

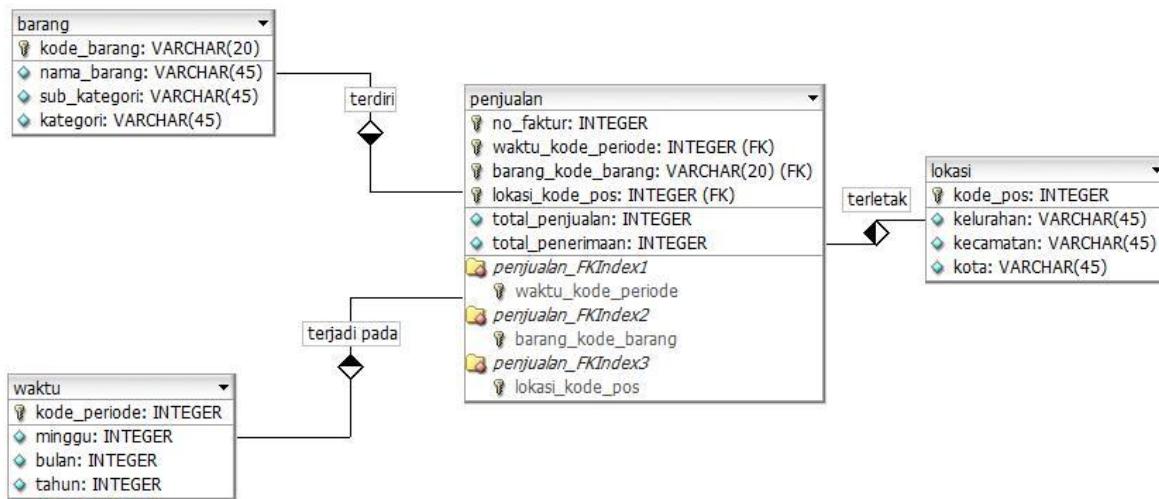
**REKAP MODUL PRAKTIKUM
DATA WAREHOUSE & DATA MINING**



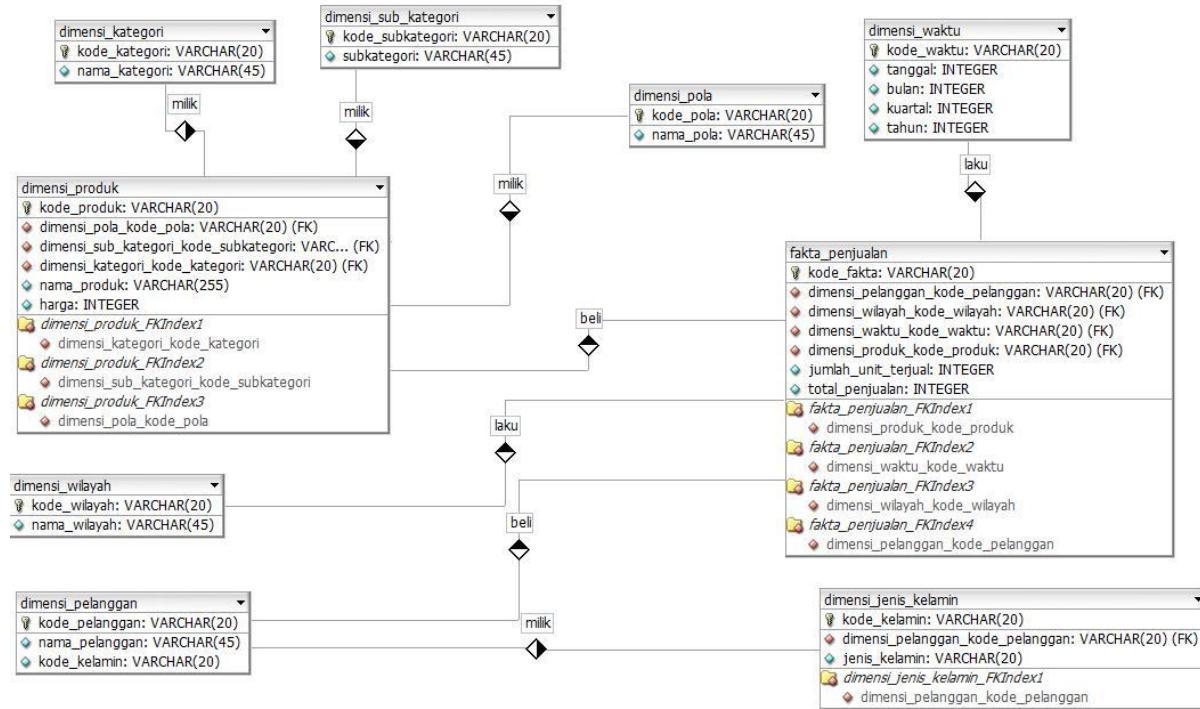
DISUSUN OLEH :
ZULFIKAR BAYU BUDIMAN (L200170052)
Kelas C

Modul 1 (Perancangan Star Schema dan Snowflake)

Star Schema



Tugas



Modul 5 (Pivot Table dan Chart)

