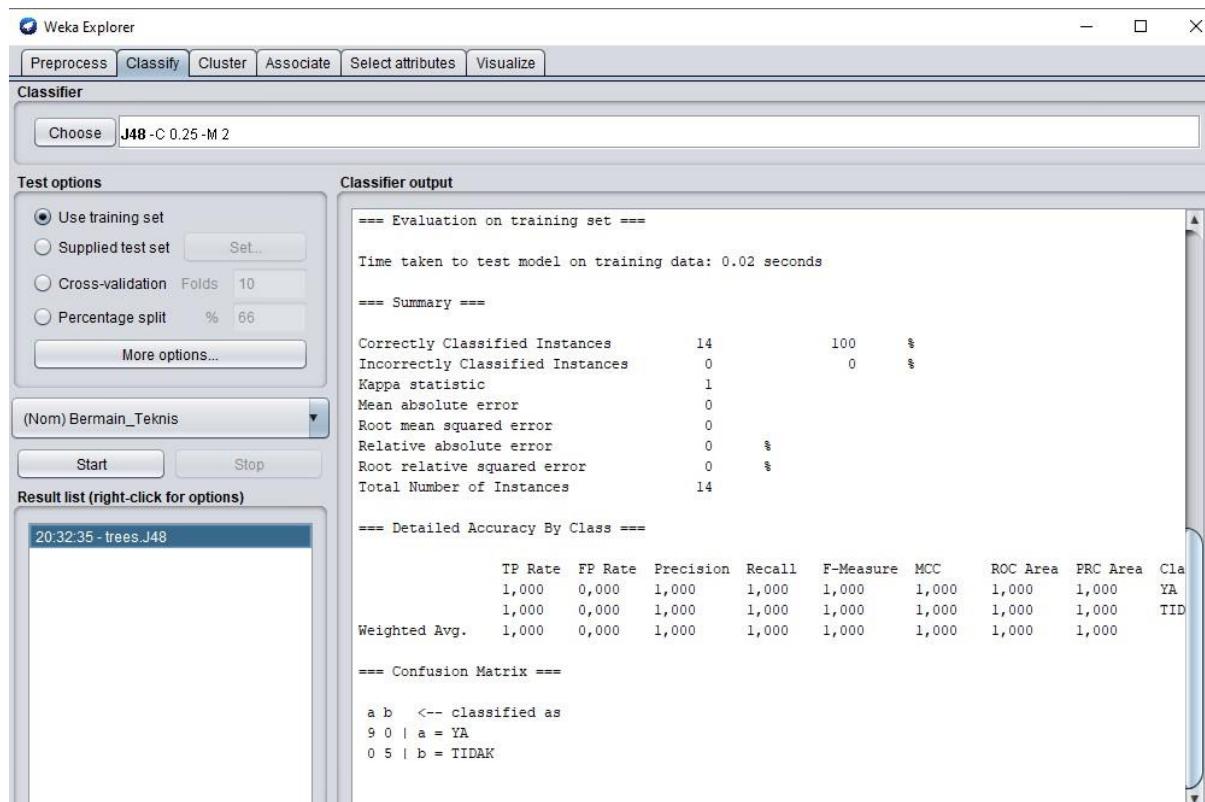
# MODUL 9

## Percobaan weka

### 1. data Cuaca.arff



2.



3.

### Classifier output

```
J48 pruned tree
-----
Cuaca = Cerah
| Kelembaban_Udara <= 75: YA (2.0)
| Kelembaban_Udara > 75: TIDAK (3.0)
Cuaca = Mendung: YA (4.0)
Cuaca = Hujan
| Berangin = YA: TIDAK (2.0)
| Berangin = TIDAK: YA (3.0)

Number of Leaves :      5

Size of the tree :      8

Time taken to build model: 0.02 seconds

==== Evaluation on training set ====

Time taken to test model on training data: 0.02 seconds

==== Summary ===

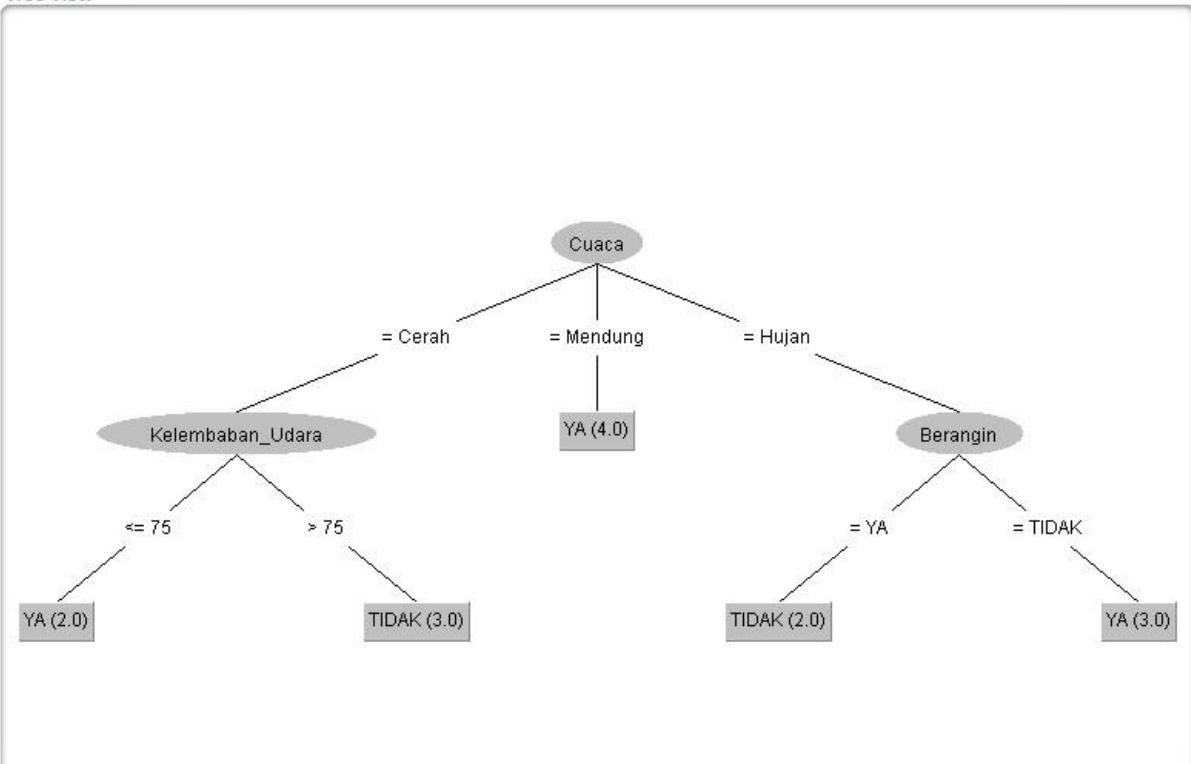
Correctly Classified Instances          14           100      %
Incorrectly Classified Instances        0            0      %
Kappa statistic                         1
Mean absolute error                    0

```

Berdasarkan hasil tersebut dapat diketahui sebagai berikut :

- a) Jumlah simpul daun pada pohon keputusan = 5
  - c) Waktu yang dibutuhkan untuk proses pelatihan = 0,02 detik
  - e) Tingkat ketidak pastian klasifikasi = 0%
- 4.

## Tree View



5. Berdsarkan pohon keputusan tersebut, dapat dilihat jenis-jenis simpul yang ada, sebagai berikut:

- a) Simpul akar = Cuaca
- b) Simpul internal = Kelembaban\_Udara, Berangin
- c) Simpul daun = YA, TIDAK

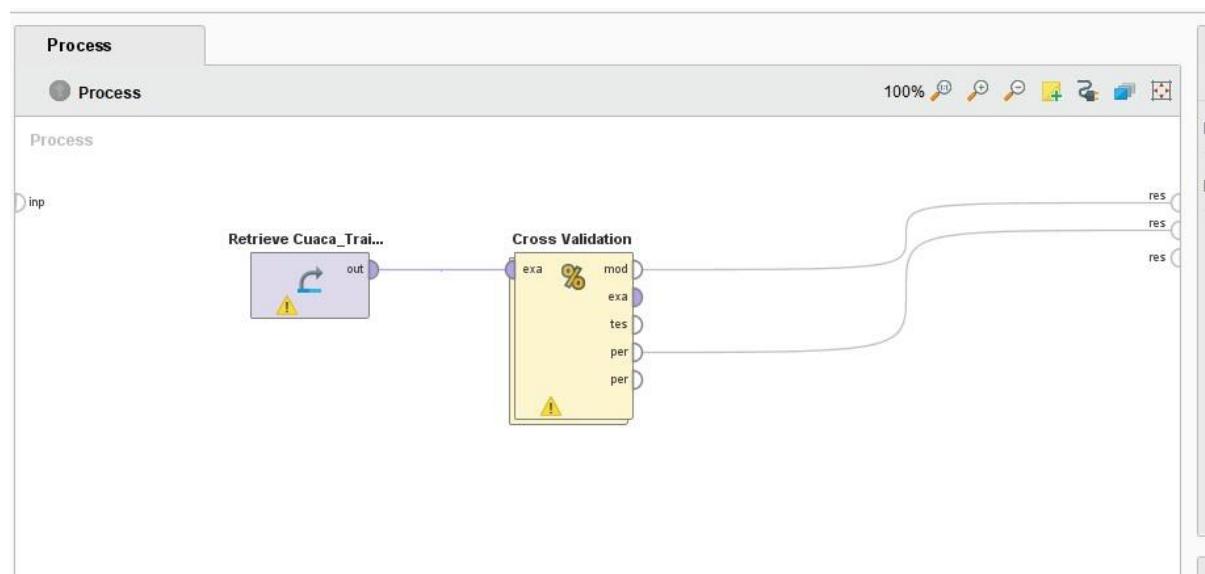
6. Klasifikasi yang terbentuk yaitu :

- a) Seseorang akan bermain tenis (YA) jika kondisi sebagai berikut:
  - i. Cuaca = Cearh, Kelembaban\_Udara <= 75, (nilai atribut lain diabaikan)
  - ii. Cuaca = Mendung, (kondisi lain diabaikan)
  - iii. Ciaca = Hujan, Berangn = TIDAK, (nilai atribut lain diabaikan)

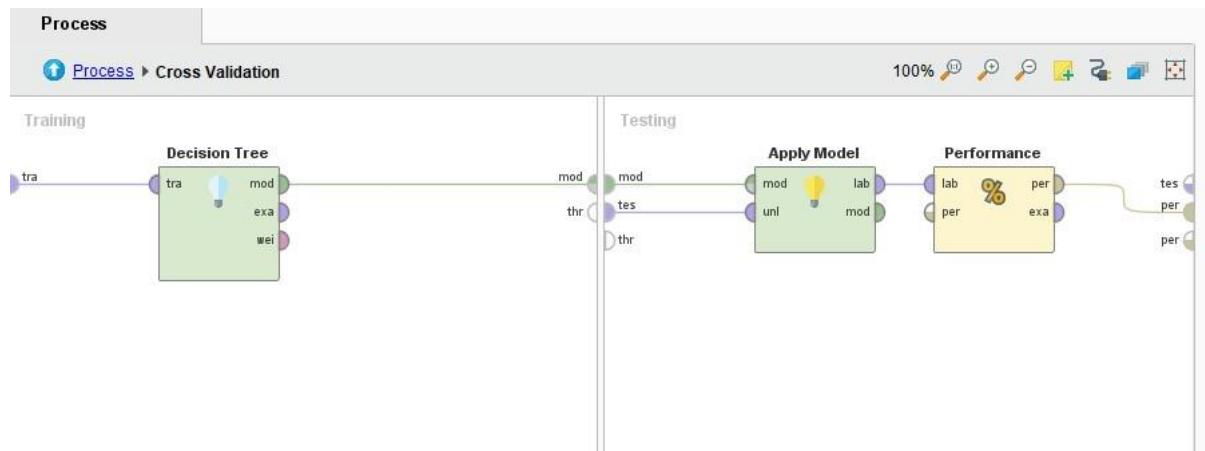
- b) Seseorang tidak akan bermain tenis (TIDAK) jika kondisi sebagai berikut :
- i. Cuaca = Cerah, Kelembabab\_Udara > 75, (nilai atribut lain diabaikan)
  - ii. Cuaca = Hujan, Berangin = YA, (nilai atribut lain diabaikan).

## Rapid Miner

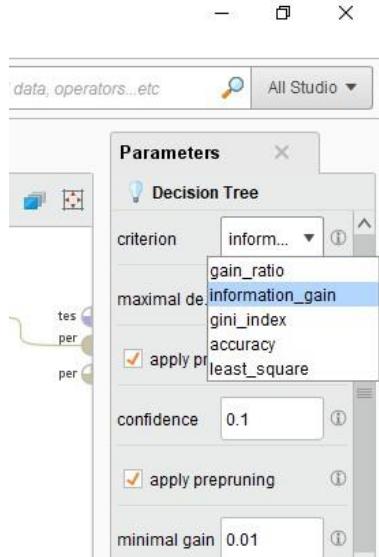
1. kita harus mengimport terlebih dahulu data Cuaca Training. Setelah kita input kita drag Cuaca\_Training ke dalam area process View.



2. Drag pula operator Cross Validation. Lalu klik 2 kali pada cross validation maka akan muncul seperti gambar dibawah ini. Lau masukkan operator Decision Tree ke dalam area Training. Operator Apply Model dan performance dalam area Testing.



3. Lalu setelah itu lik 2 kali pada Decision Tree . dan pastikan kriteria(Criterion) yang dipakai adalah Gain (J48) pada kolom parameter disebelah kanan Process View.

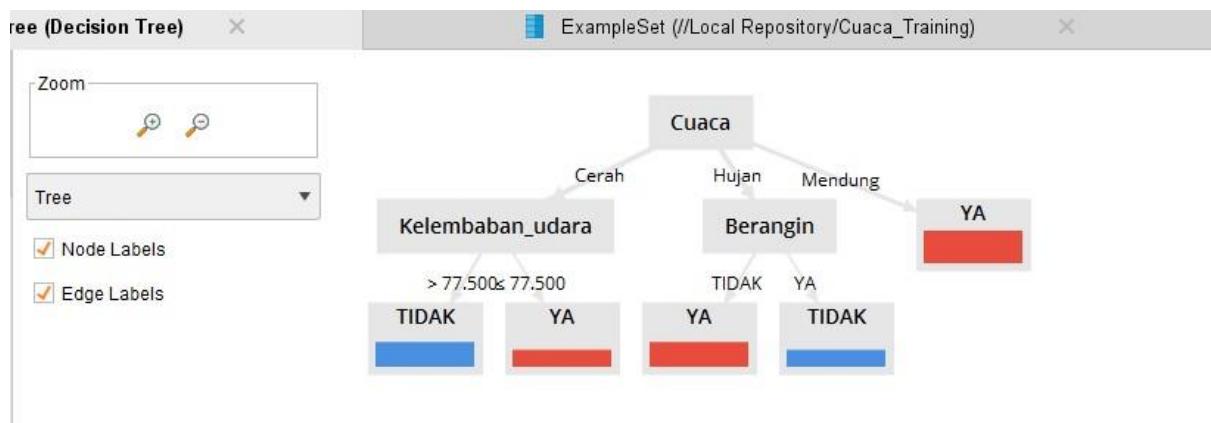


4. Lalu run hasilnya. Akan tampil seperti gambar dibawah ini

- Performance Vektor (Performance) menunjukkan tingkat akurasi, presisi, recall dan lain-lain dalam bentuk table atau pivot view.

accuracy: 60.00% +/- 45.95% (micro average: 64.29%)			
	true TIDAK	true YA	class precision
pred. TIDAK	2	2	50.00%
pred. YA	3	7	70.00%
class recall	40.00%	77.78%	

- Tree menunjukkan hasil pohon keputusan dari proses klasifikasi. Dapat dilihat bahwa pohon keputusan yang dihasilkan sama seperti saat menggunakan weka.



5. Berdasarkan pohon keputusan tersebut, dapat dilihat jenis – jenis simpul yang ada, sebagai berikut :

- Simpul akar = Cuaca
- Simpul Internal = Kelembaban\_Udara, dan Berangin
- Simpul daun = YA, TIDAK

6. Klasifikasi yang terbentuk yaitu :

- Seseorang akan bermain tenis (YA) jika kondisi sebagai kondisi :
  - i. Cuaca = Cerah, Kelembaban\_Udara  $\leq 77,5$  (nilai atribut lain diabaikan)
  - ii. Cuaca = Mendung, (nilai atribut lain diabaikan)
  - iii. Cuaca = Hujan, Brangin = TIDAK. (nilai atribut lain diabaikan)
- Seseorang tidak akan bermain tenis (TIDAK) jika kondisi sebagai berikut :
  - i. Cuaca = Cerah, Kelembabab\_udara  $> 77,5$  (nilai atribut lain diabaikan)

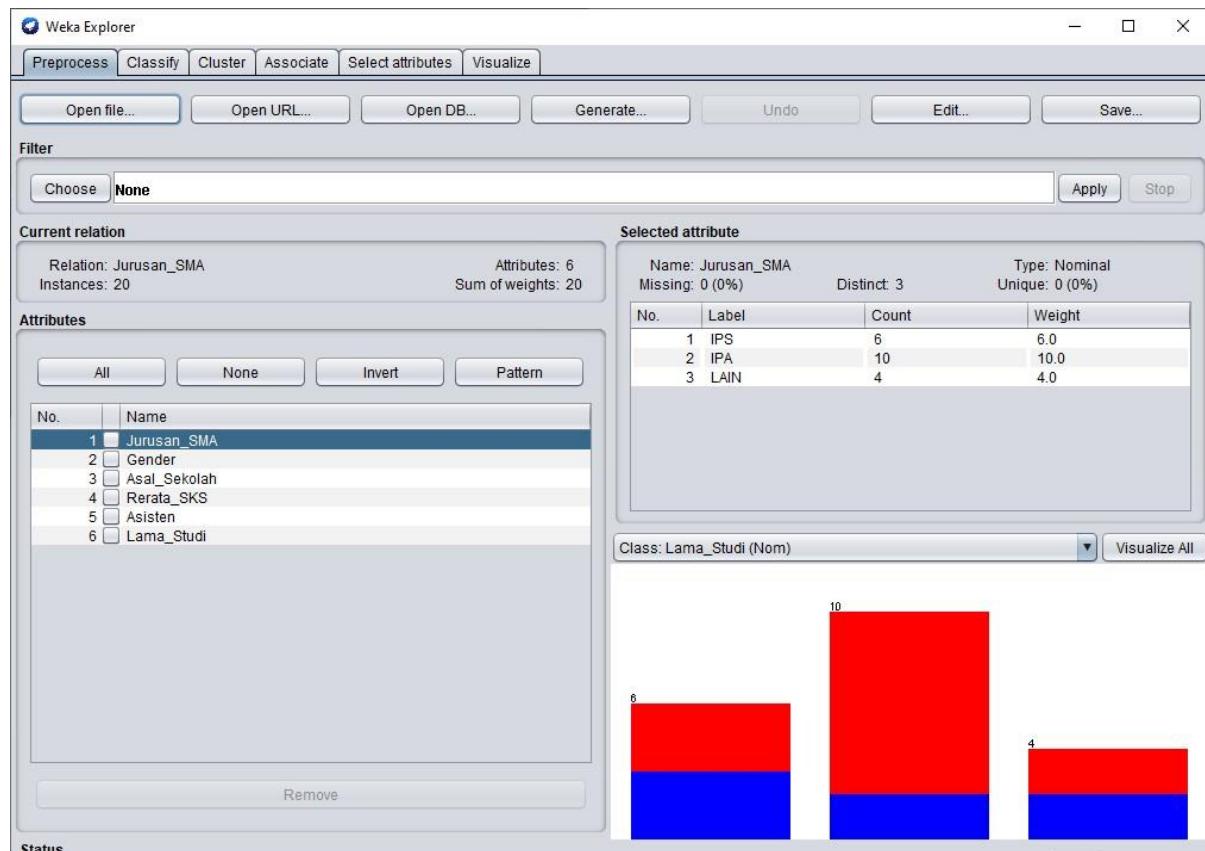
- ii. Cuaca = Hujan, Berangin = YA. (nilai atribut lain diabaikan)