1. Data Fact Table

A	B	C	D	E	F	G	H	I	J	K	L
bulan	kuartal	tauhun	nama_produk	nama_kategori	nama_subkategori	nama_pola	nama_pelanggan	jenis_kelamin	nama_wilayah	jumlah	harga
2	12	4	2011 Jarik Standar	Standar	Jarik	Print	Bapak Ketut	PRIA	Bali	2	225000
3	1	1	2012 Kaos Batik	Batik	Kaos	Cap	Ibu Harini	WANITA	Jawa Timur	14	30000
4	4	2	2012 Jarik Standar	Standar	Jarik	Tulis	Ibu Harini	WANITA	Jawa Timur	4	40000
5	4	2	2011 Hem Katun	Katun	Hem	Print	Ibu Harini	WANITA	Jawa Timur	3	70000
6	9	3	2012 Batik Standar	Standar	Batik	Cap	Bapak Heru	PRIA	Jawa Timur	1	150000
7	5	2	2012 Hem Katun	Katun	Hem	Print	Bapak Totok	PRIA	Jawa Timur	3	299000
8	12	4	2011 Bolero Standar	Standar	Bolero	Cap	Ibu Hatamah	WANITA	Jawa Timur	1	225000
9	10	4	2011 Sarimbit Standar	Standar	Sarimbit	Print	Ibu Hatamah	WANITA	Jawa Timur	1	150000
10	1	1	2011 Kaos Katun	Katun	Kaos	Print	Bapak Imron	PRIA	Jawa Barat	1	60000
11	2	1	2012 Celana Standar	Standar	Celana	Cap	Ibu Hadi Sukarni	WANITA	Jawa Barat	17	55000
12	3	1	2010 Celana Standar	Standar	Celana	Print	Ibu Hadi Sukarni	WANITA	Jawa Barat	17	55000
13	3	1	2011 Bahan Standar	Standar	Bahan	Cap	Ibu Siti Arya	WANITA	Jawa Barat	8	120000
14	12	4	2012 Rok Batik	Batik	Rok	Print	Ibu Siti Arya	WANITA	Jawa Barat	1	225000
15	1	1	2012 Jam Standar	Standar	Jam	Print	Ibu Siti Arya	WANITA	Jawa Barat	44	80000
16	9	3	2012 Hem Standar	Standar	Hem	Cap	Ibu Aini Kasmaji	WANITA	Jawa Tengah	1	100000
17	6	2	2012 Bahan Lawasan	Lawasan	Bahan	Tulis	Ibu Niken	WANITA	Jawa Tengah	1	130000
18	8	3	2011 Hem Standar	Standar	Hem	Tulis	Ibu Atik	WANITA	Jawa Tengah	5	55000
19	4	2	2012 Bahan Standar	Standar	Bahan	Cap	Ibu Tyas	WANITA	Jawa Tengah	7	135000
20	6	2	2010 Bahan Beludru	Beludru	Bahan	Cap	Ibu Tyas	WANITA	Jawa Tengah	1	500000
21	11	4	2010 Hem Sutra	Sutra	Hem	Print	Ibu Tyas	WANITA	Jawa Tengah	5	100000
22											
23											

2. Pivot Table dengan grouping field (Sub Kategori, Tahun, Jumlah)

A	B	C	D	E	
1					
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

3. Pivot Table Sum of Jumlah2 diCount/di Jumlah kan (Total jumlah produk yang terjual, Jumlah transaksi yang terjadi)

Row Labels	Column Labels	2010	2011	2012	Total Sum of jumlah	
Row Labels	Sum of jumlah	Count of jumlah2	Sum of jumlah	Count of jumlah2	Sum of jumlah	Count of jumlah2
Bahan	1	1	8	1	8	2
Batik					1	1
Bolero			1	1		1
Celana	17	1			17	34
Hem	5	1	8	2	2	5
Jam					44	1
Jarik	2	1	4	1	6	2
Kaos	1	1	14	1	15	2
Rok				1	1	1
Sarimbit		1	1		1	1
Grand Total	23	3	21	7	93	10
					137	20

4. Calculated Field (Sum of Pendapatan)

	Column Labels			2010			2011			2012			Total Sum of jumlah			Total Count of jumlah2		Total Sum of Pendapatan	
	Sum of jumlah	Count of jumlah2	Sum of Pendapatan	Sum of jumlah	Count of jumlah2	Sum of Pendapatan	Sum of jumlah	Count of jumlah2	Sum of Pendapatan	Sum of jumlah	Count of jumlah2	Sum of Pendapatan	Total Sum of jumlah	Total Count of jumlah2	Total Sum of Pendapatan				
6 Bahan	1	1	500000	8	1	960000	8	2	2120000	17	4	15045000							
7 Batik			0			0	1	1	150000	1	1	150000							
8 Bolero			0	1	1	225000			0	1	1	225000							
9 Celana	17	1	935000			0	17	1	935000	34	2	3740000							
10 Hem	5	1	500000	8	2	4960000	4	2	1596000	17	5	19023000							
11 Jam			0			0	44	1	3520000	44	1	3520000							
12 Jarik			0	2	1	450000	4	1	160000	6	2	1590000							
13 Kaos			0	1	1	60000	14	1	420000	15	2	1350000							
14 Rok			0			0	1	1	225000	1	1	225000							
15 Sarimbit			0	1	1	150000			0	1	1	150000							
16 Grand Total	23	3	15065000	21	7	2940000	93	10	115692000	137	20	451963000							

5. Operasi Roll Up dan Down urutan (nama_kategori, nama_subkategori, dan nama_produk) dan (tahun)

	Sum of Pendapatan	Column Labels	2010	2011	2012	Grand Total	25
Row Labels							
5 Bahan	500000	960000	2120000	15045000	27	Hem Standar	0
6 Beludru	500000	0	0	500000	28	Sutra	500000
7 Bahan Beludru	500000	0	0	500000	29	Hem Sutra	500000
8 Lawasan	0	0	130000	130000	30	Jam	0
9 Bahan Lawasan	0	0	130000	130000	31	Standar	0
10 Standar	0	960000	945000	3825000	32	Jam Standar	0
11 Bahan Standar	0	960000	945000	3825000	33	Jarik	0
12 Batik	0	0	150000	150000	34	Standar	0
13 Standar	0	0	150000	150000	35	Jarik Standar	0
14 Batik Standar	0	0	150000	150000	36	Kaos	0
15 Bolero	0	225000	0	225000	37	Batik	0
16 Standar	0	225000	0	225000	38	Kaos Batik	0
17 Bolero Standar	0	225000	0	225000	39	Katun	0
18 Celana	935000	0	935000	3740000	40	Kaos Katun	0
19 Standar	935000	0	935000	3740000	41	Rok	0
20 Celana Standar	935000	0	935000	3740000	42	Batik	0
21 Hem	500000	4960000	1596000	19023000	43	Rok Batik	0
22 Katun	0	210000	897000	2214000	43	Sarimbit	0
23 Hem Katun	0	210000	897000	2214000	44	Standar	0
24 Standar	0	2750000	100000	3900000	45	Sarimbit Standar	0
						Grand Total	15065000
							29400000
							115692000
							451963000

Tugas

- a. PPN (Pajak Pertambahan Nilai) sebesar 10% dari tiap pendapatan pada Pivot Table.
- b. Total Penghasilan yang dihitung dari pendapatan dikurangi dengan PPN tersebut.

	Column Labels			2012	Total Sum of Pendapatan	Total Sum of PPN(10%)	Total Sum of Total Penghasilan
Row Labels	Sum of Pendapatan	Sum of PPN(10%)	Sum of Total Penghasilan				
6 Bahan	2120000	212000	1908000	2120000	212000	1908000	
7 Batik	150000	15000	135000	150000	15000	135000	
8 Celana	935000	93500	841500	935000	93500	841500	
9 Hem	1596000	159600	1436400	1596000	159600	1436400	
10 Jam	3520000	352000	3168000	3520000	352000	3168000	
11 Jarik	160000	16000	144000	160000	16000	144000	
12 Kaos	420000	42000	378000	420000	42000	378000	
13 Rok	225000	22500	202500	225000	22500	202500	
14 Grand Total	115692000	11569200	10412800	11569200	11569200	10412800	

2. Buatlah PivotTable untuk melihat PPN dan Total Penghasilan tersebut selama tahun 2010 – 2012.

Kategori produk apakah yang memberikan nilai penghasilan terbanyak selama 3 tahun tersebut? (Hem : memiliki total penghasilan sebesar 17120700).

3	2012			Total Sum of Pendapatan	Total Sum of PPN(10%)	Total Sum of Total Penghasilan
4	Sum of Pendapatan	Sum of PPN(10%)	Sum of Total Penghasilan			
6	212000	212000	1908000	1504500	1504500	13540500
7	150000	15000	135000	150000	15000	135000
8	0	0	0	225000	22500	202500
9	935000	93500	841500	3740000	374000	3366000
10	1596000	159600	1436400	19023000	1902300	17120700
11	3520000	352000	3168000	3520000	352000	3168000
12	160000	16000	144000	1590000	159000	1431000
13	420000	42000	378000	1350000	135000	1215000
14	225000	22500	202500	225000	22500	202500
15	0	0	0	150000	15000	135000
16	11569200	11569200	104122800	451963000	45196300	406766700

Modul 6 (Pengenalan Data Mining)

	A	B	C	D	E	F
1	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten	Lama_Studi
2	IPS	WANITA	SURAKARTA	18	TIDAK	TERLAMBAT
3	IPA	PRIA	SURAKARTA	19	YA	TEPAT
4	LAIN	PRIA	SURAKARTA	19	TIDAK	TERLAMBAT
5	IPA	PRIA	LUAR	17	TIDAK	TERLAMBAT
6	IPA	WANITA	SURAKARTA	17	TIDAK	TEPAT
7	IPA	WANITA	LUAR	18	YA	TEPAT
8	IPA	PRIA	SURAKARTA	18	TIDAK	TERLAMBAT
9	IPA	PRIA	SURAKARTA	19	TIDAK	TEPAT
10	IPS	PRIA	LUAR	18	TIDAK	TERLAMBAT
11	LAIN	WANITA	SURAKARTA	18	TIDAK	TEPAT
12	IPA	WANITA	SURAKARTA	19	TIDAK	TEPAT
13	IPS	PRIA	SURAKARTA	20	TIDAK	TEPAT
14	IPS	PRIA	SURAKARTA	19	TIDAK	TEPAT
15	IPA	PRIA	SURAKARTA	19	TIDAK	TEPAT
16	IPA	PRIA	LUAR	22	YA	TEPAT
17	LAIN	PRIA	SURAKARTA	16	TIDAK	TERLAMBAT
18	IPS	PRIA	LUAR	20	TIDAK	TEPAT
19	LAIN	PRIA	LUAR	23	YA	TEPAT
20	IPA	PRIA	SURAKARTA	21	YA	TEPAT
21	IPS	PRIA	SURAKARTA	19	TIDAK	TERLAMBAT