### TUGAS :

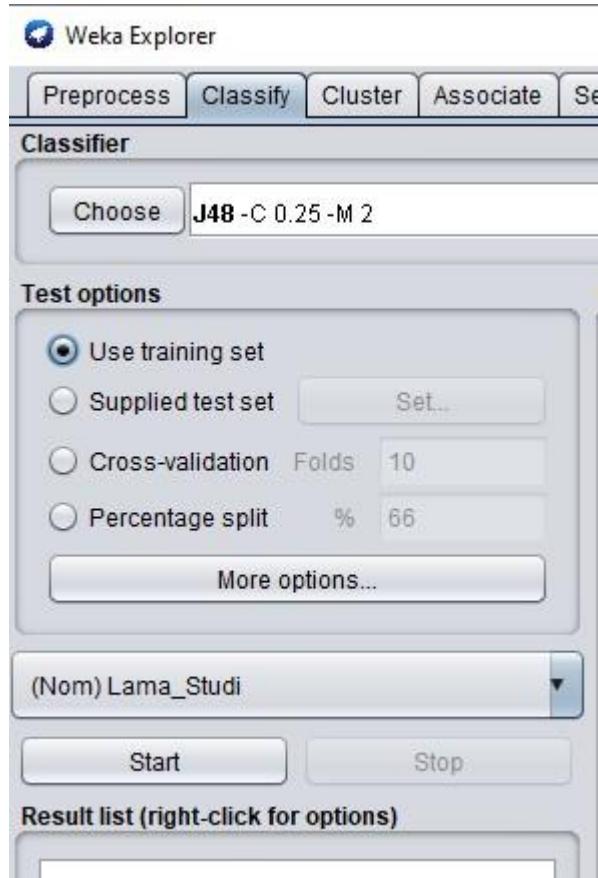
- Berdasarkan pohon keputusan pada percobaan yang kita buat tadi (menggunakan rapid miner). Isikan nilai kelas atribut Bermain\_Tenis pada table Testing berikut :

Cuaca	Suhu	Kelembaban_Udara	Berangin	Bermain_Tenis
Cerah	75	65	TIDAK	YA
Cerah	80	68	YA	YA
Cerah	83	87	YA	TIDAK
Mendung	70	96	TIDAK	YA
Mendung	68	81	TIDAK	YA
Hujan	65	75	TIDAK	YA
Hujan	64	85	YA	TIDAK

- buka file Jurusan\_SMA dengan weka Explorer



3. setelah itu buka tab Classify. Tekan tombol choose-Trees-J48. Lalu pilihlah Test Options, pilih Use Training set. Data Jurusan\_SMA.arff kita gunakan sebagai data training untuk menciptakan klasifikasi



4. Berdasarkan hasil tersebut dapat diketahui sebagai berikut :

- a) Jumlah simpul daun pada pohon keputusan = 3
- b) Jumlah simpul keseluruhan pada pohon keputusan = 5
- c) Waktu yang dibutuhkan untuk proses pelatihan = 0,02 detik
- d) Tingkat ketepatan klasifikasi = 85%
- e) Tingkat ketidaktepatan klasifikasi = 15%

```
Classifier output

J48 pruned tree
-----
Gender = PRIA
|   Rerata_SKS <= 18: TERLAMBAT (4.0)
|   Rerata_SKS > 18: TEPAT (11.0/2.0)
Gender = WANITA: TEPAT (5.0/1.0)

Number of Leaves :      3

Size of the tree :      5

Time taken to build model: 0.02 seconds

==== Evaluation on training set ====

Time taken to test model on training data: 0 seconds

==== Summary ===

Correctly Classified Instances          17           85    %
Incorrectly Classified Instances        3            15    %
Kappa statistic                         0.6341
Mean absolute error                     0.2436
Root mean squared error                 0.349
Relative absolute error                  53.0693 %
Root relative squared error             73.1456 %
Total Number of Instances                20


```

**Classifier output**

```

    === Evaluation on training set ===

    Time taken to test model on training data: 0 seconds

    === Summary ===

    Correctly Classified Instances      17          85      %
    Incorrectly Classified Instances   3           15      %
    Kappa statistic                   0.6341
    Mean absolute error              0.2436
    Root mean squared error          0.349
    Relative absolute error          53.0693 %
    Root relative squared error     73.1456 %
    Total Number of Instances        20

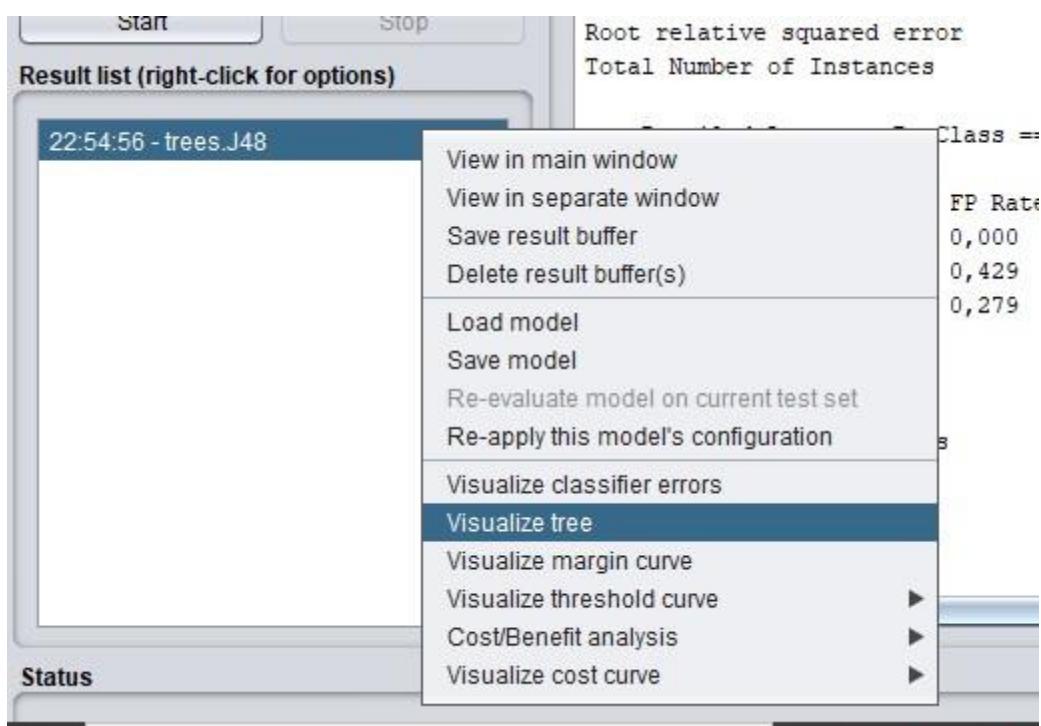
    === Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC   ROC Area   PRC Area   Class
      0,571     0,000     1,000      0,571     0,727      0,681   0,791     0,751     TER
      1,000     0,429     0,813      1,000     0,897      0,681   0,791     0,816     TEP
    Weighted Avg.   0,850     0,279     0,878      0,850     0,837      0,681   0,791     0,793

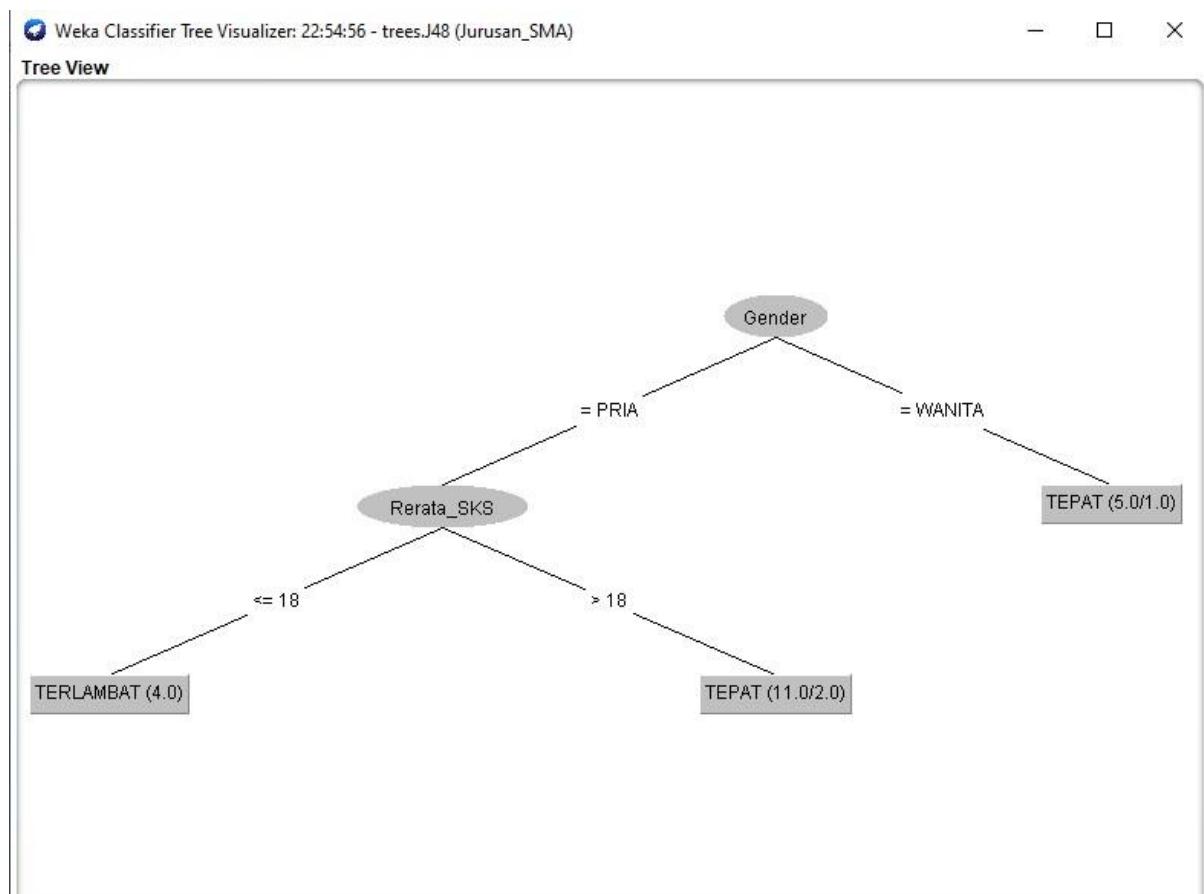
    === Confusion Matrix ===

      a   b   <-- classified as
      4   3   |   a = TERLAMBAT
      0   13  |   b = TEPAT
  
```

5. untuk melihat hasil skema pohon keputusan, kembali ke kolom Result List. Klik kanan pada hasil trees.j480-visualize tree



7. sebuah jendela weka Classifier Tree Visualizer akan ditampilkan. Pada jendela ini akan tampak Tree View, hasil klasifikasi pohon keputusan.



8. berdasarkan pohon keputusan tersebut, dapat dilihat jenis-jenis simpul yang ada, sebagai berikut:

- a) Simpul akar = Gender
- b) Simpul internal = Rerata\_SKS
- c) Simpul daun = TERLAMBAT, TEPAT

9. klasifikasi yang terbentuk yaitu :

- a) Seseorang akan Lama Studi (TERLAMBAT) jika kondisi sebagai berikut :
  - i. Gender = PRIA, Rerata\_SKS <= 18 (nilai atribut lain diabaikan)
- b) Seseorang akan Lama Studi (TEPAT) jika kondisi sebagai berikut :
  - i. Gender = PRIA, Rerata\_SKS > 18 (nilai atribut lain diabaikan)
  - ii. Gender = WANITA ( kondisi lain diabaikan)

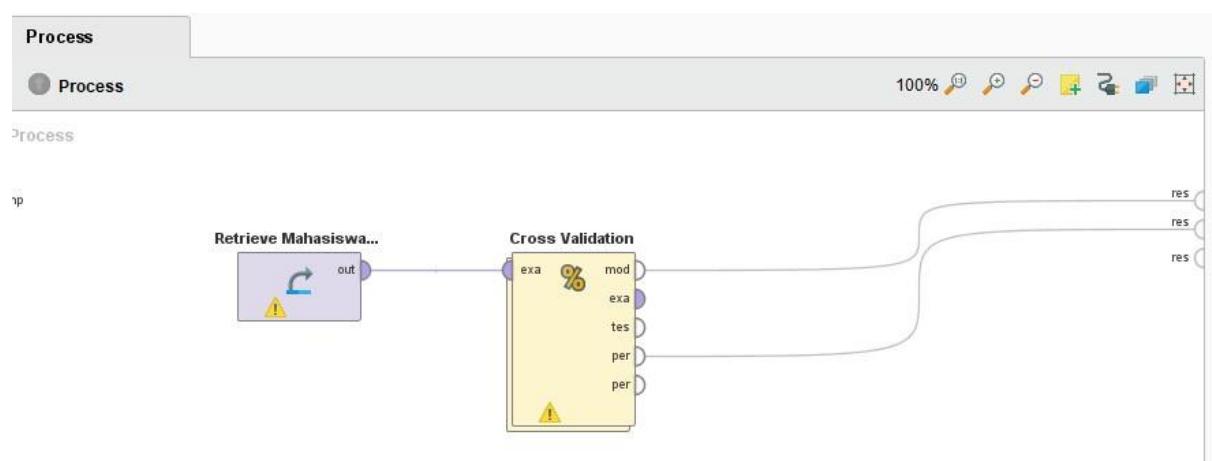
## Rapid Miner

1. Pada area process view, kita akan mendesain model proses pohon keputusan menggunakan data Mahasiswa\_Training. Yang telah kita buat di modul 7

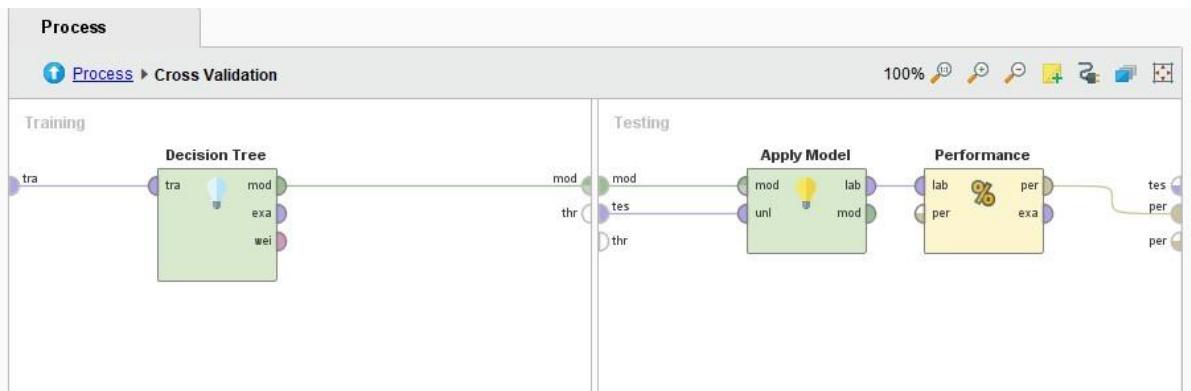
Row No.	Lama_Studi	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	TERLAMBAT	IPS	WANITA	SURAKARTA	18	TIDAK
2	TEPAT	IPA	PRIA	SURAKARTA	19	YA
3	TERLAMBAT	LAIN	PRIA	SURAKARTA	19	TIDAK
4	TERLAMBAT	IPA	PRIA	LUAR	17	TIDAK
5	TEPAT	IPA	WANITA	SURAKARTA	17	TIDAK
6	TEPAT	IPA	WANITA	LUAR	18	YA
7	TERLAMBAT	IPA	PRIA	SURAKARTA	18	TIDAK
8	TEPAT	IPA	PRIA	SURAKARTA	19	TIDAK
9	TERLAMBAT	IPS	PRIA	LUAR	18	TIDAK
10	TEPAT	LAIN	WANITA	SURAKARTA	18	TIDAK
11	TEPAT	IPA	WANITA	SURAKARTA	19	TIDAK
12	TEPAT	IPS	PRIA	SURAKARTA	20	TIDAK
13	TEPAT	IPS	PRIA	SURAKARTA	19	TIDAK
14	TEPAT	IPA	PRIA	SURAKARTA	19	TIDAK

ExampleSet (20 examples, 1 special attribute, 5 regular attributes)

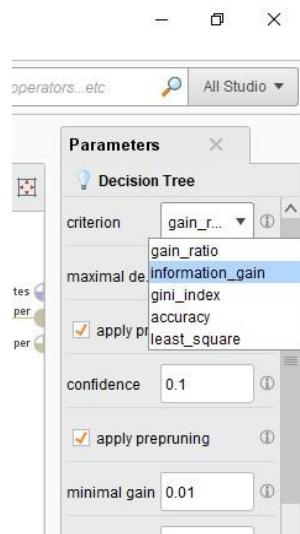
2. Drag Mahasiswa\_Training ke area process view, drag pula operator Cross Validation ke area process view. Lalu hubungan tiap portnya sesuai gambar dibawah ini :



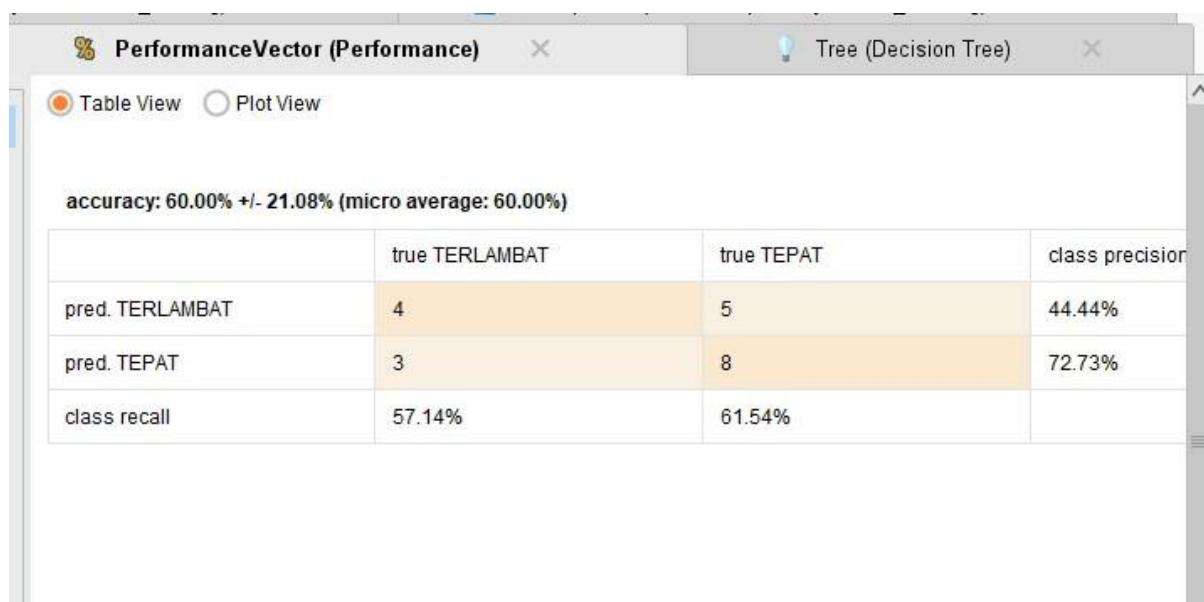
3. klik 2 kali operator Cross validation. Sehingga masuk ke jendela process-Validation. Masukkan operator Decision Tree dalam area Training, operator Apply Model dan performance dalam area Testing. Hubungkan port input dan output masing- masing operator seperti gambar dibawah ini :



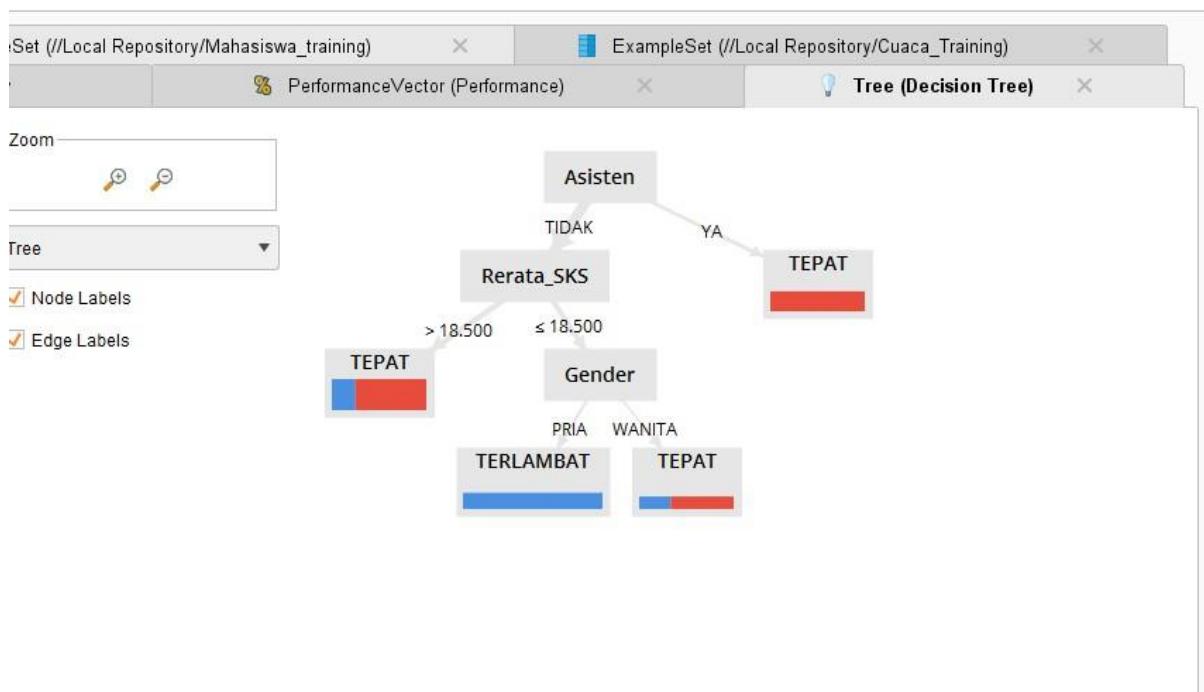
4. klik decision tree pada area training, dan pastikan kriteria (Criterion) yang dipakai adalah information Gain (J48) pada kolom parameter disebelah kanan process view.