- a. **Jumlah Data IPA, IPS, dan LAIN** [=COUNTIF(A2:A21;A3)] | [=COUNTIF(A2:A21;A2)] | [=COUNTIF(A2:A21;A4)]
- b. **Jumlah Data Lama_Studi (Tepat dan Terlambat)** [=COUNTIF(F2:F21;F3)] | [=COUNTIF(F2:F21;F2)]
- c. **Pada data Rerata_SKS, mencari Max, Min, Mean, dan Standart Deviation**
=MAX(D2:D21) | =MIN(D2:D21) | =AVERAGE(D2:D21) | =STDEV.S(D2:D21)
- d. **Jumlah Data Gabungan pada atribut Jurusan_SMA = IPA, Gender = PRIA, ASISTEN=YA, Lama_Studi = TEPAT**
[=COUNTIFS(A2:A21;A3;B2:B21;B3:E2:E21;E3:F2:F21;F3)]

Jumlah Data IPA	Jumlah Data IPS	Jumlah Data Lain	
10	6	4	
Lama Studi(TEPAT)	Lama Studi(Terlambat)		
13	7		
Nilai Max	Nilai Min	Mean	Standar Deviasi
23	16	18,95	1,66938375
Jumlah Data Gabungan			
3			
(IPA,PRIA,ASISTEN(YA),LAMA_STUDI(TEPAT))			

Modul 7 (Data Preprocessing)

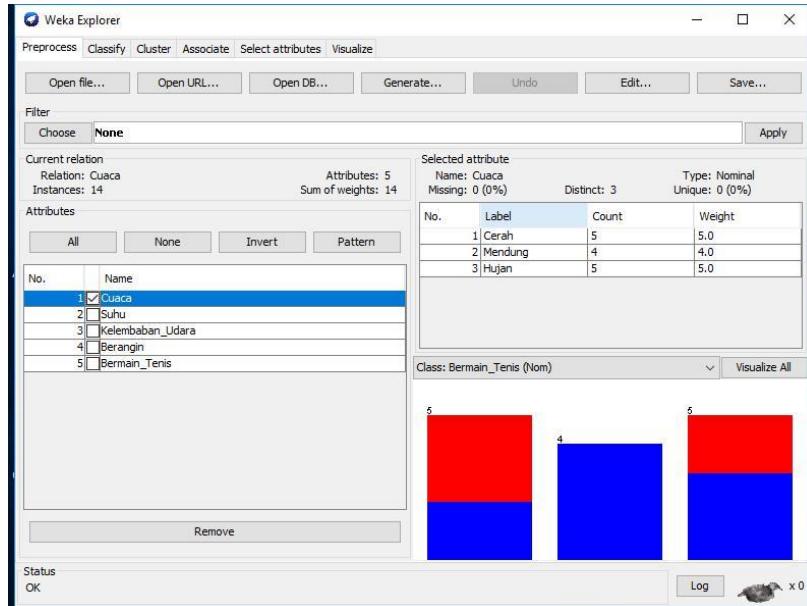
Soal Latihan Percobaan

```
Untitled - Notepad
File Edit Format View Help
@relation Cuaca

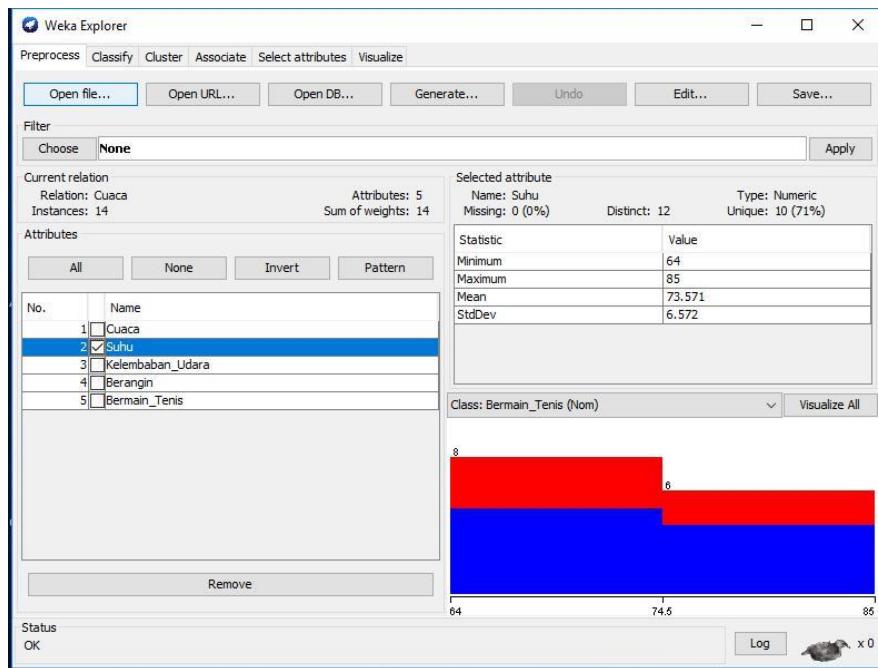
@attribute Cuaca {Cerah, Mendung, Hujan}
@attribute Suhu real
@attribute Kelembaban_Udara real
@attribute Berangin {YA, TIDAK}
@attribute Bermain_Tenis {YA, TIDAK}

@data
Cerah,85,85,TIDAK,TIDAK
Cerah,80,90,YA,TIDAK
Mendung,83,86,TIDAK,YA
Hujan,78,96,TIDAK,YA
Hujan,68,80,TIDAK,YA
Hujan,65,70,YA,TIDAK
Mendung,64,65,YA,YA
Cerah,72,95,TIDAK,TIDAK
Cerah,69,70,TIDAK,YA
Hujan,75,80,TIDAK,YA
Cerah,75,70,YA,YA
Mendung,72,90,YA,YA
Mendung,81,75,TIDAK,YA
Hujan,71,91,YA,TIDAK
```

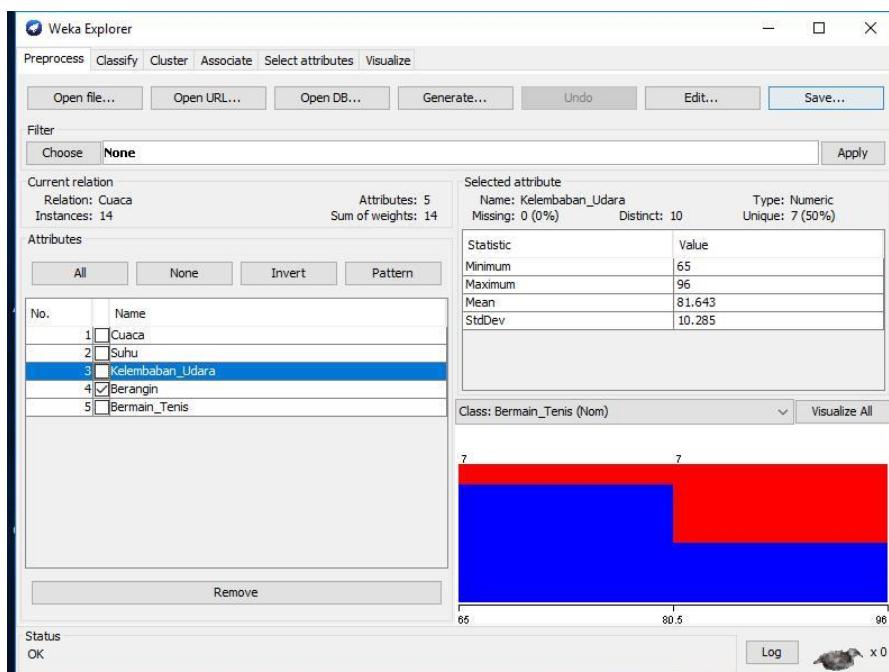
Grafik Cuaca



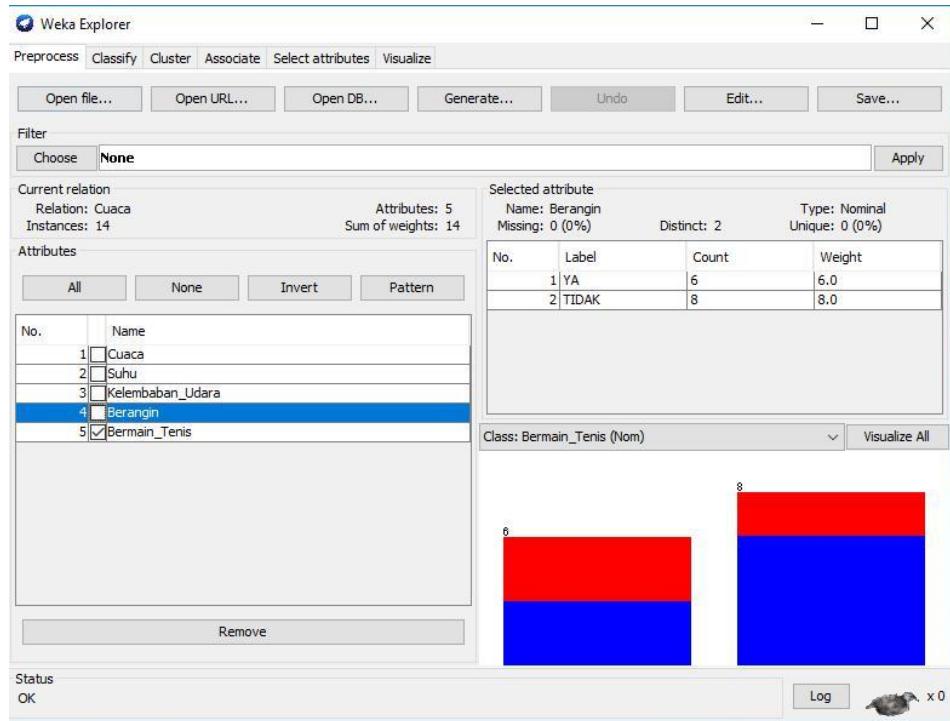
Grafik Suhu



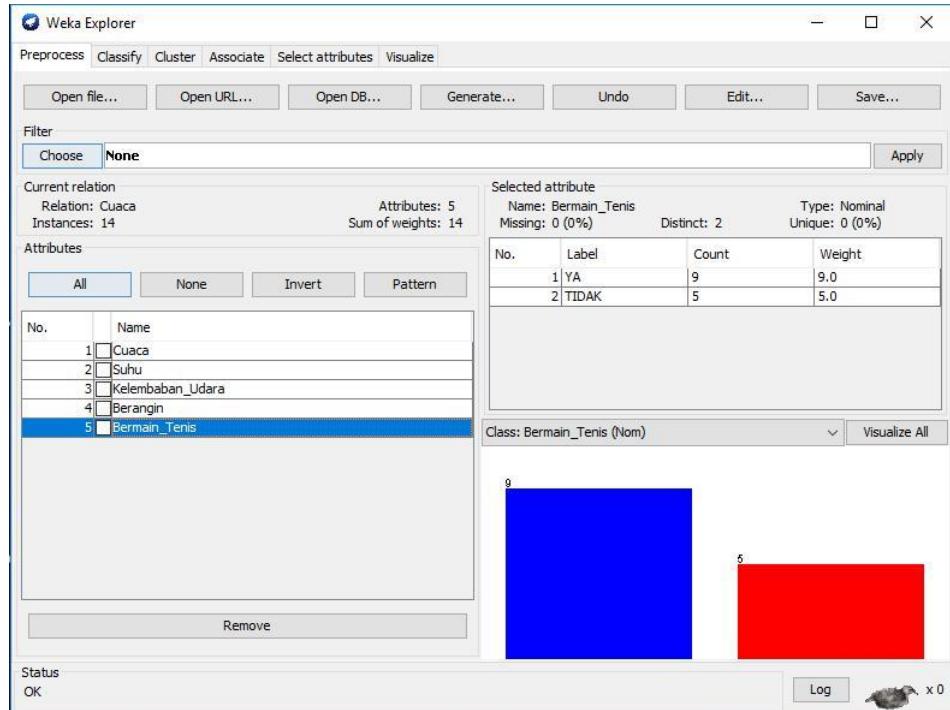
Grafik Kelembaban_Udara



Grafik Berangin



Grafik Bermain_Tenis



Tugas

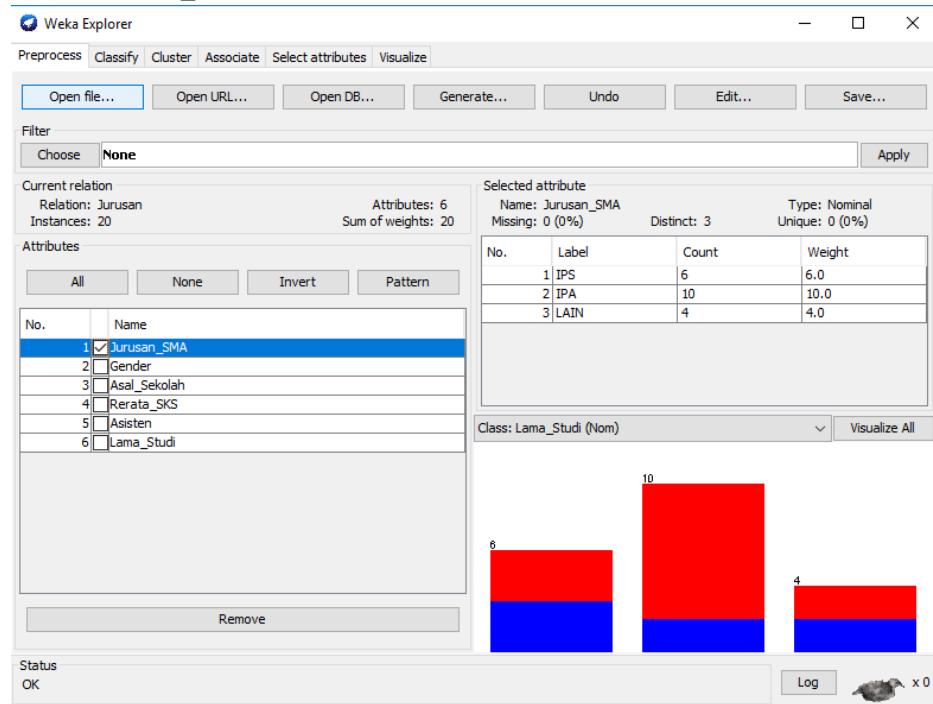
Soal Tugas

Jurusan - Notepad
File Edit Format View Help
@relation Jurusan

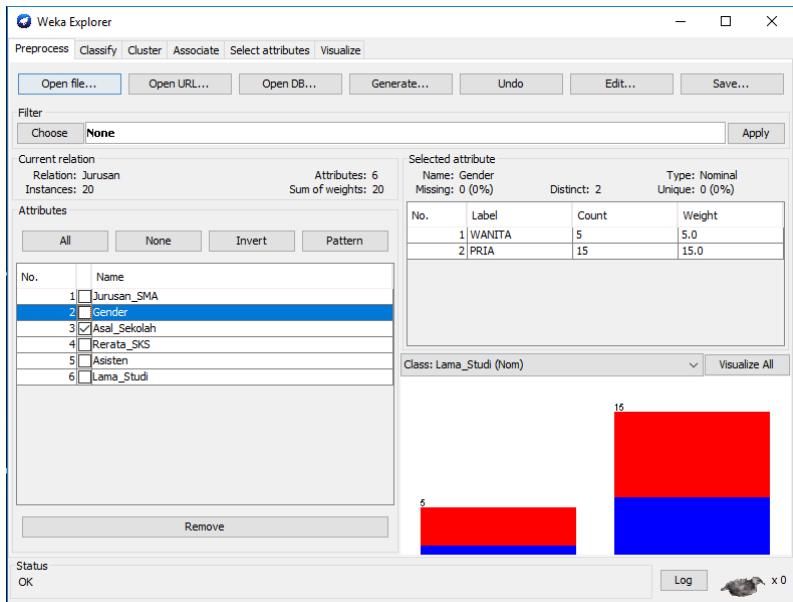
```
@attribute Jurusan_SMA {IPS, IPA, LAIN}