## 5. hasilnya performanceVector



## 6. hasil Tree (Decision Tree) menunjukan hasil pohon keputusan dari proses klasifikasi.



7. berdasarkan pohon keputusan tersebut, dapat dilihat jenis-jenis simpul yang ada sebagai berikut :

- Simpul akar = Asisten
- Simpul internal = Rarata\_SKS, Gender
- Simpul daun = TEPAT, TERLAMBAT

8. klasifikasi yang terbentuk adalah :

- Seseorang akan Lama Studi (TEPAT) jika kondisi sebagai berikut :
  - i. Asisten = TIDAK, Rerata\_SKS > 18,500.( nilai atribut lain diabaikan)
  - ii. Asisten = TIDAK, Rerata\_SKS <= 18.500 Gender(WANITA) = TEPAT  
(nilai atribut lain diabaikan)
  - iii. Asisten = YA.(nilai atribut lain diabaikan)
- Seseorang akan Lama Studi (TERLAMBAT) jika kondisi sebagai berikut :
  - i. Asisten = TIDAK, Rerata\_SKS <= 18.500 Gender (PRIA) = TERLAMBAT (nilai atribut lain diabaikan)

## MODUL 10

### Percobaan

1. membuat table data nilai ujian siswa dan disimpan dengan nama file Tabel\_NilaiUjian.xls

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

2. masukan data ke dalam rapid miner

Import Data - Format your columns.

**Format your columns.**

Replace errors with missing values (i)

	NAMA polynominal id	B.IND real	B.ING real
1	JOKO	8.540	8.400
2	AGUS	9.980	6.810
3	SUSI	6.200	9.150
4	DYAH	5.240	7.260
5	WATI	5.700	5.710
6	IKA	8.570	5.870
7	EKO	7.700	7.710
8	YANTO	6.600	5.700
9	WAWAN	9.000	8.120
10	MAHMUD	9.810	9.580

 no problems.

 Previous  Next  Cancel

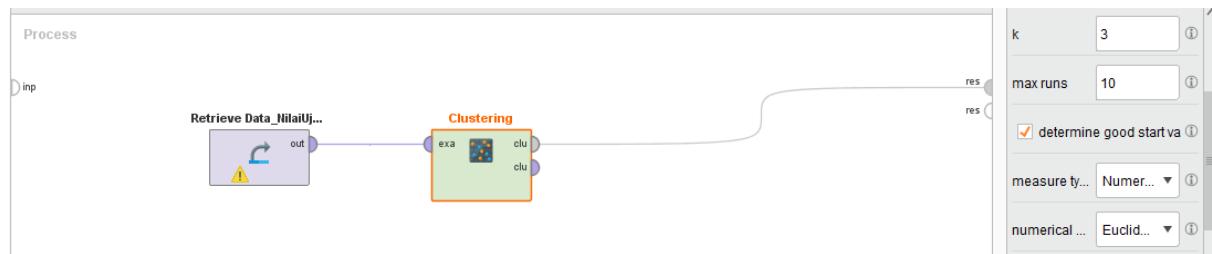
3. beri nama Data\_NilaiUjian dan masukan data pada repositories. Kemudian klik finish

Dan hasilnya :

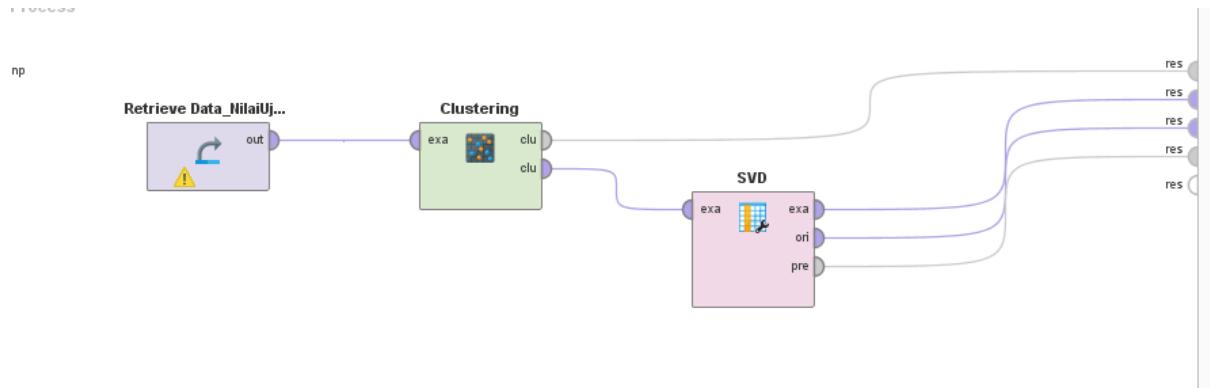
Row No.	NAMA	B.IND	B.ING
1	JOKO	8.540	8.400
2	AGUS	9.980	6.810
3	SUSI	6.200	9.150
4	DYAH	5.240	7.260
5	WATI	5.700	5.710
6	IKA	8.570	5.870
7	EKO	7.700	7.710
8	YANTO	6.600	5.700
9	WAWAN	9	8.120
10	MAHMUD	9.810	9.580

4. Gunakan Data\_NilaiUjian ini dan masukkan ke dalam area proses

5. tambahkan operator –means. Hubungkan output operator retrieve ke entry exa operator ini dan output clu(cluster model) dihubungkan ke connector res panel. Ubah nilai parameter k =3 pada operator ini .



6. tambahkan operator SVD. Lalu hubungkan output clu ke-2 operator clustering (k-means) kedalam entry exa operator SVD dan 3 port output exa,ori, dan pre terhadap konektor



7. hasil proses clustering dengan algoritma K-means

a) SVD

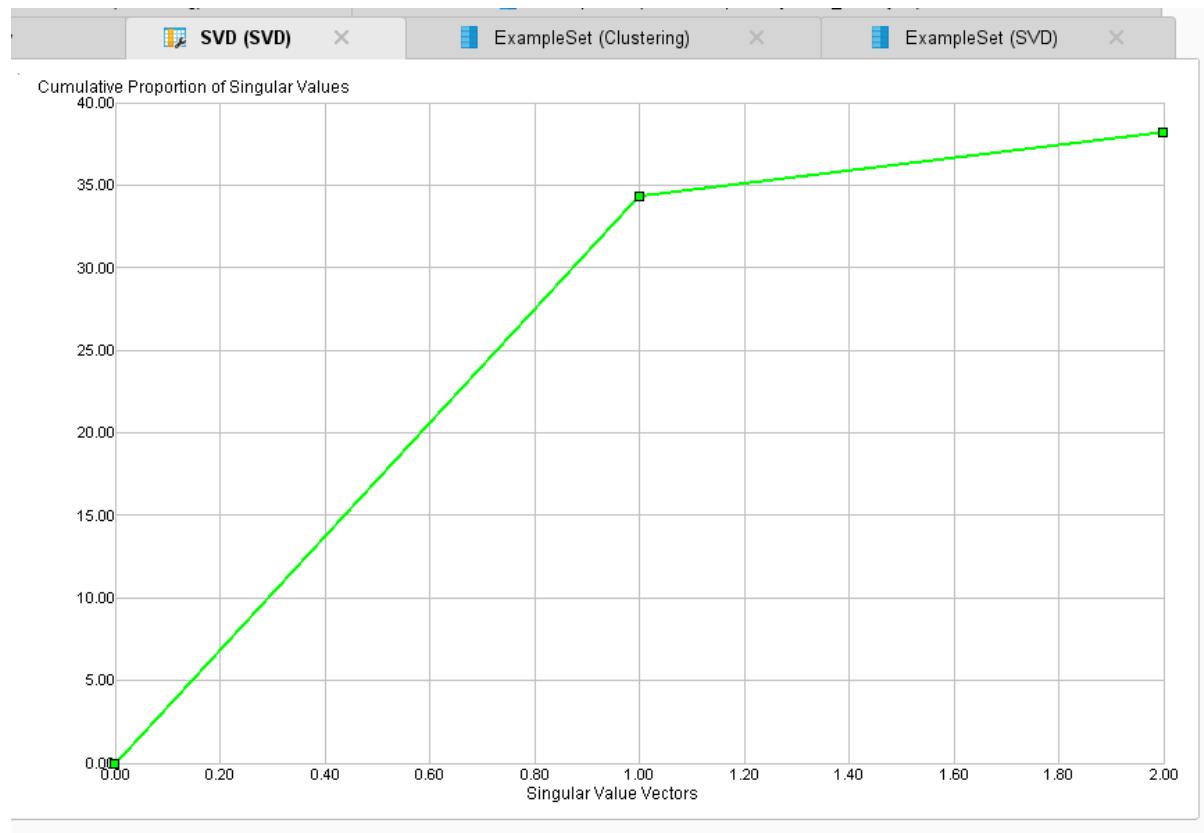
i. nilai Eigenvalue

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

8. Nilai Svd vector

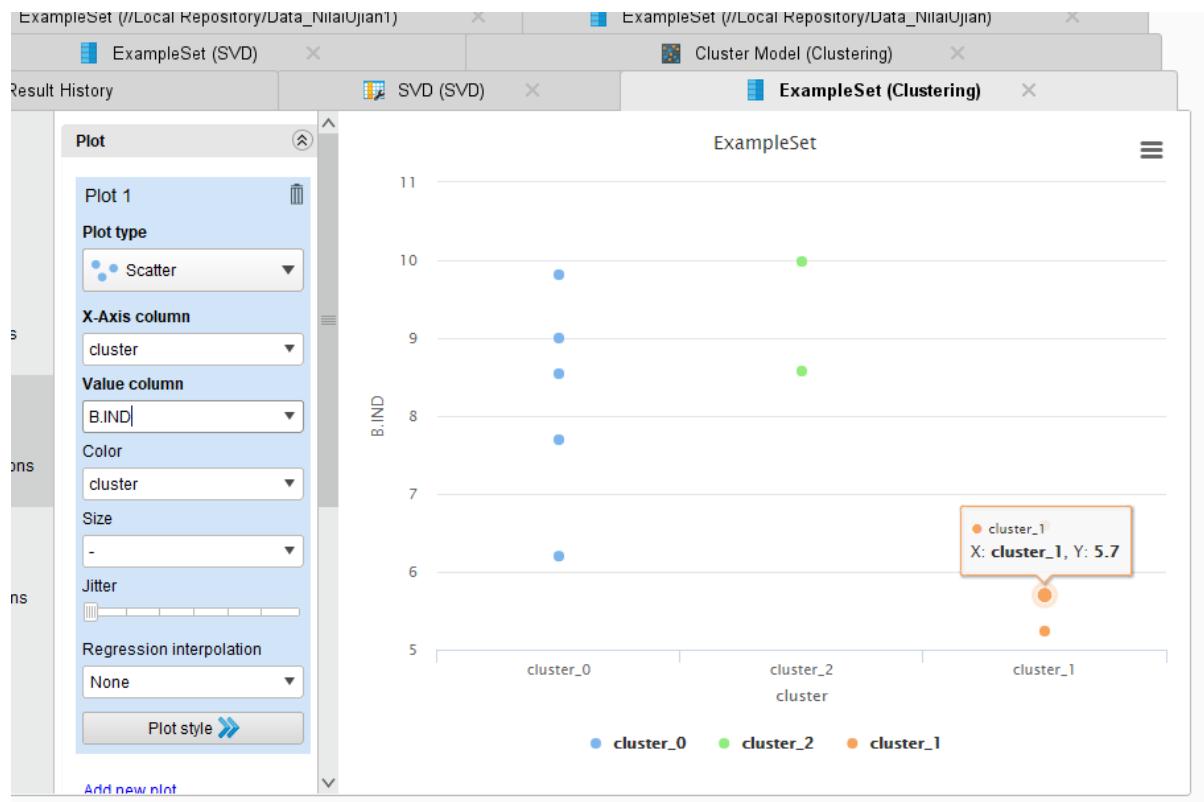
SVD (SVD)		ExampleSet (Clustering)	ExampleSet (SVD)
<b>SVD Vector 1</b>			
B.IND			0.723
B.ING			0.690

9. nilai cumulative variance

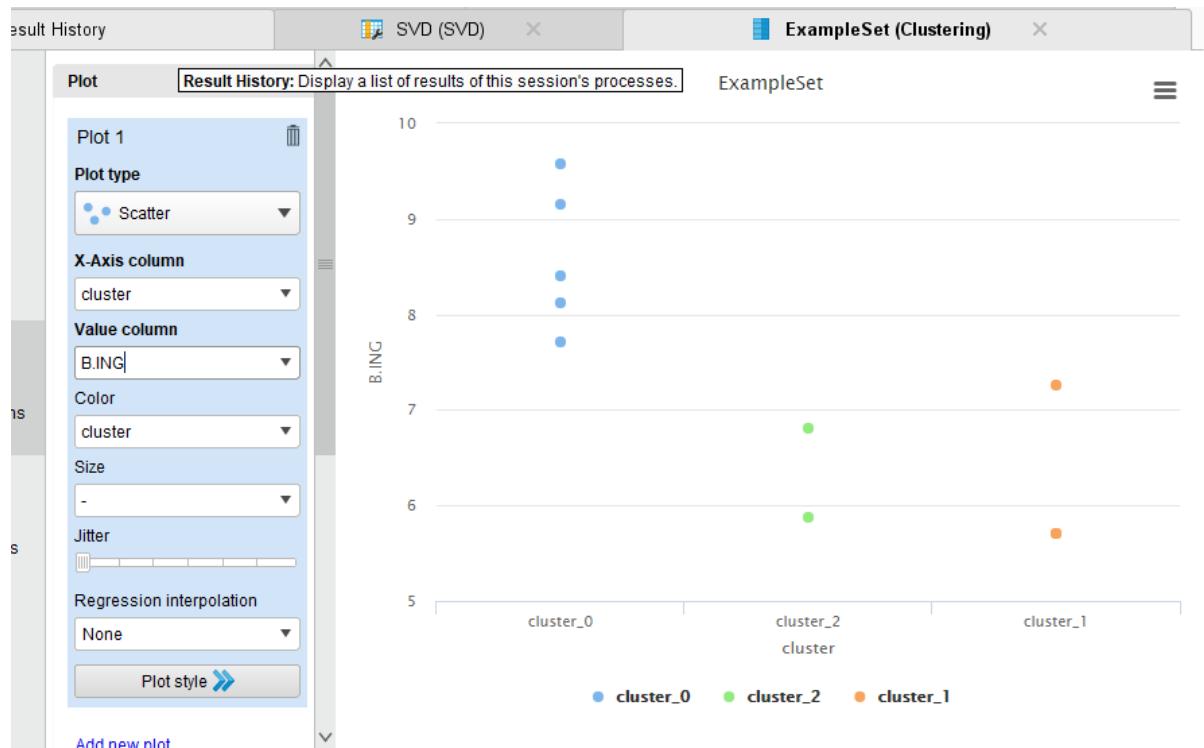


## 10. exampleSet K-means

Kelompok siswa B.Indonesia



## 11. Kelompok siswa B. Inggris



## 12. hasil ExampleSet (SVD)

ExampleSet (SVD) X

Open in Turbo Prep Auto Model

Row No.	NAMA	cluster	svd_1
1	JOKO	cluster_0	0.349
2	AGUS	cluster_2	0.347
3	SUSI	cluster_0	0.315
4	DYAH	cluster_1	0.256
5	WATI	cluster_1	0.235
6	IKA	cluster_2	0.299
7	EKO	cluster_0	0.317
8	YANTO	cluster_1	0.254
9	WAWAN	cluster_0	0.353
10	MAHMUD	cluster_0	0.399

### 13. Cluster Model(Clustering)

ExampleSet (SVD) X Clu

## Cluster Model

on

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

### TUGAS

1. membuat table Data Nilai ujian 30 siswa

NO_SISWA	NAMA	B.IND	B.ING	MTK	IPA
S_101	JOKO	9,83	5,16	6,75	7,24
S_102	AGUS	7,15	8,54	6,37	8,62
S_103	SUSI	8,33	9,83	9,01	8,21
S_104	DYAH	8,88	7,72	7,48	5,25
S_105	WATI	6,34	8,90	9,34	6,43
S_106	IKA	7,92	8,22	7,77	9,19
S_107	EKO	9,80	7,47	9,08	5,48
S_108	YANTO	7,33	5,71	7,65	7,29
S_109	WAWAN	5,25	8,68	8,76	6,74
S_110	MAHMUD	6,89	5,89	9,96	8,50
S_111	BUDI	8,87	8,26	9,76	8,99
S_112	SANTI	8,83	5,42	9,35	8,33
S_113	DIAN	6,49	5,45	7,36	9,85
S_114	DANI	7,80	9,24	7,64	6,11
S_115	AHMAD	7,53	5,52	7,48	8,54
S_116	BAYU	9,41	9,99	7,91	6,36
S_117	RISA	8,98	8,63	9,51	6,07
S_118	RANI	6,00	8,82	9,07	7,91
S_119	YANI	7,62	8,70	5,68	6,37
S_120	RATIH	6,19	7,49	7,13	5,24
S_121	INDAH	8,04	6,86	6,74	5,74
S_122	JONO	7,23	9,69	7,40	9,15
S_123	SARAH	6,99	5,92	8,32	6,02
S_124	RAMA	5,36	6,52	7,73	9,92

8	S_107	EKO	9,80	7,47	9,08	5,48
9	S_108	YANTO	7,33	5,71	7,65	7,29
10	S_109	WAWAN	5,25	8,68	8,76	6,74
11	S_110	MAHMUD	6,89	5,89	9,96	8,50
12	S_111	BUDI	8,87	8,26	9,76	8,99
13	S_112	SANTI	8,83	5,42	9,35	8,33
14	S_113	DIAN	6,49	5,45	7,36	9,85
15	S_114	DANI	7,80	9,24	7,64	6,11
16	S_115	AHMAD	7,53	5,52	7,48	8,54
17	S_116	BAYU	9,41	9,99	7,91	6,36
18	S_117	RISA	8,98	8,63	9,51	6,07
19	S_118	RANI	6,00	8,82	9,07	7,91
20	S_119	YANI	7,62	8,70	5,68	6,37
21	S_120	RATIH	6,19	7,49	7,13	5,24
22	S_121	INDAH	8,04	6,86	6,74	5,74
23	S_122	JONO	7,23	9,69	7,40	9,15
24	S_123	SARAH	6,99	5,92	8,32	6,02
25	S_124	RAMA	5,36	6,52	7,73	9,92
26	S_125	BAMBANG	9,04	7,85	9,40	7,52
27	S_126	HADI	6,18	8,03	9,85	9,29
28	S_127	NANA	7,37	6,17	7,97	9,75
29	S_128	FEBRI	8,43	8,73	9,27	5,94
30	S_129	DENI	6,64	7,04	8,60	8,77
31	S_130	TONI	6,76	7,41	5,76	9,85
32						

2. import data ke rapid miner.

Import Data - Format your columns.

X

## Format your columns.

Replace errors with missing values ⓘ

	NAMA polynomial id	B.IND real	B.ING real	MTK real	IPA real
1	JOKO	8.069	9.172	9.326	6.006
2	AGUS	9.810	8.509	5.971	6.008
3	SUSI	8.602	9.403	6.409	9.258
4	DYAH	5.619	6.698	9.828	7.557
5	WATI	8.196	5.608	6.702	9.261
6	IKA	5.448	5.949	7.499	6.025
7	EKO	6.745	7.907	8.642	7.304
8	YANTO	6.912	7.544	6.446	7.986
9	WAWAN	8.444	9.490	7.766	7.840
10	MAHMUD	8.038	6.142	8.208	7.577
11	BUDI	7.736	8.671	9.367	7.444
12	SANTI	5.144	6.170	5.125	6.519
13	DIAN	7.268	7.051	9.014	7.205

 no problems.