@attribute Gender {WANITA, PRIA}
@attribute Asal_Sekolah {SURAKARTA, LUAR}
@attribute Rerata_SKS real
@attribute Asisten {YA, TIDAK}
@attribute Lama_Studi {TERLAMBAT, TEPAT}

@data
IPS,WANITA,SURAKARTA,18,TIDAK,TERLAMBAT
IPA,PRIA,SURAKARTA,19,YA,TEPAT
LAIN,PRIA,SURAKARTA,19,TIDAK,TERLAMBAT
IPA,PRIA,LUAR,17,TIDAK,TERLAMBAT
IPA,WANITA,SURAKARTA,17,TIDAK,TEPAT
IPA,WANITA,LUAR,18,YA,TEPAT
IPA,PRIA,SURAKARTA,18,TIDAK,TERLAMBAT
IPA,PRIA,SURAKARTA,19,TIDAK,TEPAT
IPA,PRIA,LUAR,18,TIDAK,TERLAMBAT
LAIN,WANITA,SURAKARTA,18,TIDAK,TEPAT
IPA,WANITA,SURAKARTA,19,TIDAK,TEPAT
IPS,PRIA,SURAKARTA,20,TIDAK,TEPAT
IPS,PRIA,SURAKARTA,19,TIDAK,TEPAT
IPA,PRIA,SURAKARTA,19,TIDAK,TEPAT
IPA,PRIA,LUAR,22,YA,TEPAT
LAIN,PRIA,SURAKARTA,16,TIDAK,TERLAMBAT
IPS,PRIA,LUAR,20,TIDAK,TEPAT
LAIN,PRIA,LUAR,23,YA,TEPAT
IPA,PRIA,SURAKARTA,21,YA,TEPAT
IPS,PRIA,SURAKARTA,19,TIDAK,TERLAMBAT
```