 Previous

Next 

 Cancel

	NAMA polynomial id	B.IND real	B.ING real	MTK real	IPA real
11	BUDI	7.736	8.671	9.367	7.444
12	SANTI	5.144	6.170	5.125	6.519
13	DIAN	7.268	7.051	9.014	7.205
14	DANI	9.713	8.568	9.556	7.519
15	AHMAD	7.272	7.127	9.235	5.901
16	BAYU	9.329	7.103	6.515	9.815
17	RISA	8.153	7.100	6.944	9.377
18	RANI	8.738	7.510	9.480	5.283
19	YANI	7.627	9.513	9.009	9.249
20	RATIH	9.298	7.837	7.032	7.433
21	INDAH	7.441	7.737	7.633	6.473
22	JONO	5.378	9.710	9.955	6.616

 no problems.

Replace errors with missing values ⓘ

	NAMA polynomial id	B.IND real	B.ING real	MTK real	IPA real
18	RANI	8.738	7.510	9.480	5.283
19	YANI	7.627	9.513	9.009	9.249
20	RATIH	9.298	7.837	7.032	7.433
21	INDAH	7.441	7.737	7.633	6.473
22	JONO	5.378	9.710	9.955	6.616
23	SARAH	9.351	7.632	9.889	8.594
24	RAMA	6.287	5.718	8.925	5.443
25	BAMBANG	8.692	6.366	6.494	5.608
26	HADI	8.697	9.233	8.457	9.572
27	NANA	5.716	7.498	8.760	9.800
28	FEBRI	5.491	8.904	7.519	8.144
29	DENI	7.990	8.652	5.651	5.987
30	TONI	8.194	6.335	5.298	7.201

 no problems.

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

3.

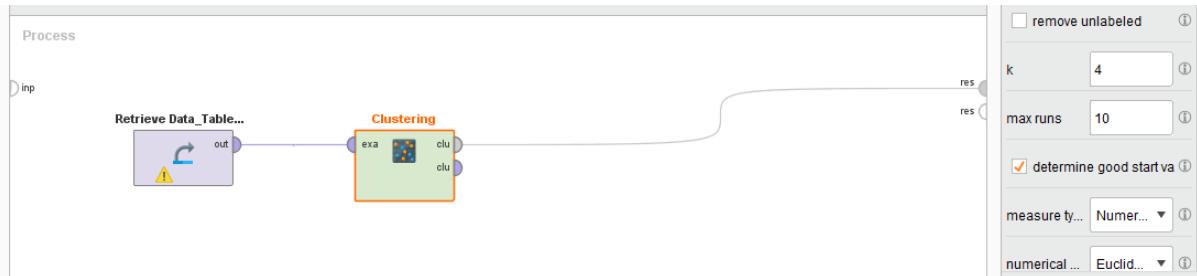
Row No.	NAMA	B.IND	B.ING	MTK	IPA
1	JOKO	8.069	9.172	9.326	6.006
2	AGUS	9.810	8.509	5.971	6.008
3	SUSI	8.602	9.403	6.409	9.258
4	DYAH	5.619	6.698	9.828	7.557
5	WATI	8.196	5.608	6.702	9.261
6	IKA	5.448	5.949	7.499	6.025
7	EKO	6.745	7.907	8.642	7.304
8	YANTO	6.912	7.544	6.446	7.986
9	WAWAN	8.444	9.490	7.766	7.840
10	MAHMUD	8.038	6.142	8.208	7.577
11	BUDI	7.736	8.671	9.367	7.444
12	SANTI	5.144	6.170	5.125	6.519
13	DIAN	7.268	7.051	9.014	7.205

Row No.	NAMA	B.IND	B.ING	MTK	IPA
13	DIAN	7.200	7.031	9.014	7.200
14	DANI	9.713	8.568	9.556	7.519
15	AHMAD	7.272	7.127	9.235	5.901
16	BAYU	9.329	7.103	6.515	9.815
17	RISA	8.153	7.100	6.944	9.377
18	RANI	8.738	7.510	9.480	5.283
19	YANI	7.627	9.513	9.009	9.249
20	RATIH	9.298	7.837	7.032	7.433
21	INDAH	7.441	7.737	7.633	6.473
22	JONO	5.378	9.710	9.955	6.616
23	SARAH	9.351	7.632	9.889	8.594
24	RAMA	6.287	5.718	8.925	5.443
25	BAMBANG	8.692	6.366	6.494	5.608
26	HADI	8.607	9.233	8.457	9.572

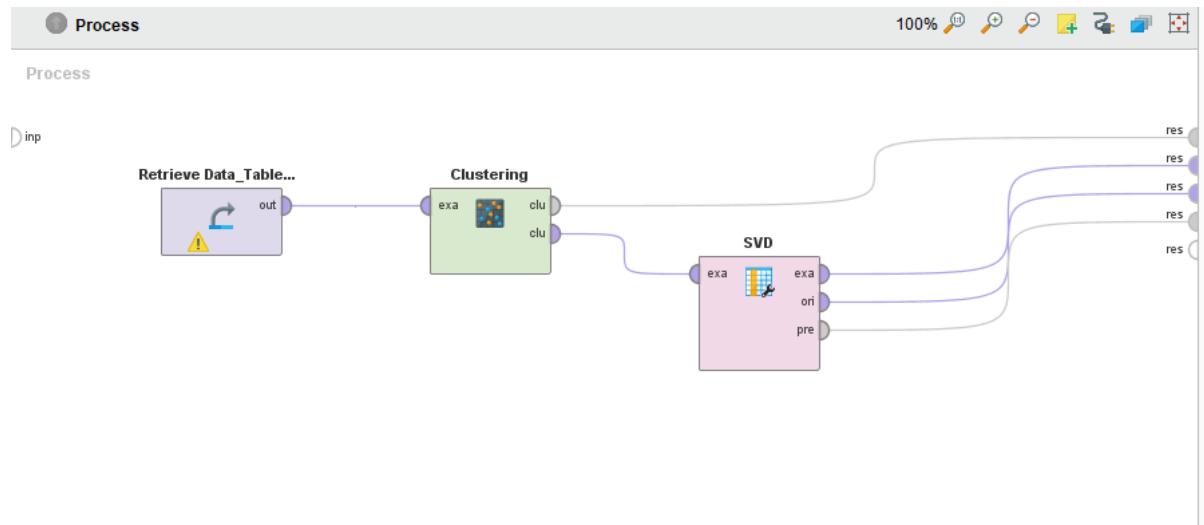
ExampleSet (30 examples, 1 special attribute, 4 regular attributes)

Row No.	NAMA	B.IND	B.ING	MTK	IPA
18	RANI	8.738	7.510	9.480	5.283
19	YANI	7.627	9.513	9.009	9.249
20	RATIH	9.298	7.837	7.032	7.433
21	INDAH	7.441	7.737	7.633	6.473
22	JONO	5.378	9.710	9.955	6.616
23	SARAH	9.351	7.632	9.889	8.594
24	RAMA	6.287	5.718	8.925	5.443
25	BAMBANG	8.692	6.366	6.494	5.608
26	HADI	8.697	9.233	8.457	9.572
27	NANA	5.716	7.498	8.760	9.800
28	FEBRI	5.491	8.904	7.519	8.144
29	DENI	7.990	8.652	5.651	5.987
30	TONI	8.194	6.335	5.298	7.201

4. tambahkan operator –means. Hubungkan output operator retrieve ke entry exa operator ini dan output clu(cluster model) dihubungkan ke connector res panel. Ubah nilai parameter k =3 pada operator ini



5. tambahkan operator SVD. Lalu hubungkan output clu ke-2 operator clustering (k-means) kedalam entry exa operator SVD dan 3 port output exa,ori, dan pre terhadap konektor



6. hasil proses clustering dengan algoritma K-means

a) SVD

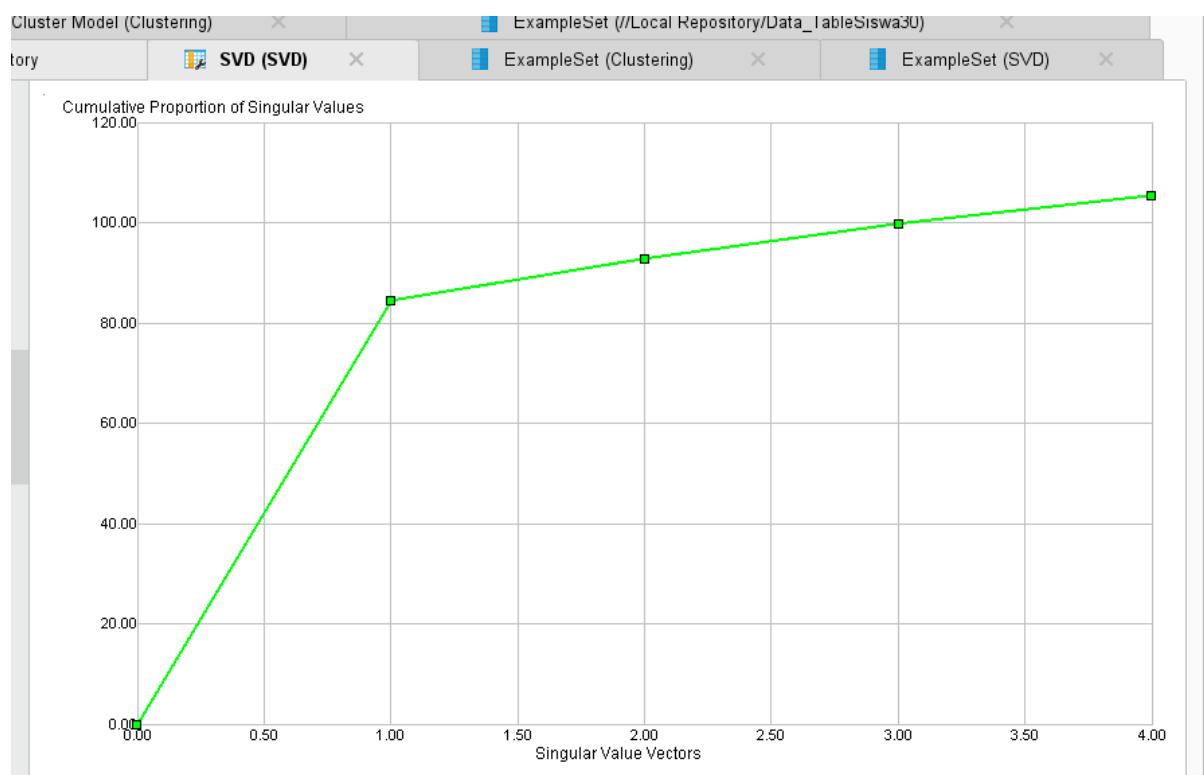
i. nilai Eigenvalue

Component	Singular Value	Proportion of Singular V...	Cumulative Singular Val...	Cumulative Proportion o...
SVD 1	84.502	0.801	84.502	0.801
SVD 2	8.430	0.080	92.933	0.881
SVD 3	6.944	0.066	99.876	0.947
SVD 4	5.599	0.053	105.475	1.000

## 7. Nilai Svd vector

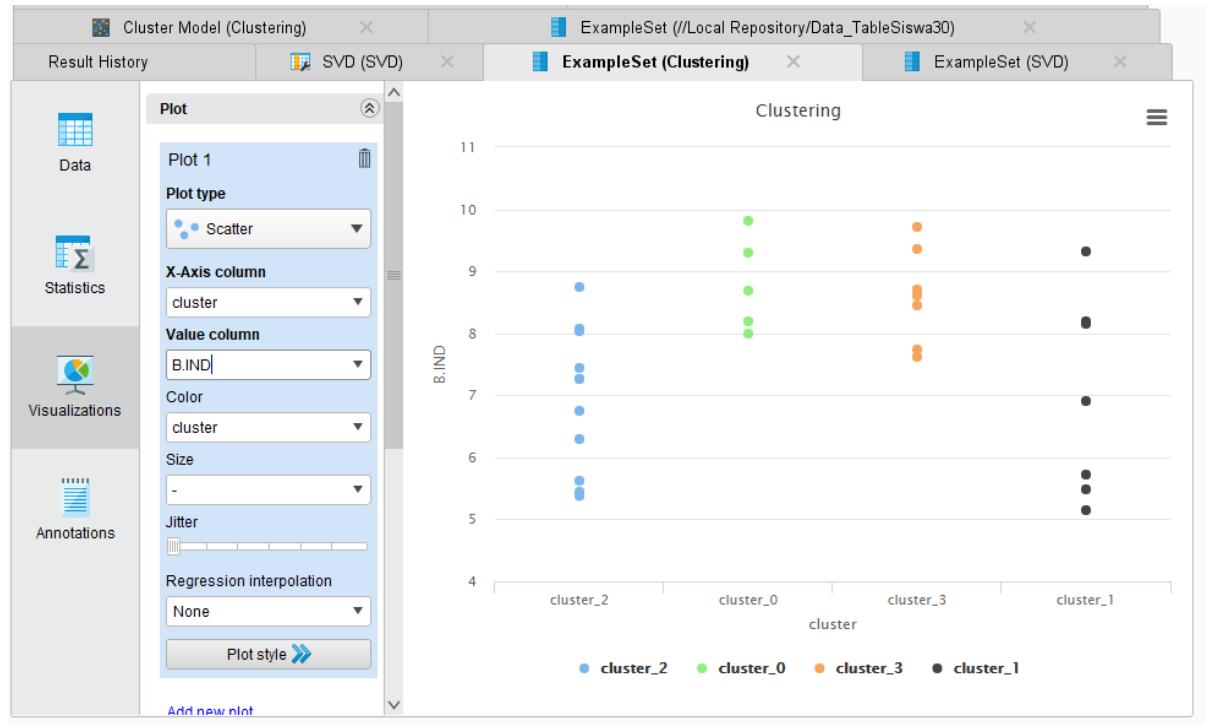
Attribute	SVD Vector 1	SVD Vector 2	SVD Vector 3
B.IND	0.498	-0.522	0.615
B.ING	0.502	0.068	0.095
MTK	0.514	0.774	0.051
IPA	0.487	-0.353	-0.781

## 8. nilai cumulative variance

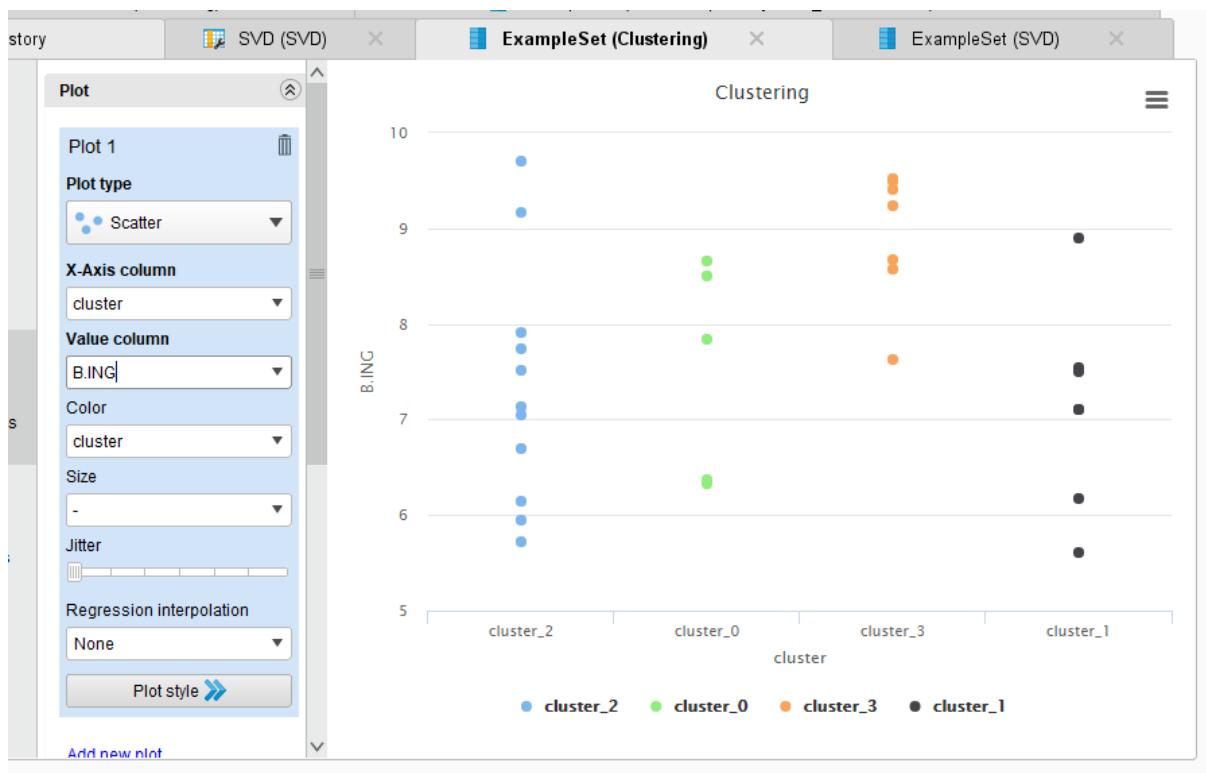


## 9. exampleSet K-means

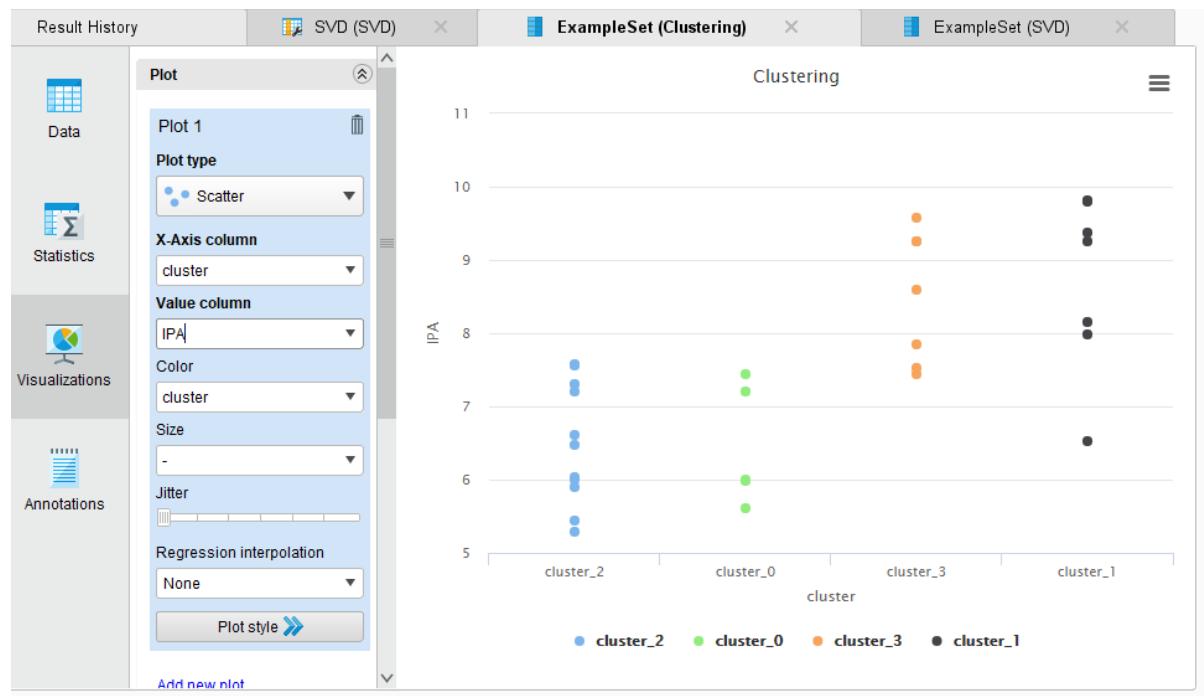
Kelompok siswa B.Indonesia



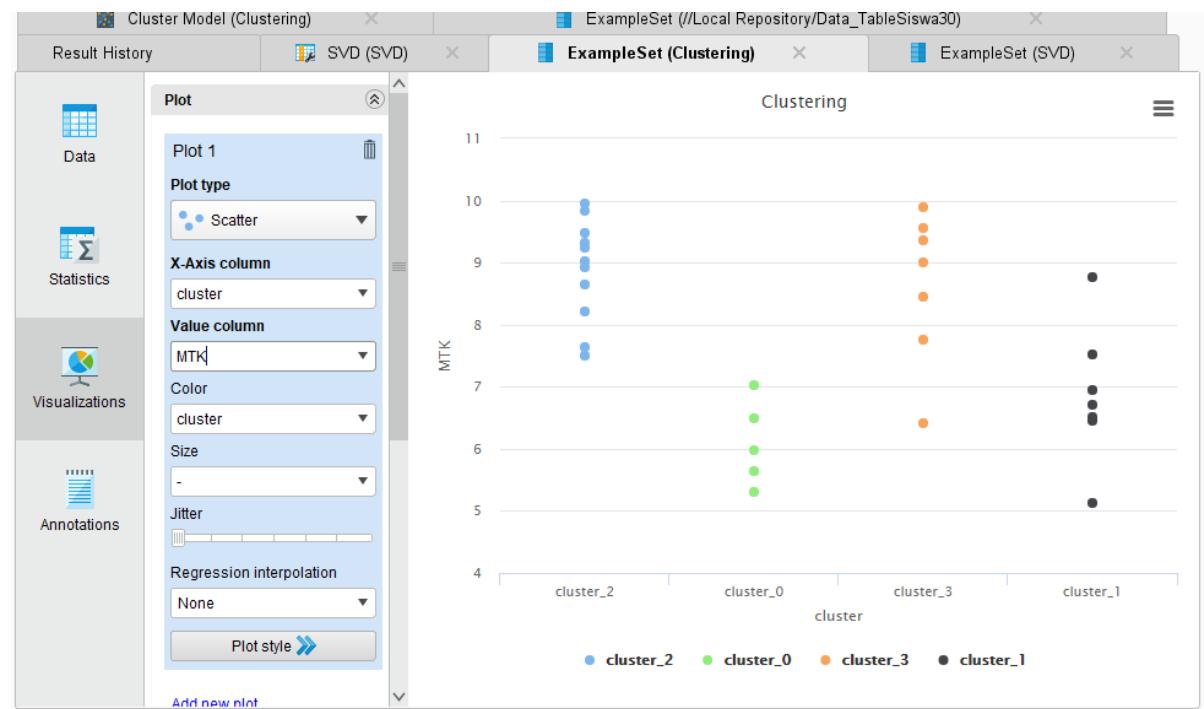
## 10. Kelompok siswa B. Inggris



## 9. Kelompok siswa IPA



## 10. Kelompok siswa MTK



## 11. hasil ExampleSet (SVD)

**Data**

**Statistics**

**Visualizations**

**Annotations**

Open in [Turbo Prep](#) [Auto Model](#)

Row No.	NAMA	cluster	svd_1
1	JOKO	cluster_2	0.193
2	AGUS	cluster_0	0.179
3	SUSI	cluster_3	0.199
4	DYAH	cluster_2	0.176
5	WATI	cluster_1	0.176
6	IKA	cluster_2	0.148
7	EKO	cluster_2	0.181
8	YANTO	cluster_1	0.171
9	WAWAN	cluster_3	0.198
10	MAHMUD	cluster_2	0.177
11	BUDI	cluster_3	0.197
12	SANTI	cluster_1	0.136
13	DIAN	cluster_2	0.181

Result History

SVD (SVD)

ExampleSet (Clustering)



Data

Open in [Turbo Prep](#)[Auto Model](#)

Statistics



Visualizations



Annotations

Row No.	NAMA	cluster	svd_1
12	DANI	cluster_1	0.150
13	DIAN	cluster_2	0.181
14	DANI	cluster_3	0.209
15	AHMAD	cluster_2	0.175
16	BAYU	cluster_1	0.193
17	RISA	cluster_1	0.186
18	RANI	cluster_2	0.184
19	YANI	cluster_3	0.209
20	RATIH	cluster_0	0.187
21	INDAH	cluster_2	0.173
22	JONO	cluster_2	0.188
23	SARAH	cluster_3	0.210
24	RAMA	cluster_2	0.157
25	PAMPRANG	cluster_0	0.161

ExampleSet (30 examples, 2 special attributes, 1 regular attribute)

Screenshot of the KNIME interface showing the SVD (SVD) node results.

The title bar shows "Result History" and "SVD (SVD)".

The left sidebar has tabs: Data, Statistics, Visualizations, and Annotations. The "Data" tab is selected.

Buttons above the table: "Open in Turbo Prep" and "Auto Model".

The table displays the following data:

Row No.	NAMA	cluster	svd_1
18	RANI	cluster_2	0.184
19	YANI	cluster_3	0.209
20	RATIH	cluster_0	0.187
21	INDAH	cluster_2	0.173
22	JONO	cluster_2	0.188
23	SARAH	cluster_3	0.210
24	RAMA	cluster_2	0.157
25	BAMBANG	cluster_0	0.161
26	HADI	cluster_3	0.213
27	NANA	cluster_1	0.188
28	FEBRI	cluster_1	0.178
29	DENI	cluster_0	0.167
30	TONI	cluster_0	0.160

Text below the table: "ExampleSet (30 examples, 2 special attributes, 1 regular attribute)"

## 12. Cluster Model(Clustering)

Screenshot of the KNIME interface showing the Cluster Model (Clustering) node results.

The title bar shows "Cluster Model (Clustering)".

The left sidebar has tabs: Description and Folder View. The "Description" tab is selected.

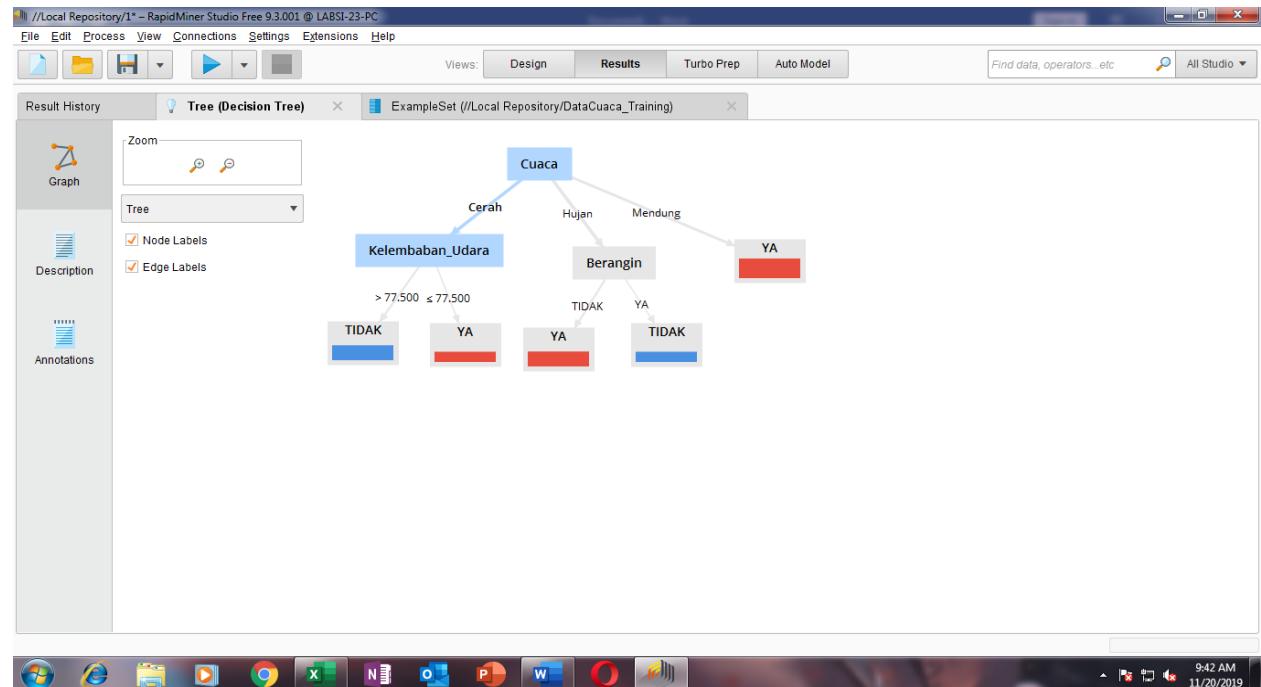
The main area displays the following text:

**Cluster Model**

Cluster 0: 5 items  
 Cluster 1: 7 items  
 Cluster 2: 11 items  
 Cluster 3: 7 items  
 Total number of items: 30

# MODUL 11

## Percobaan dan tugas



The screenshot shows the RapidMiner Studio interface with the 'RuleModel (Rule Induction)' results. It displays a table of 14 examples from the 'DataCuaca\_Training' dataset. The columns are: Row No., Bermain\_Te... (Bernain\_Te...), Cuaca (Weather), Suhu (Temperature), Kelembaban... (Humidity), and Berangin (Wind). The data is as follows:

Row No.	Bermain_Te...	Cuaca	Suhu	Kelembaban...	Berangin
1	TIDAK	Cerah	85	85	TIDAK
2	TIDAK	Cerah	80	90	YA
3	YA	Mendung	83	86	TIDAK
4	YA	Hujan	70	96	TIDAK
5	YA	Hujan	68	80	TIDAK
6	TIDAK	Hujan	65	70	YA
7	YA	Mendung	64	65	YA
8	TIDAK	Cerah	72	95	TIDAK
9	YA	Cerah	69	70	TIDAK
10	YA	Hujan	75	80	TIDAK
11	YA	Cerah	75	70	YA
12	YA	Mendung	72	90	YA
13	YA	Mendung	81	75	TIDAK
14	TIDAK	Hujan	71	91	YA

//Local Repository/1 - RapidMiner Studio Free 9.3.001 @ LABSI-23-PC

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model

Find data, operators...etc All Studio

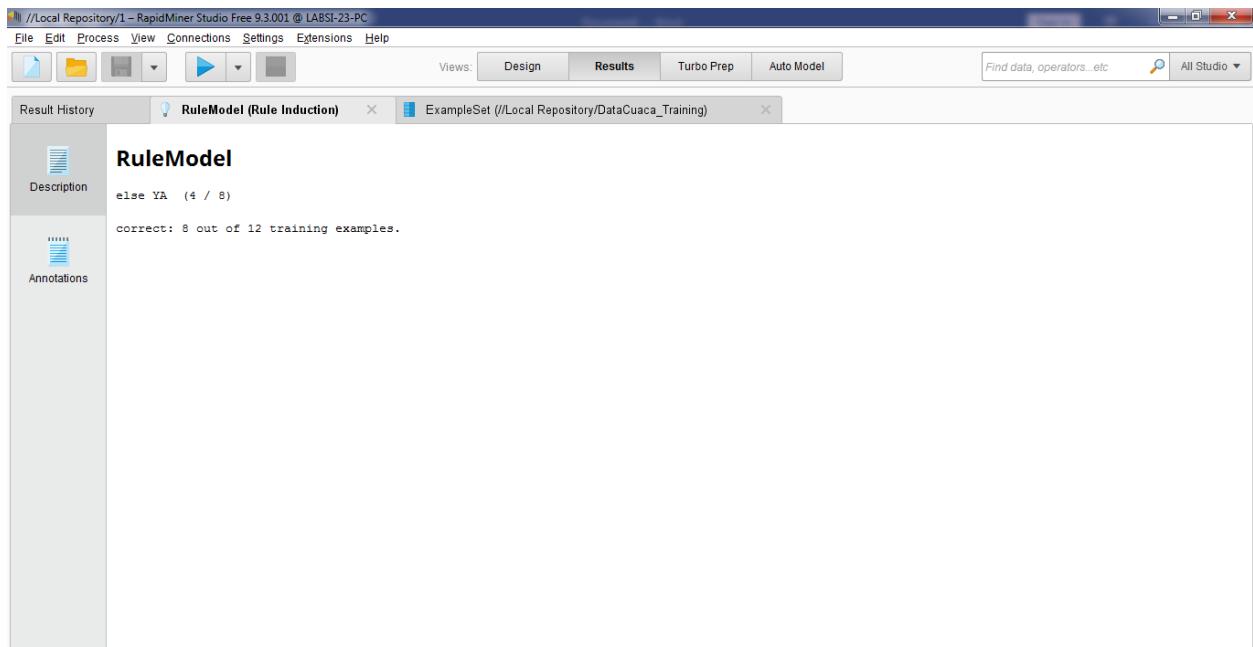
Result History RuleModel (Rule Induction) ExampleSet (/Local Repository/DataCuaca\_Training)

**RuleModel**

Description

Annotations

else YA (4 / 8)  
correct: 8 out of 12 training examples.



//Local Repository/2 – RapidMiner Studio Free 9.3.001 @ LABSI-23-PC

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model

Find data, operators...etc All Studio

Result History ExampleSet (Nominal to Binomial) FrequentItemSets (FP-Growth) ExampleSet (/Local Repository/DataCuaca\_Training) AssociationRules (Create Association Rules)

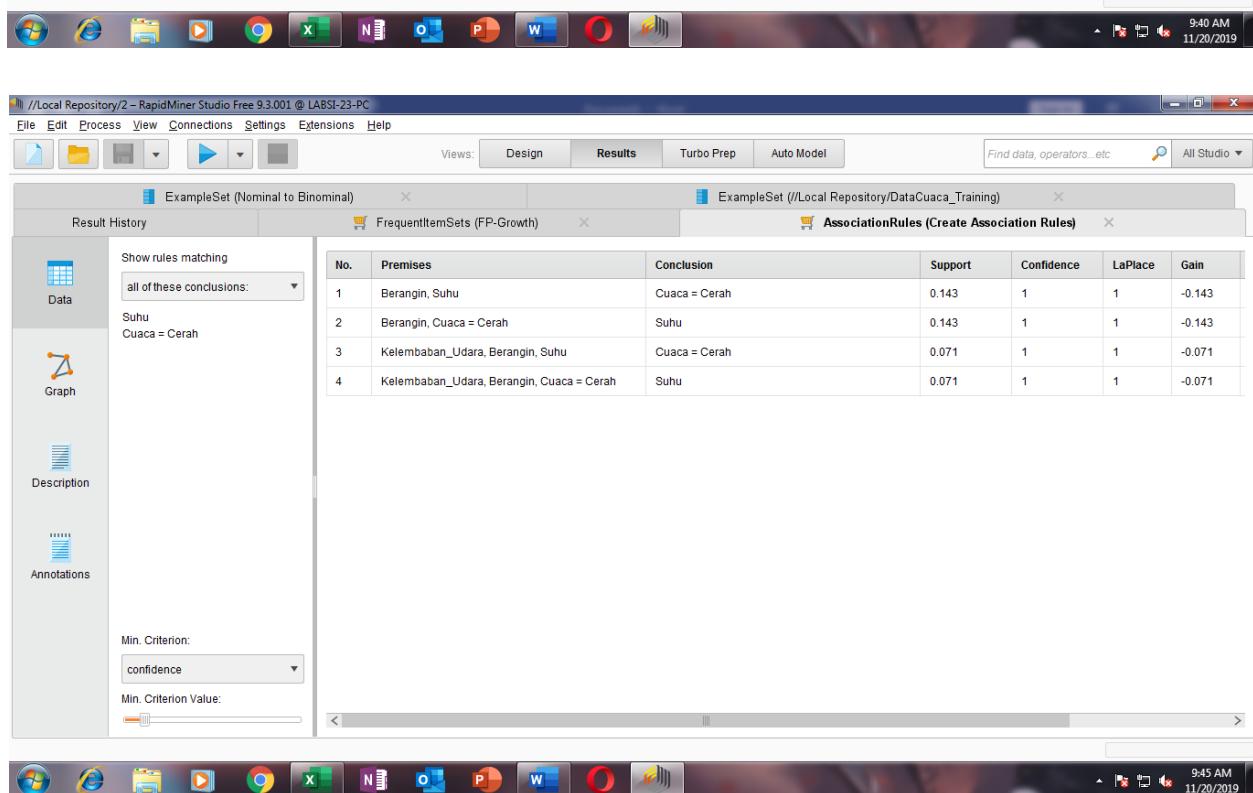
Data Graph Description Annotations

Show rules matching all of these conclusions:  
Suhu Cuaca = Cerah

No. Premises Conclusion Support Confidence LaPlace Gain

1	Berangin, Suhu	Cuaca = Cerah	0.143	1	1	-0.143
2	Berangin, Cuaca = Cerah	Suhu	0.143	1	1	-0.143
3	Kelembaban_Udara, Berangin, Suhu	Cuaca = Cerah	0.071	1	1	-0.071
4	Kelembaban_Udara, Berangin, Cuaca = Cerah	Suhu	0.071	1	1	-0.071

Min. Criterion: confidence  
Min. Criterion Value:



File Edit Process View Connections Settings Extensions Help

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

ExampleSet (Nominal to Binomial) ExampleSet (/Local Repository/DataCuaca\_Training) AssociationRules (Create Association Rules)

Result History FrequentItemSets (FP-Growth)

No. of Sets: 26	Size	Support	Item 1	Item 2	Item 3	Item 4
Total Max. Size: 4	1	0.500	Kelembaban_Udara			
Min. Size: 1	1	0.429	Berangin			
Max. Size: 4	1	0.429	Suhu			
Contains Item:	1	0.357	Cuaca = Cerah			
	1	0.357	Cuaca = Hujan			
	1	0.286	Cuaca = Mendung			
	2	0.214	Kelembaban_Udara	Berangin		
	2	0.214	Kelembaban_Udara	Suhu		
	2	0.214	Kelembaban_Udara	Cuaca = Cerah		
	2	0.143	Kelembaban_Udara	Cuaca = Hujan		
	2	0.143	Kelembaban_Udara	Cuaca = Mendung		
	2	0.143	Berangin	Suhu		
	2	0.143	Berangin	Cuaca = Cerah		
	2	0.143	Berangin	Cuaca = Hujan		
	2	0.143	Berangin	Cuaca = Mendung		

Data Annotations

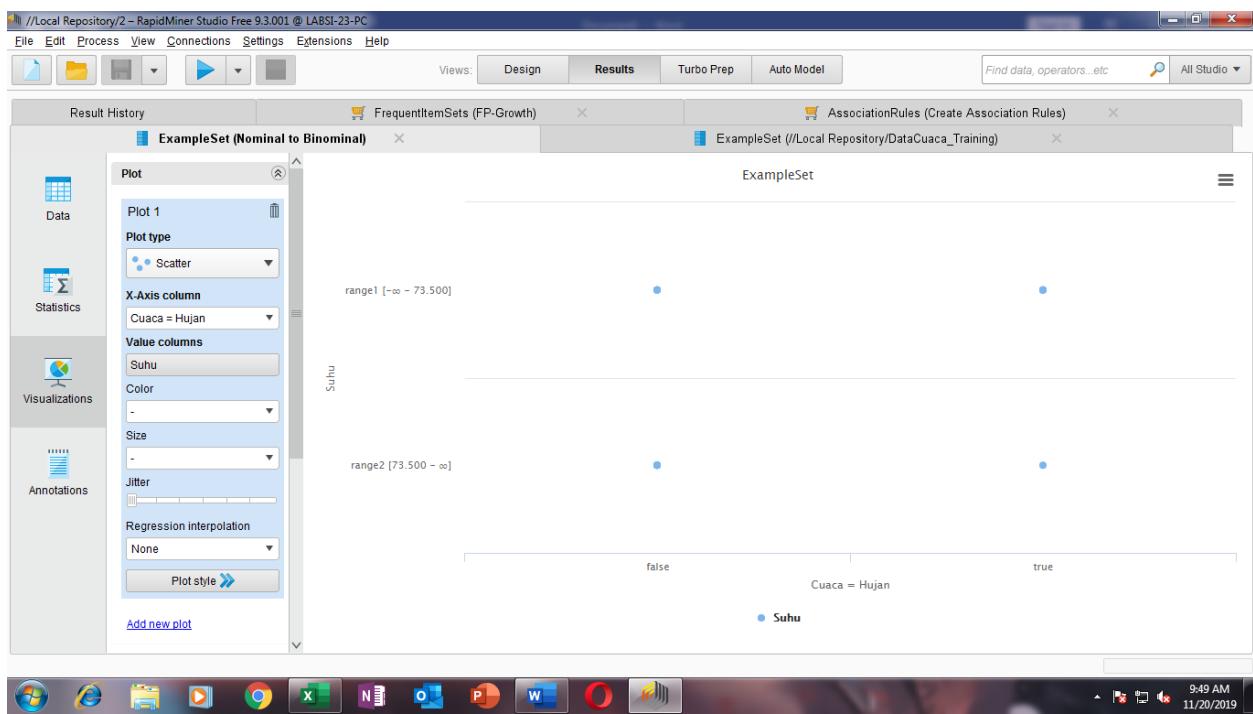
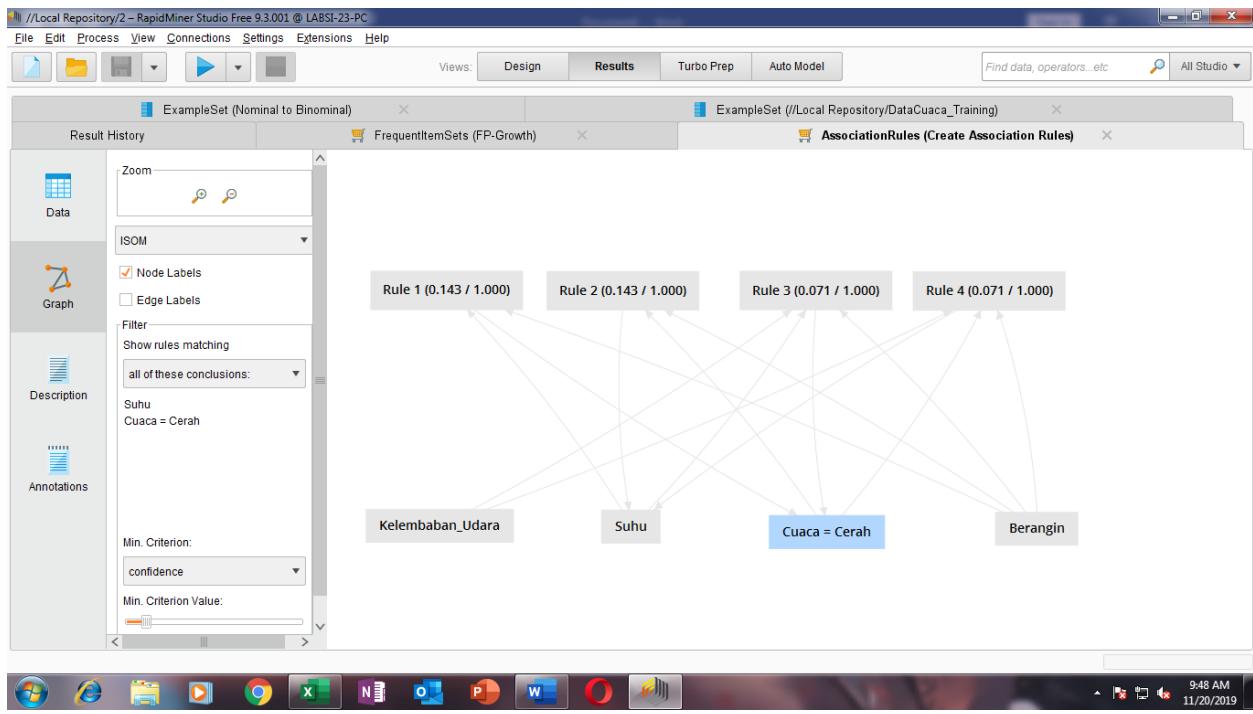
Update View



	Size	Support	Item 1	Item 2	Item 3	Item 4
	2	0.214	Suhu	Cuaca = Cerah		
	2	0.071	Suhu	Cuaca = Hujan		
	2	0.143	Suhu	Cuaca = Mendung		
	3	0.071	Kelembaban_Udara	Berangin	Suhu	
	3	0.071	Kelembaban_Udara	Berangin	Cuaca = Cerah	
	3	0.071	Kelembaban_Udara	Berangin	Cuaca = Hujan	
	3	0.071	Kelembaban_Udara	Berangin	Cuaca = Mendung	
	3	0.143	Kelembaban_Udara	Suhu	Cuaca = Cerah	
	3	0.071	Kelembaban_Udara	Suhu	Cuaca = Mendung	
	3	0.143	Berangin	Suhu	Cuaca = Cerah	
	4	0.071	Kelembaban_Udara	Berangin	Suhu	Cuaca = Cerah

Update View





## Tugas 1

The screenshot shows the RapidMiner Studio interface with the title bar //Local Repository/tugas 1 – RapidMiner Studio Free 9.3.001 @ LABSI-23-PC. The main window displays the RuleModel (Rule Induction) results. The rule defined is:

```
if Asisten = YA then TEPAT (0 / 5)
if Rerata_Sekolah > 19.500 then TEPAT (0 / 2)
else TERLAMBAT (6 / 5)

correct: 13 out of 18 training examples.
```

The screenshot shows the RapidMiner Studio interface with the title bar //Local Repository/tugas 1 – RapidMiner Studio Free 9.3.001 @ LABSI-23-PC. The main window displays the PerformanceVector (Performance) results. The accuracy is 60.00% +/- 31.62% (micro average: 60.00%). The confusion matrix is as follows:

	true TERLAMBAT	true TEPAT	class precision
pred. TERLAMBAT	0	1	0.00%
pred. TEPAT	7	12	63.16%
class recall	0.00%	92.31%	

## Tugas 2

### A number bin 2

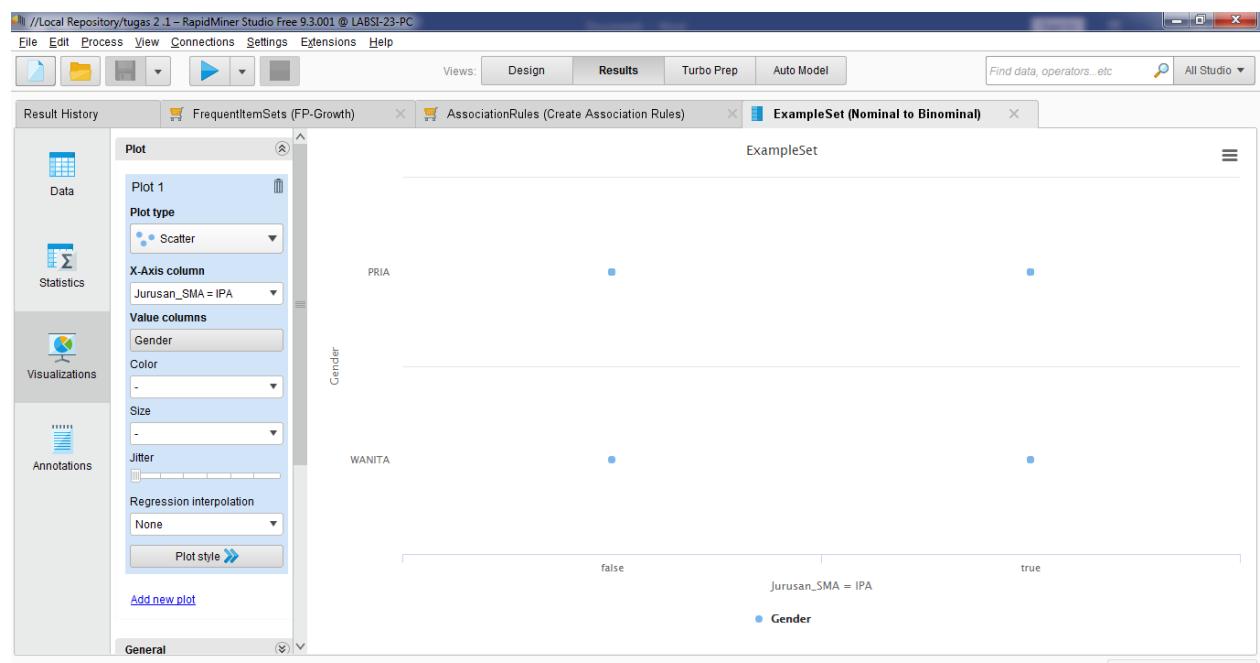
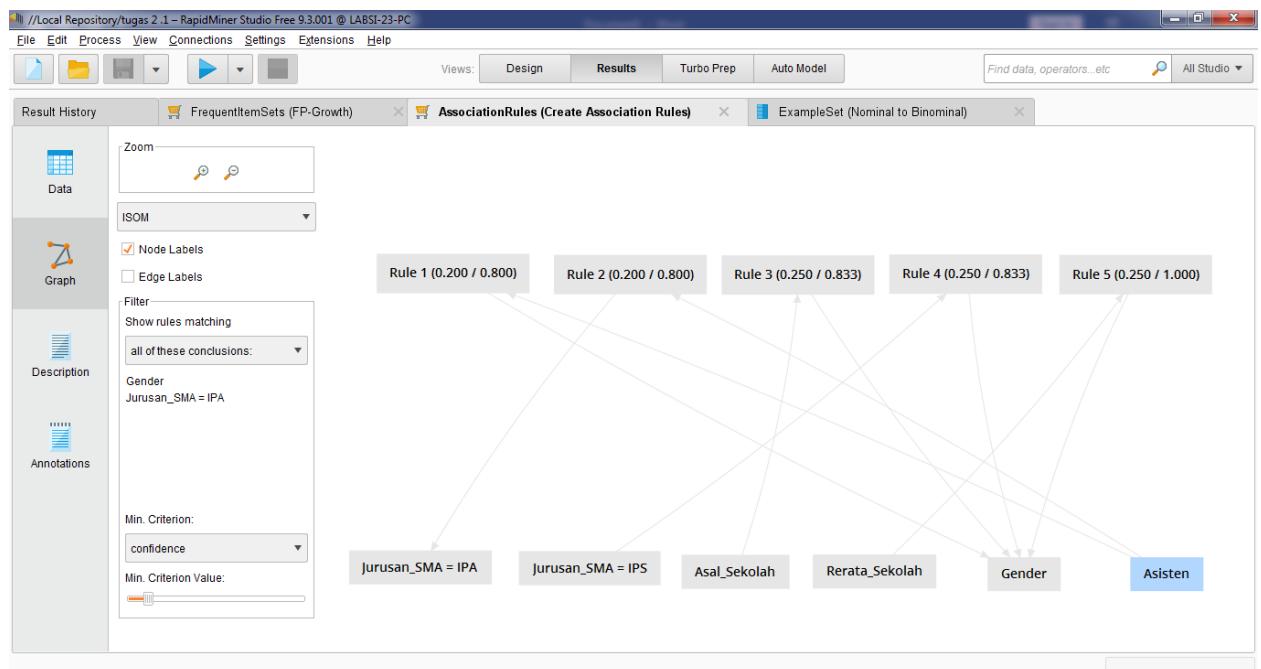
Result History      FrequentItemSets (FP-Growth)      AssociationRules (Create Association Rules)      ExampleSet (Nominal to Binomial)

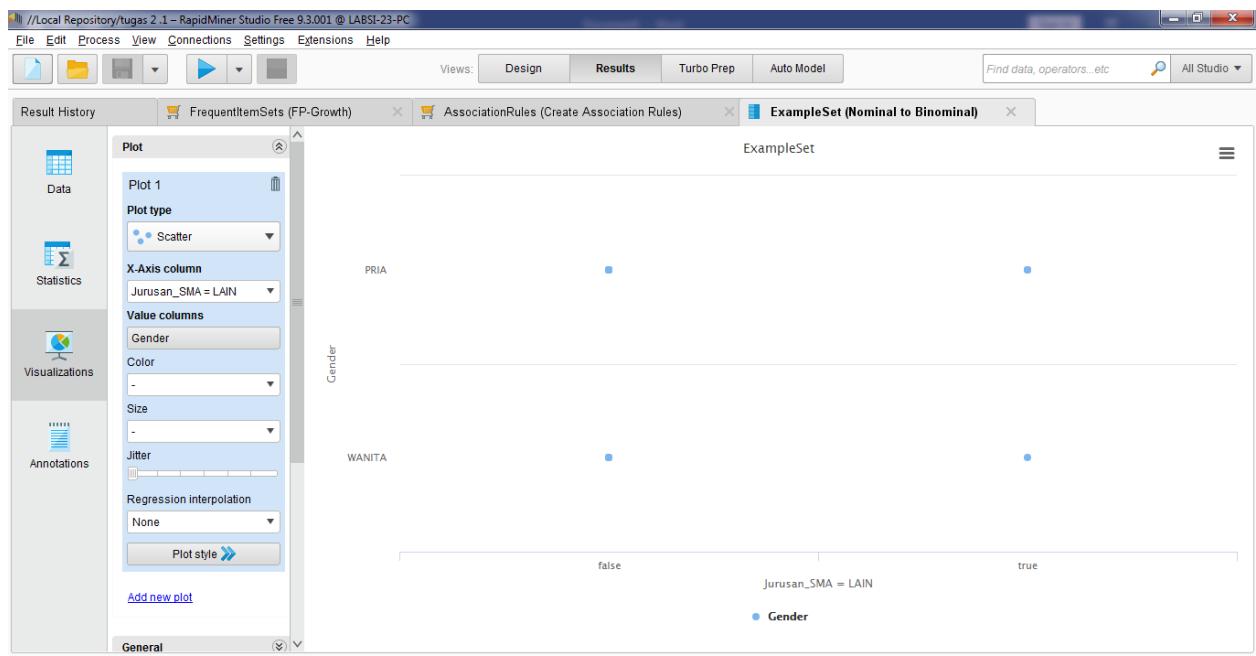
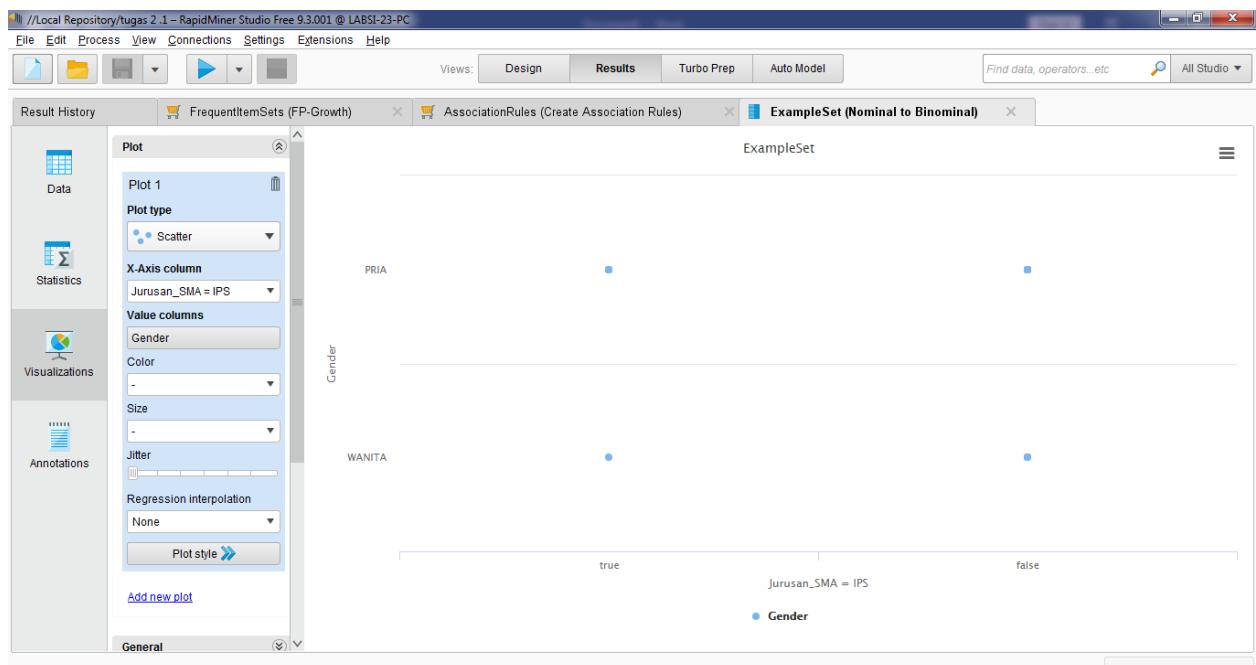
	No. of Sets: 13 Total Max. Size: 2	Size	Support	Item 1	Item 2
Data		1	0.750	Gender	
Annotations	Min. Size: 1 Max. Size: 2 Contains Item: <input type="text"/> Update View	1	0.500	Jurusan_SMA = IPA	
		1	0.300	Asal_Sekolah	
		1	0.300	Jurusan_SMA = IPS	
		1	0.250	Asisten	
		1	0.250	Rerata_Sekolah	
		1	0.200	Jurusan_SMA = LAIN	
		2	0.350	Gender	Jurusan_SMA = IPA
		2	0.250	Gender	Asal_Sekolah
		2	0.250	Gender	Jurusan_SMA = IPS
		2	0.200	Gender	Asisten
		2	0.250	Gender	Rerata_Sekolah
		2	0.200	Jurusan_SMA = IPA	Asisten

Result History      FrequentItemSets (FP-Growth)      AssociationRules (Create Association Rules)      ExampleSet (Nominal to Binomial)

No.	Premises	Conclusion	Support	Confidence	LaPlace	Gain	p-s
3	Asal_Sekolah	Gender	0.250	0.833	0.962	-0.350	0.025
4	Jurusan_SMA = IPS	Gender	0.250	0.833	0.962	-0.350	0.025
5	Rerata_Sekolah	Gender	0.250	1	1	-0.250	0.062

p-s	Lift	Convicti...
0.025	1.111	1.500
0.025	1.111	1.500
0.062	1.333	$\infty$





## Tugas 2

### B number 3

//Local Repository/tugas 2.b – RapidMiner Studio Free 9.3.001 @ LABSI-23-PC

File Edit Process View Connections Settings Extensions Help

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

AssociationRules (Create Association Rules) ExampleSet (Nominal to Binomial) PerformanceVector (Performance)

Result History FrequentItemSets (FP-Growth)

No. of Sets: 19 Total Max. Size: 2

Data Annotations

Size Support Item 1 Item 2

Size	Support	Item 1	Item 2
1	0.750	Gender	
1	0.500	Jurusan_SMA = IPA	
1	0.400	Rerata_Sekolah = range1 [-∞ - 18.500]	
1	0.350	Rerata_Sekolah = range2 [18.500 - 19.500]	
1	0.300	Asal_Sekolah	
1	0.300	Jurusan_SMA = IPS	
1	0.250	Asisten	
1	0.250	Rerata_Sekolah = range3 [19.500 - ∞]	
1	0.200	Jurusan_SMA = LAIN	
2	0.350	Gender	Jurusan_SMA = IPA
2	0.200	Gender	Rerata_Sekolah = range1 [-∞ - 18.500]
2	0.300	Gender	Rerata_Sekolah = range2 [18.500 - 19.500]
2	0.250	Gender	Asal_Sekolah
2	0.250	Gender	Jurusan_SMA = IPS
2	0.200	Gender	Asisten

Sending statistics

2	0.250	Gender	Rerata_Sekolah = range3 [19.500 - ∞]
2	0.200	Jurusan_SMA = IPA	Rerata_Sekolah = range1 [-∞ - 18.500]
2	0.200	Jurusan_SMA = IPA	Rerata_Sekolah = range2 [18.500 - 19.500]
2	0.200	Jurusan_SMA = IPA	Asisten

//Local Repository/tugas 2.b – RapidMiner Studio Free 9.3.001 @ LABSI-23-PC

File Edit Process View Connections Settings Extensions Help

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

Result History

AssociationRules (Create Association Rules) x FrequentItemSets (FP-Growth) x ExampleSet (Nominal to Binominal) x PerformanceVector (Performance) x

Data Graph Description Annotations

Show rules matching: all of these conclusions: Gender Jurusan\_SMA = IPA

No. Premises Conclusion Support Confidence LaPlace Gain p-s

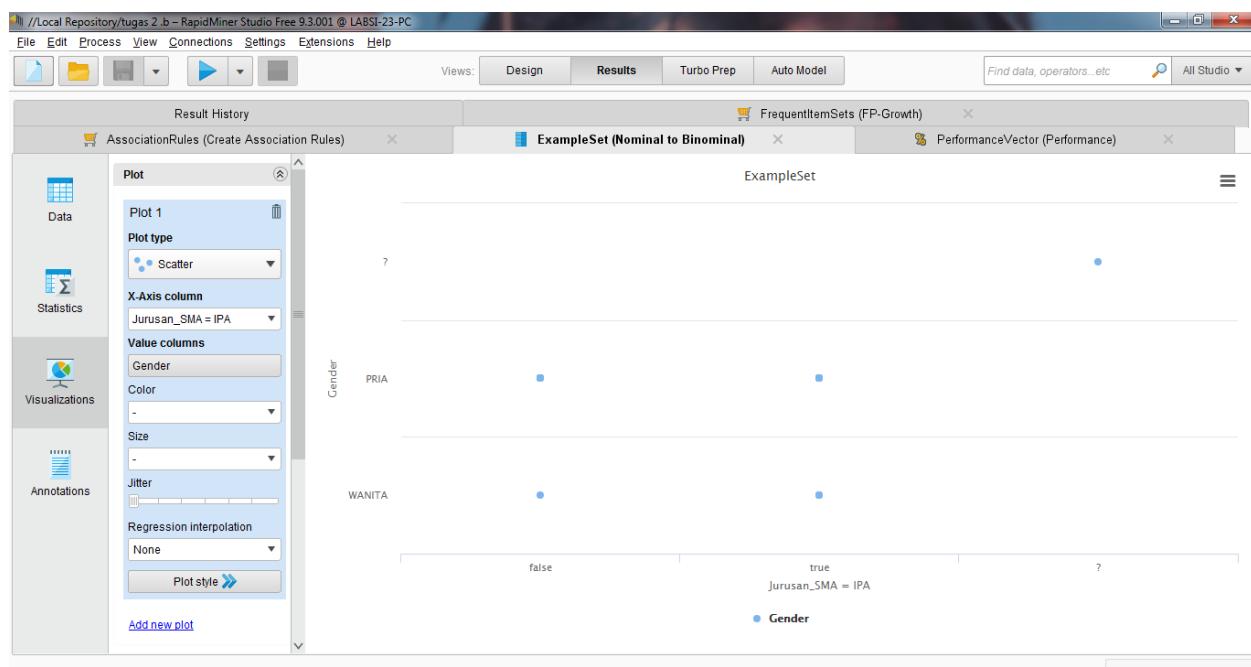
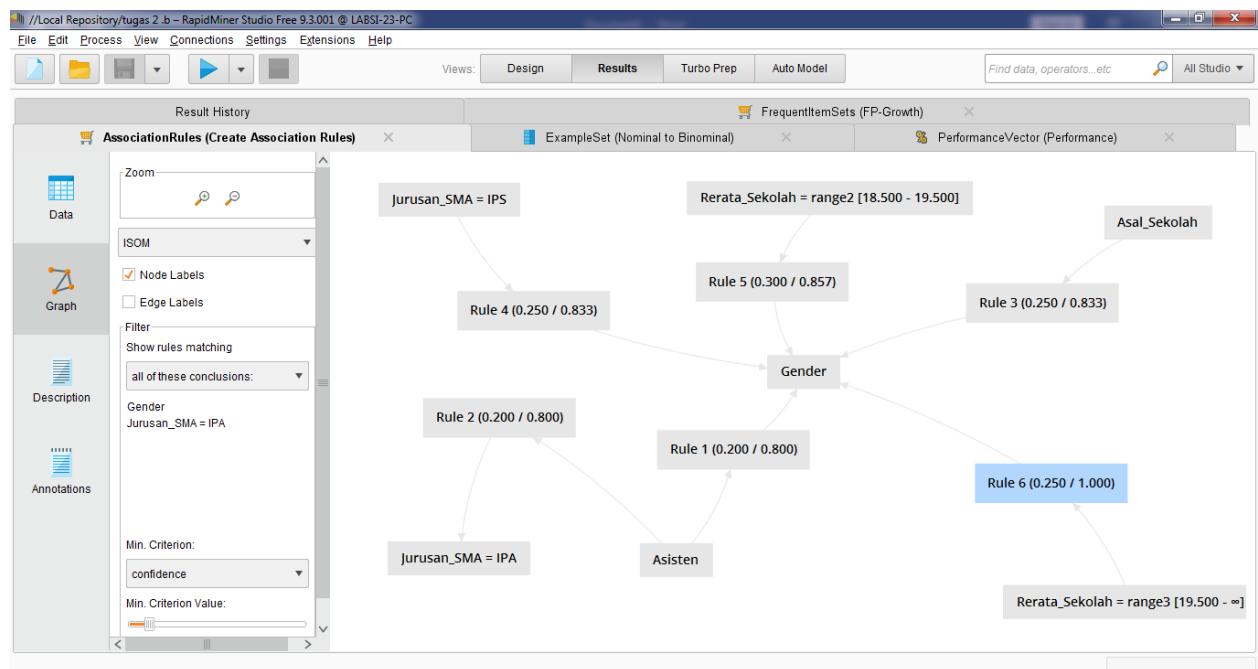
3	Asal_Sekolah	Gender	0.250	0.833	0.962	-0.350	0.025
4	Jurusan_SMA = IPS	Gender	0.250	0.833	0.962	-0.350	0.025
5	Rerata_Sekolah = range2 [18.500 - 19.500]	Gender	0.300	0.857	0.963	-0.400	0.038
6	Rerata_Sekolah = range3 [19.500 - ∞]	Gender	0.250	1	1	-0.250	0.062

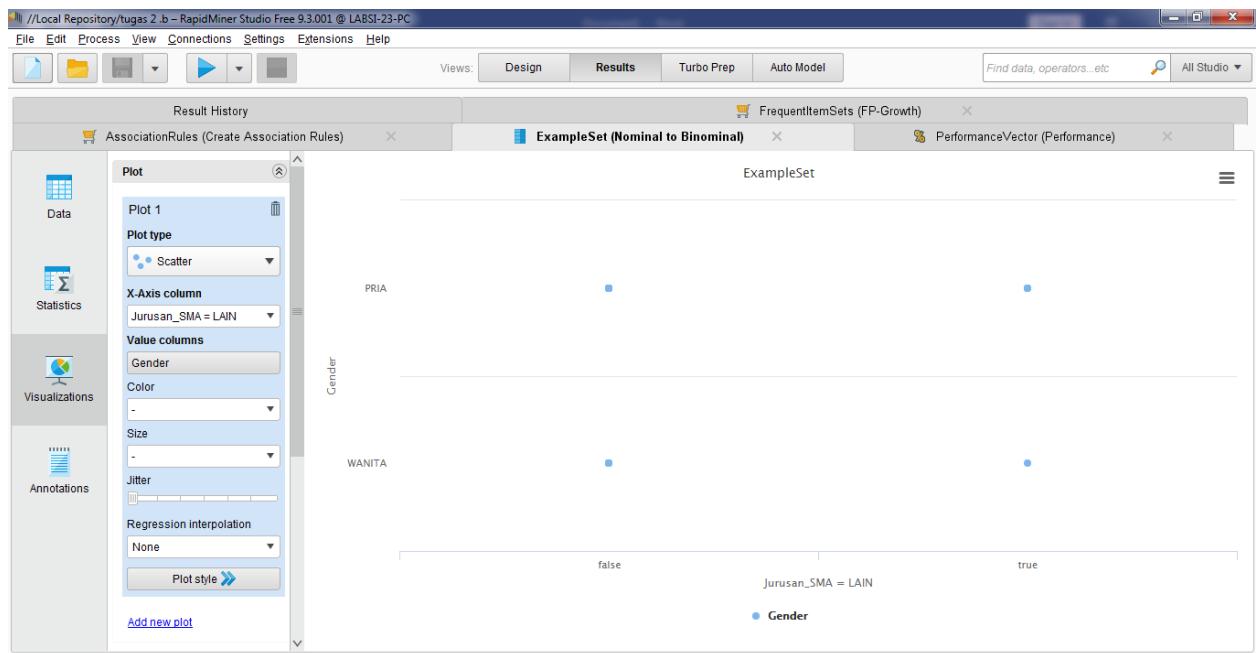
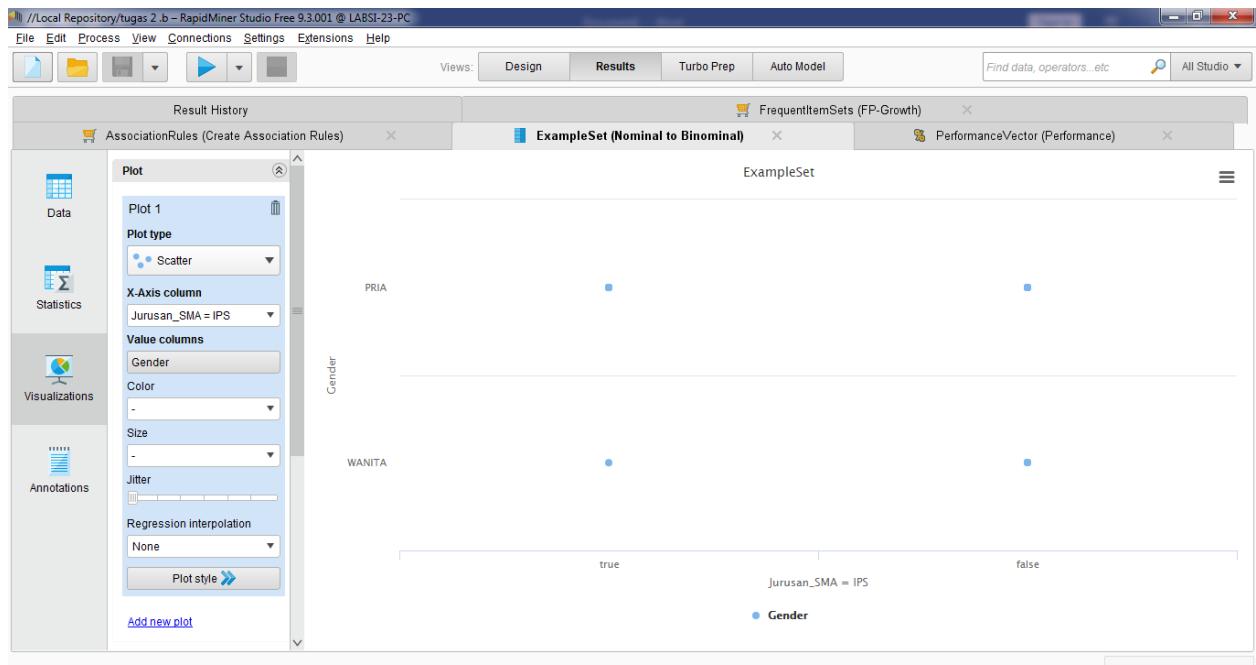
Min. Criterion: confidence  
Min. Criterion Value:

All Studio

Lift Convicti...

1.111	1.500
1.111	1.500
1.143	1.750
1.333	∞





## MODUL 12

Percobaan :

1	NO_SISWA	NAMA	LAMA BELAJAR	NILAI
2	S-101	JOKO	15	783
3	S-102	AGUS	18	877
4	S-103	SUSI	7	505
5	S-104	DYAH	9	860
6	S-105	WATI	15	968
7	S-106	IKA	17	793
8	S-107	EKO	10	752
9	S-108	YANTO	5	571
10	S-109	WAWAN	8	667
11	S-110	MAHMUD	15	723
12				

1. import data diatas ke rapid miner
2. ubah tipe data dan jenis masing-masing atribut sebagai berikut .
  - a) No\_SISWA : polynomial,id
  - b) NAMA : pilih Exclude column
  - c) LAMA JAM BELAJAR : integer
  - d) NILAI : integer, label

Import Data - Format your columns.

**Format your columns.**

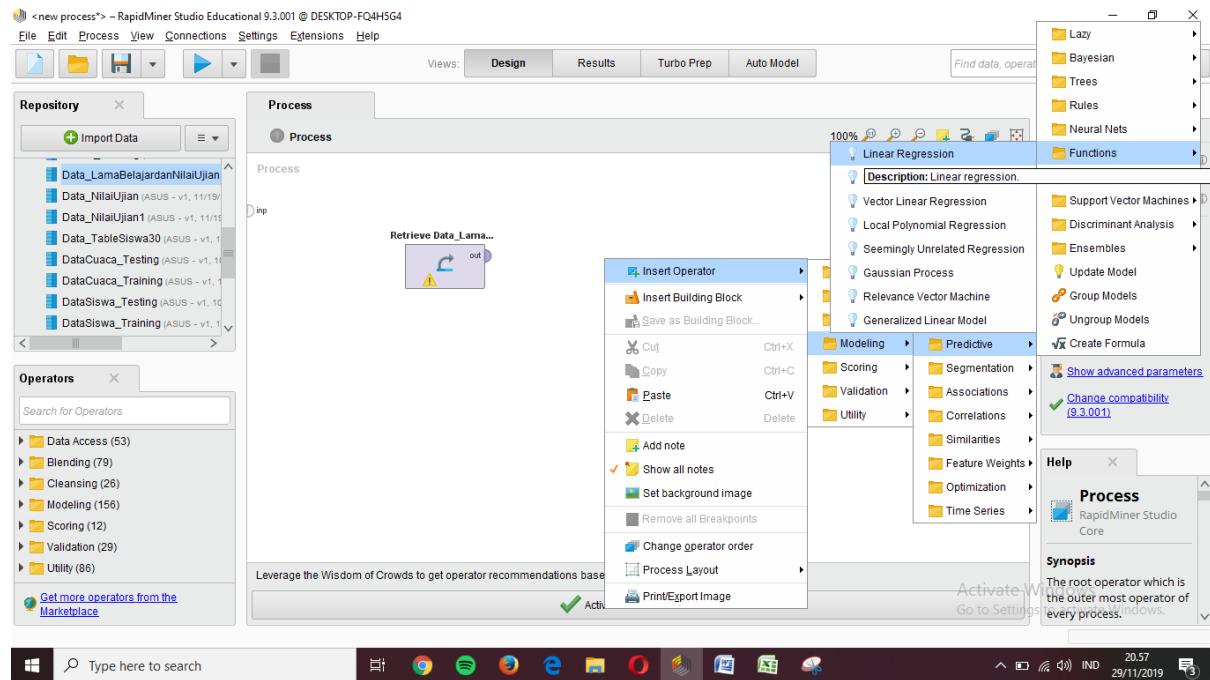
Replace errors with missing values  ⓘ

NO_SISWA <i>polynominal id</i>	NAMA <i>polynominal</i>	LAMA BELAJAR <i>integer</i>	NILAI <i>integer label</i>
1 S-101	JOKO	15	783
2 S-102	AGUS	18	877
3 S-103	SUSI	7	505
4 S-104	DYAH	9	860
5 S-105	WATI	15	968
6 S-106	IKA	17	793
7 S-107	EKO	10	752
8 S-108	YANTO	5	571
9 S-109	WAWAN	8	667
10 S-110	MAHMUD	15	723

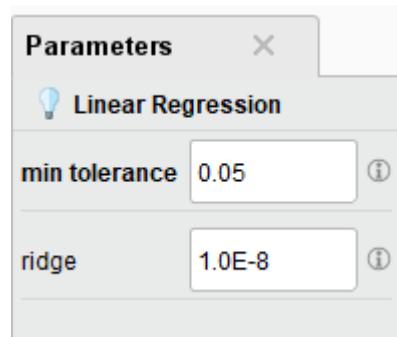
 no problems.

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

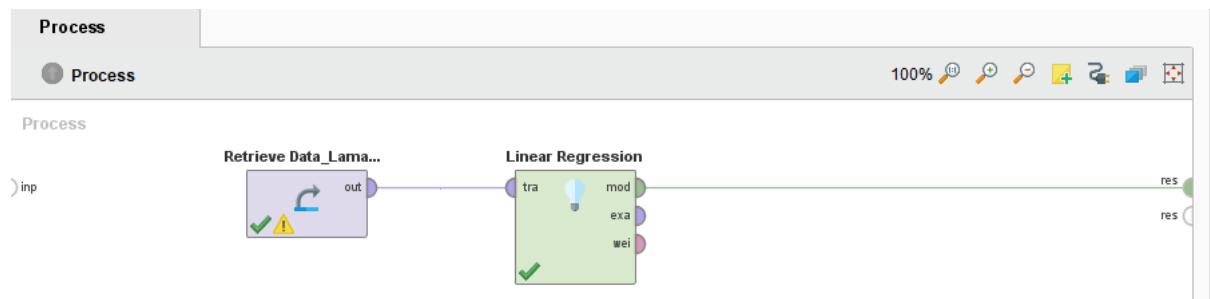
3. lalu beri nama LamaBelajarNilaiUjian kemudian klik finish
4. gunakan data tersebut dan masukan ke dalam area process
5. Tambahkan operator Modeling → Predictive → Functions → Linear Regression. Lalu hubungkan port out dan 3 outputnya ke connector res process

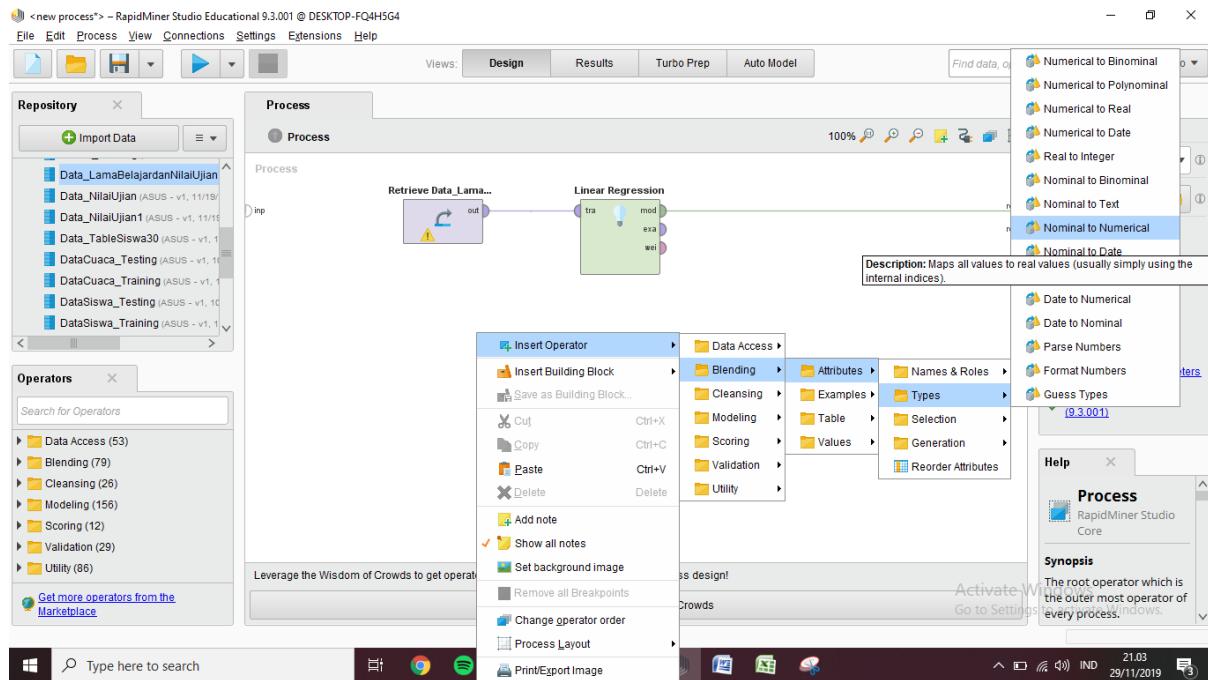


6. Klik pada Linear Regression, tentukan parameter min tolerance = 0.05



7. jika data input bertipe nominal atau polynomial tambahkan operator Blending → atribut  
→ Types → Nominal to Numerical tepat setelah data di input, sebelum operator linear  
Reggression





8. lalu jalankan proses dengan menekan run

9. hasil proses regresi linier :

a) Table View (mencari besarnya nilai t-hitung)

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

Dapat dilihat nilai t-statistic (t-hitung) sebesar 2,827

b) Text view (mencari model regresi)

Result History	LinearRegression (Linear Regression)
	<pre> ExampleSet //Local Repository/Data_LamaBelajardanNilaiUjian LinearRegression LinearRegression 21.608 * LAMA BELAJAR + 492.769 </pre>

Dari hasil text view diatas terlihat persamaan berikut :

$$= 21.608 * \text{LAMA JAM BELAJAR} + 492.769$$

Berikut model regresi linier yang terbentuk :

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

## Percobaan ke 2

	A	B	C	
1	NO_SISWA	NAMA	LAMA BELAJAR (JAM)	
2	S-111	BUDI	12	
3	S-112	SANTI	13	
4	S-113	DIAN	14	
5	S-114	DANI	11	
6	S-115	AHMAD	5	
7	S-116	BAYU	13	
8	S-117	RISA	9	
9	S-118	RANI	10	
10	S-119	YANI	10	
11	S-120	RATIH	9	
12				

1. gunakan file ini sebagai testing. Import pada rapid miner
2. ubah tipe data dan jenis masing-masing atribut sebagai berikut .
  - a) No\_SISWA : polynomial,id
  - b) NAMA : pilih Exclude column
  - c) LAMA JAM BELAJAR : integer

Replace errors with missing values. ⓘ

NO_SISWA polynomial id	NAMA polynomial	LAMA BELAJAR (JAM) integer
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. beri nama Data\_PrediksiNilaiUjian dan masukkan pada repositori, kemudian klik finish
4. tambahkan operator Scoring → Apply Model letakkan setelah operator linear reggresin. Hubungkan port port output dan input

<new process> – RapidMiner Studio Educational 9.3.001 @ DESKTOP-FQ4H5G4

File Edit Process View Connections Settings Extensions Help

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

Repository X Import Data

Process X Process

Parameters X

Process logverbosity init logfile

Operators X Search for Operators

Operators Data Access (53) Blending (79) Cleansing (26) Modeling (156) Scoring (12) Validation (29) Utility (66) Get more operators from the Marketplace

Process

Retrieve Data\_Pred...

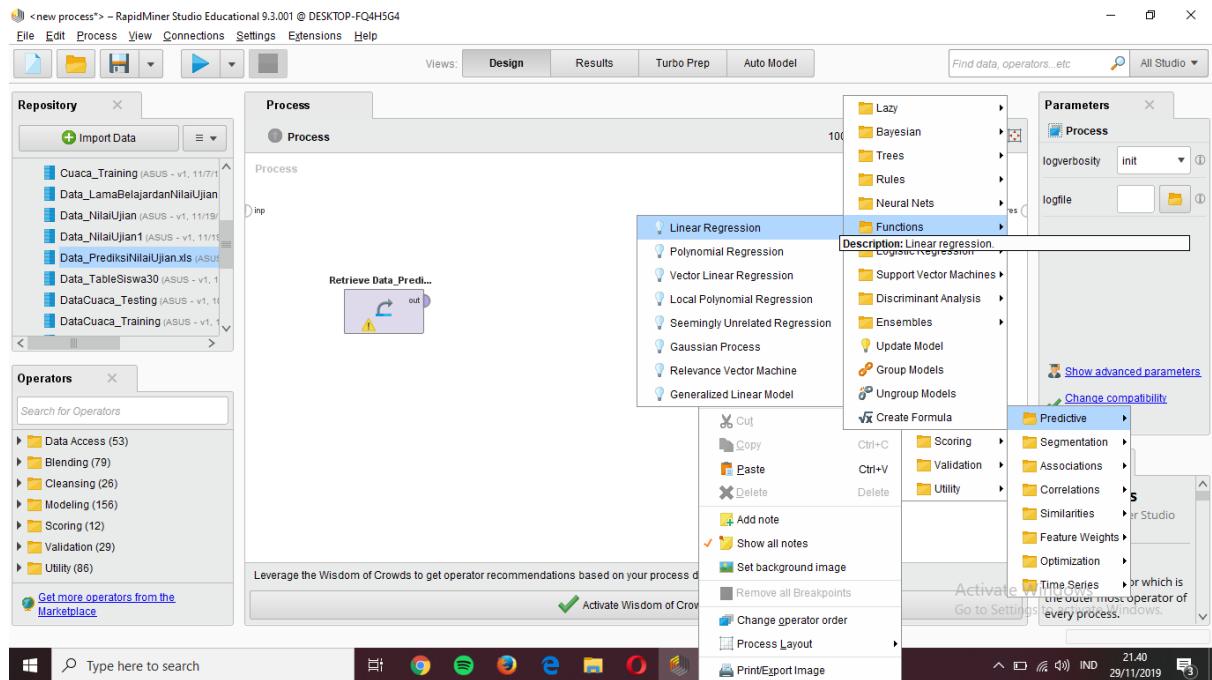
Insert Operator Data Access Blending Cleansing Modeling Scoring Confidences Validation Utility Explain Predictions Prescriptive Analytics

Description: Applies a model to an example set. This might be a prediction or another data transformation model.

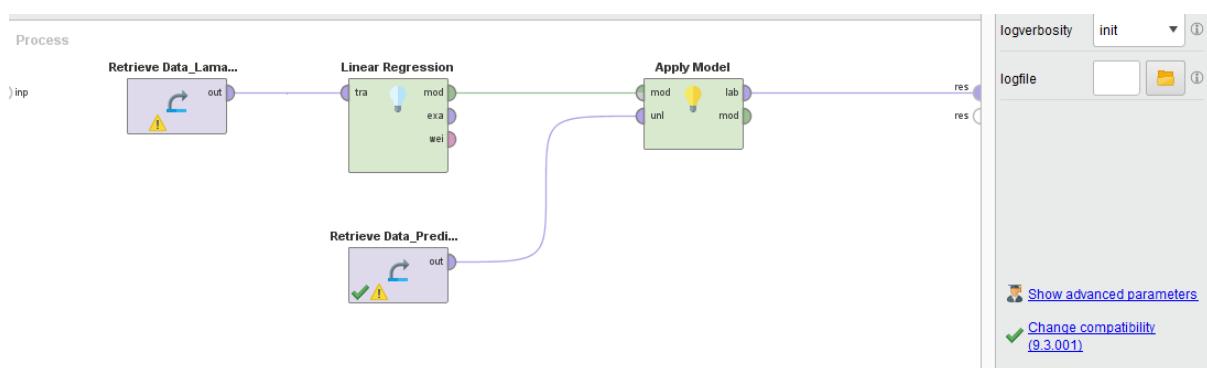
Scoring Apply Model

Process RapidMiner Studio Core Synopsis The root operator which is the outer most operator of every process.

Type here to search



## 5. lalu run



## 6. hasil proses prediksi terhadap data testing menggunakan regresi linier

- a) Data view (hasil prediksi nilai ujian )

**RapidMiner Studio Free 9.3.001 @ LABSI-22-PC**

File Edit Process View Connections Settings Extensions Help

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

Result History ExampleSet (/Local Repository/Data\_LamaBelajarNilaiUjian) ExampleSet (Apply Model) ExampleSet (/Local Repository/Data\_PrediksiNilaiUjian)

Data Statistics Visualizations Annotations

Open in Turbo Prep Auto Model Filter (10 / 10 examples): all

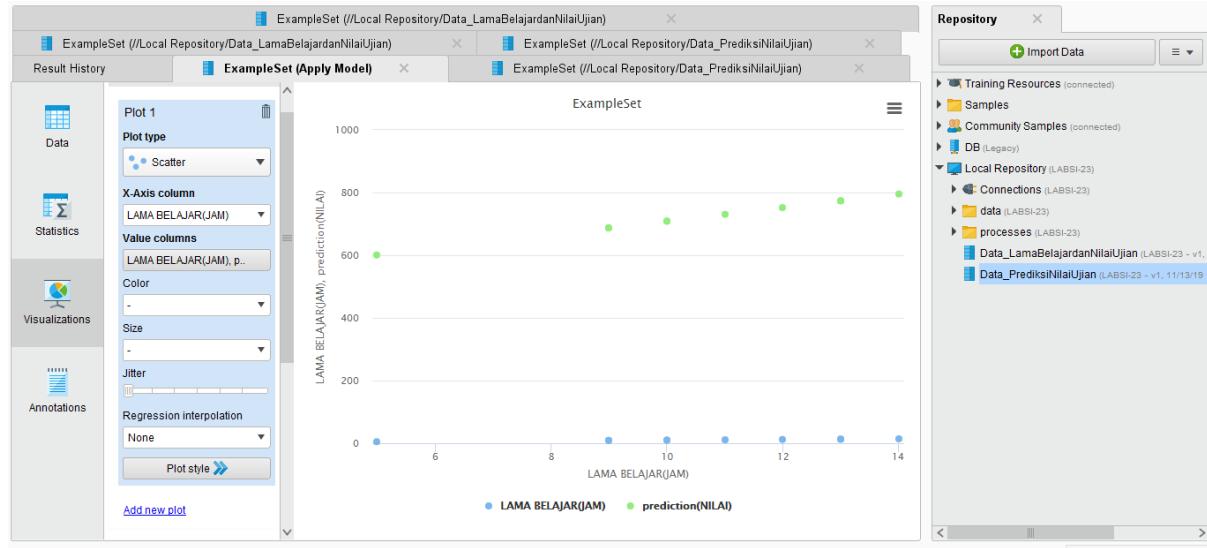
Row No.	NO_SISWA	prediction(NILAI)	LAMA BELAJAR(JAM)
1	S-111	752.061	12
2	S-112	773.668	13
3	S-113	795.276	14
4	S-114	730.453	11
5	S-115	600.807	5
6	S-116	773.668	13
7	S-117	687.238	9
8	S-118	708.845	10
9	S-119	708.845	10
10	S-120	687.238	9

Repository Import Data

- Training Resources (connected)
- Community Samples (connected)
- Samples
- DB (Legacy)
- Local Repository (LABSI-22)
  - Connections (LABSI-22)
  - data (LABSI-22)
  - processes (LABSI-22)
    - Data\_LamaBelajarNilaiUjian (LABSI-22 - v1, 11/13/19, 1 example)
    - Data\_PrediksiNilaiUjian (LABSI-22 - v1, 11/13/19, 1 example)

Hasil prediksi nilai ujian terhadap 10 siswa lainnya

### b) Charts View (Scatter Plot)



## Tugas 1

Import Data - Format your columns.

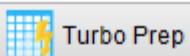
Replace errors with missing values  ⓘ

**Format your columns.**

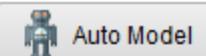
NO. RESPONDEN	PENDAPATAN (RUPIAH)	JUMLAH ANGGOTA K...	DAYA BELI (RUPIAH)
1	1000000	6	834000
2	1400000	7	1200000
3	200000	3	134000
4	1400000	6	1167000
5	500000	3	334000
6	1700000	5	1360000
7	400000	3	267000
8	1900000	5	1520000
9	300000	3	200000
10	500000	4	375000
11	700000	7	600000
12	1900000	3	1267000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

no problems.

Open in



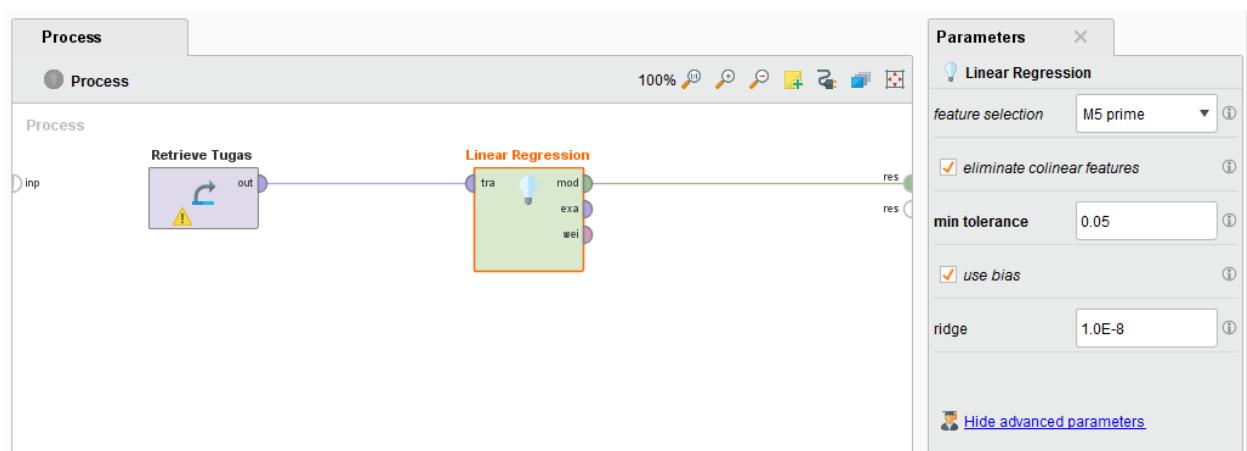
Turbo Prep



Auto Model

Row No.	NO. RESPON...	DAYA BELI (...)	PENDAPATA...	JUMLAH AN...
1	1	834000	1000000	6
2	2	1200000	1400000	7
3	3	134000	200000	3
4	4	1167000	1400000	6
5	5	334000	500000	3
6	6	1360000	1700000	5
7	7	267000	400000	3
8	8	1520000	1900000	5
9	9	200000	300000	3
10	10	375000	500000	4
11	11	600000	700000	7
12	12	1267000	1900000	3
13	13	600000	800000	4

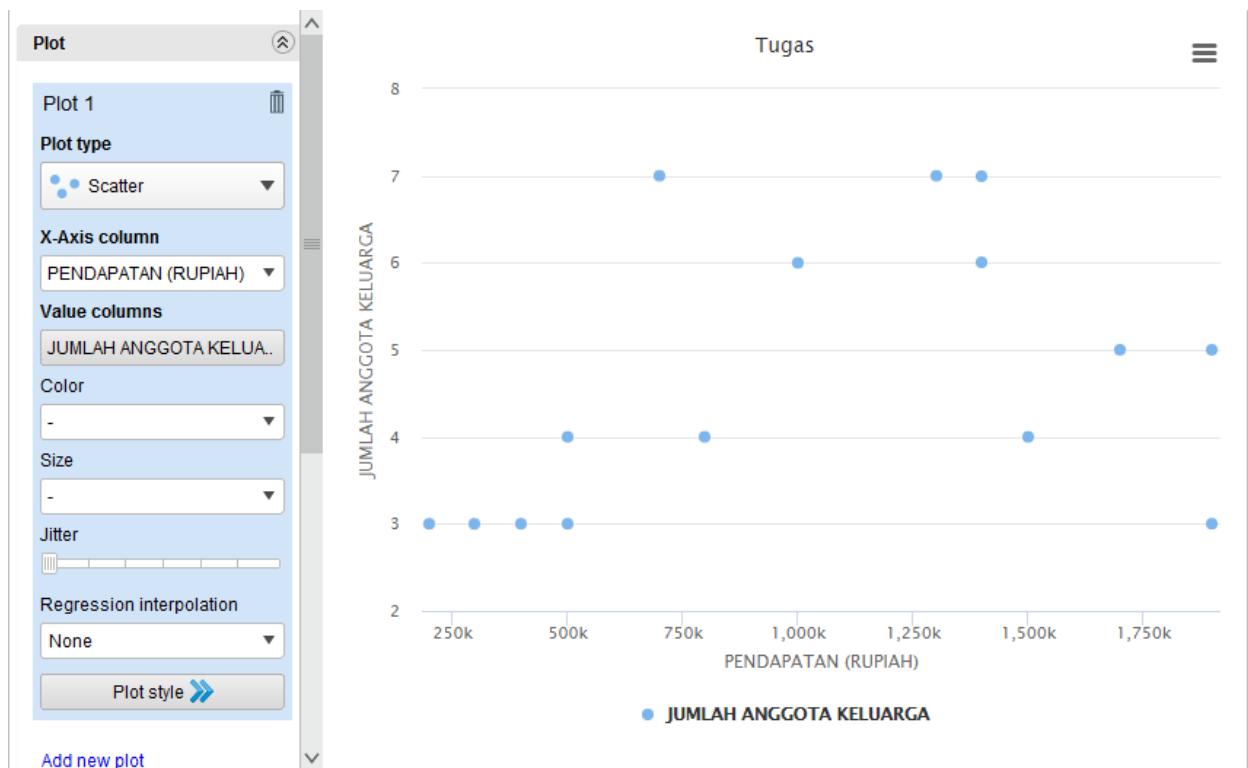
ExampleSet (15 examples, 2 special attributes, 2 regular attributes)



Attribute	Coefficient	Std. Error	Std. Coeffici...	Tolerance	t-Stat	p-Value	Code
PENDAPATAN...	0.739	0.021	0.924	0.857	35.037	0.000	****
JUMLAH ANG...	47807.624	7833.319	0.161	0.857	6.103	0.000	****
(Intercept)	-180222.487	36497.284	?	?	-4.938	0.000	****

# Linear Regression

0.739 \* PENDAPATAN (RUPIAH)  
+ 47807.624 \* JUMLAH ANGGOTA KELUARGA  
- 180222.487



## Tugas 2

Import Data - Format your columns.

Replace errors with missing values  ⓘ

**Format your columns.**

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

✓ no problems.

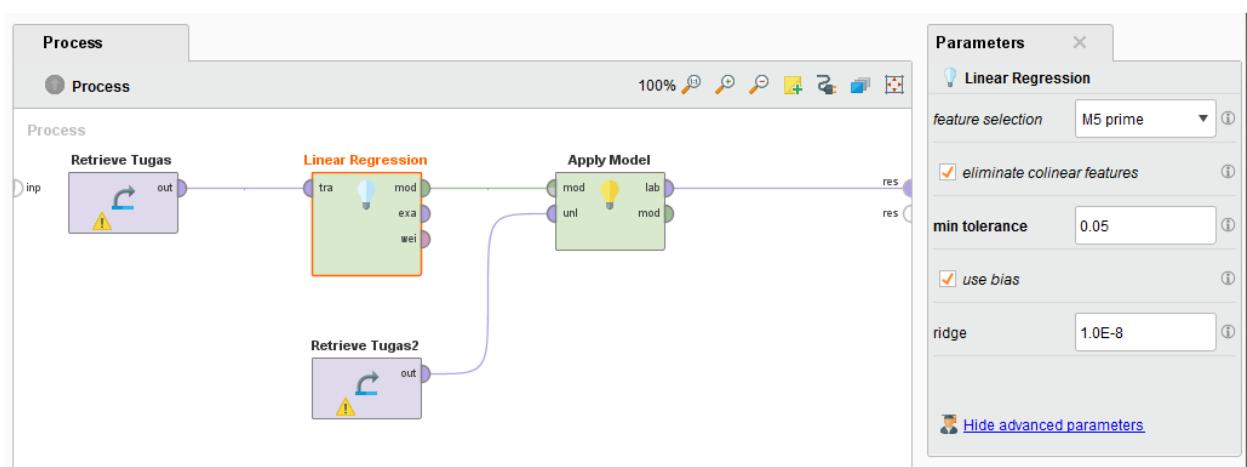
← Previous → Next X Cancel

Open in

Turbo Prep

Auto Model

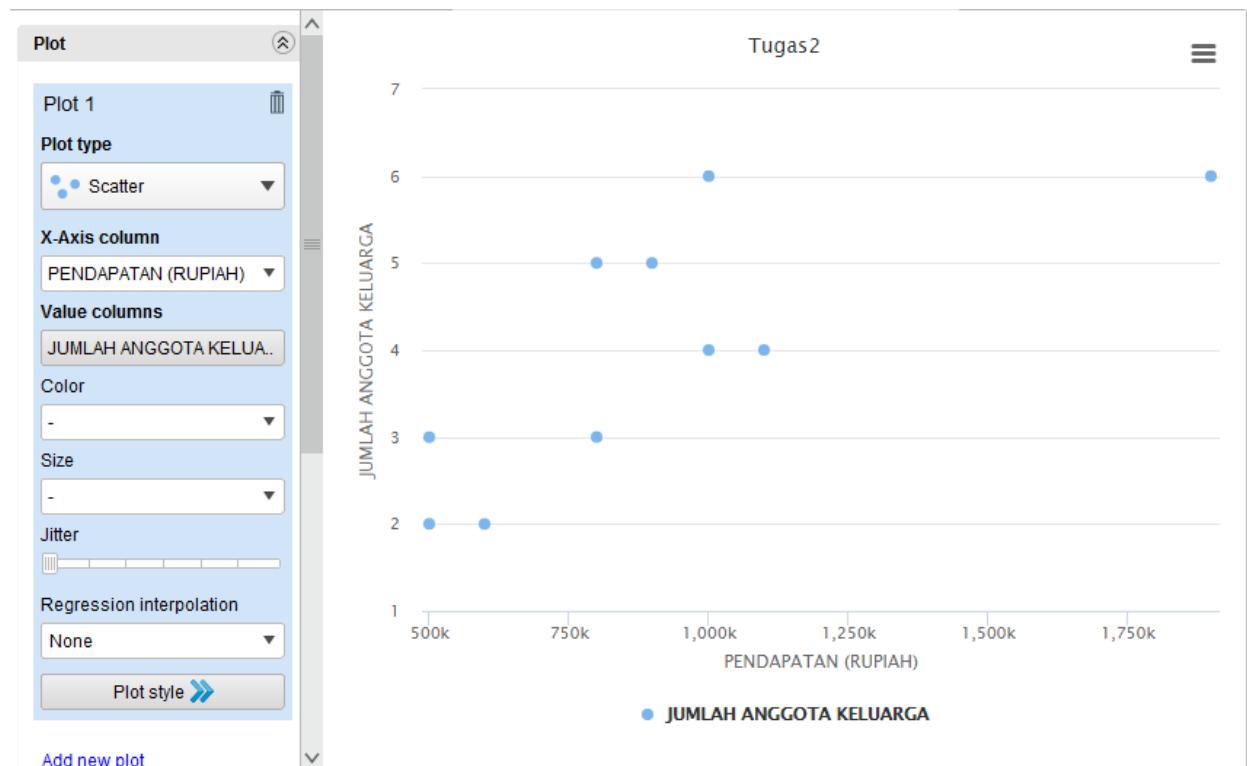
Row No.	NO. RESPON...	JUMLAH AN...	PENDAPATA...
1	1	5	900000
2	2	3	800000
3	3	2	500000
4	4	6	1900000
5	5	2	600000
6	6	5	800000
7	7	6	1000000
8	8	4	1100000
9	9	4	1000000
10	10	3	500000



Open in [Turbo Prep](#)

[!\[\]\(b41fb99fc024f49cd078ae6414d5de01\_img.jpg\) Auto Model](#)

Row No.	NO. RESPON...	JUMLAH AN...	prediction(D...	PENDAPATA...
1	1	5	723933.263	900000
2	2	3	554416.056	800000
3	3	2	284902.556	500000
4	4	6	1510760.476	1900000
5	5	2	358804.515	600000
6	6	5	650031.304	800000
7	7	6	845642.845	1000000
8	8	4	823929.557	1100000
9	9	4	750027.598	1000000
10	10	3	332710.179	500000



NO. RESPONDEN	PENDAPATAN (RUPIAH)	JUMLAH ANGGOTA KELUARGA	Y
1	900000	5	723915.633
2	800000	3	554400.385
3	500000	2	284892.761
4	1900000	6	1510723.257
5	600000	2	358792.761
6	800000	5	650015.633
7	1000000	6	845623.257
8	1100000	4	823908.009
9	1000000	4	750008.009
10	500000	3	332700.385