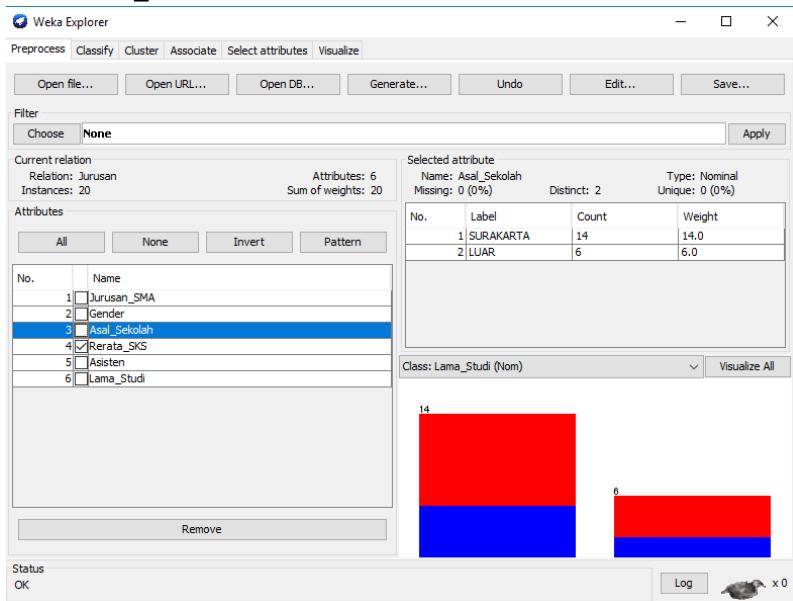
Grafik Jurusan_SMA



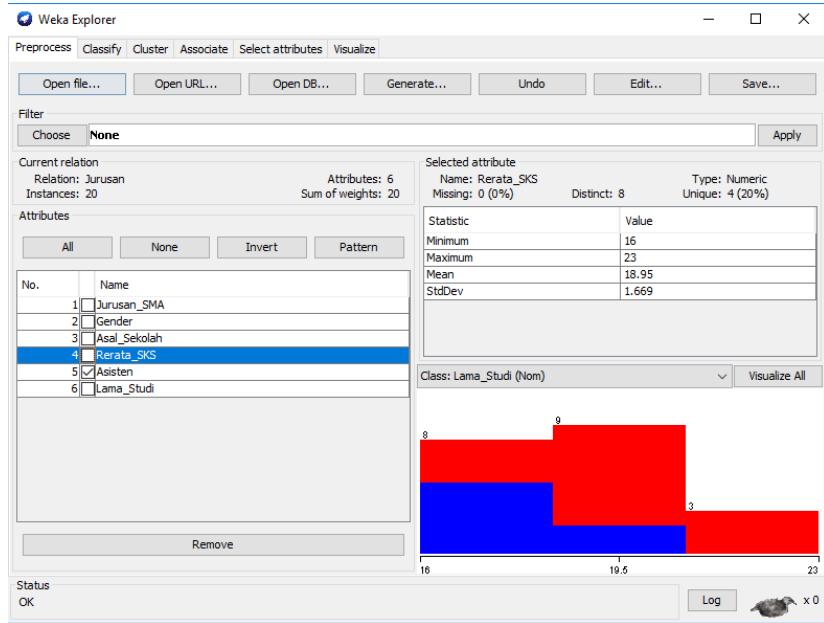
Grafik Gender



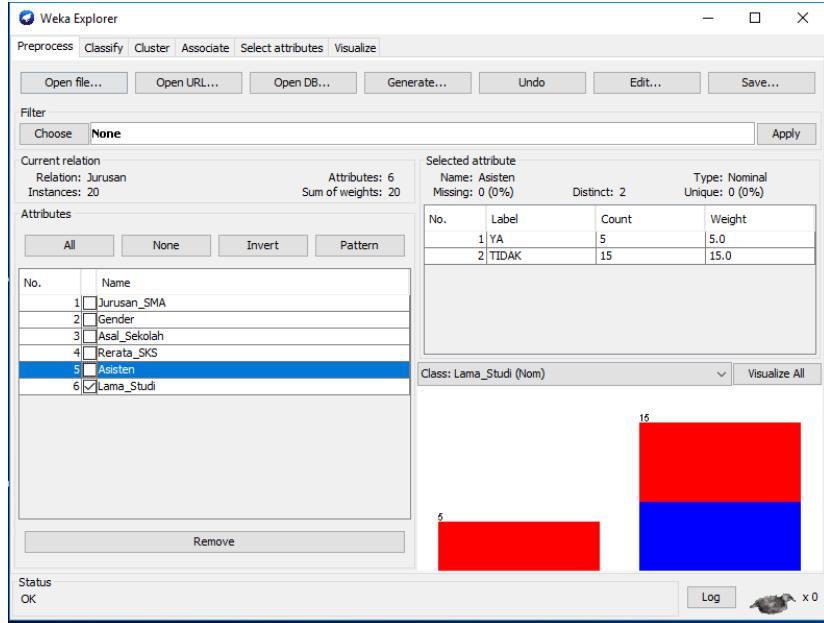
Grafik Asal_Sekolah



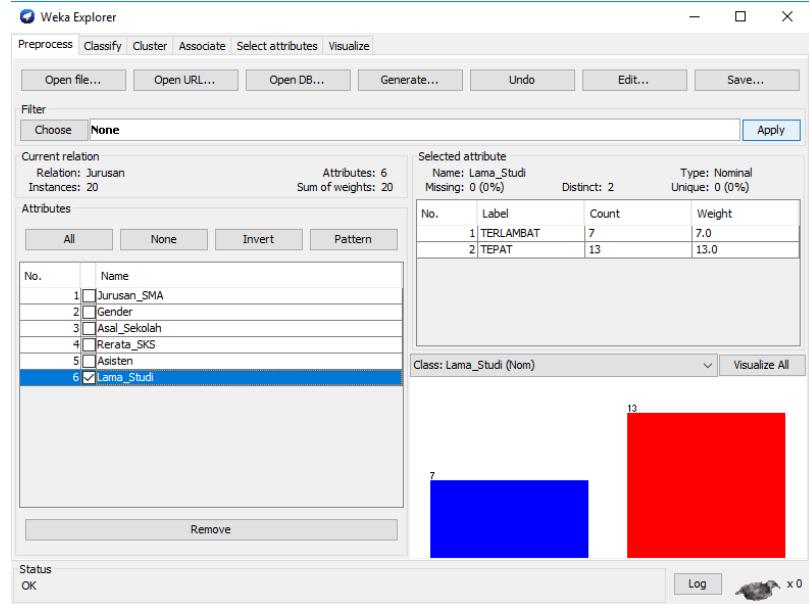
Grafik Rerata_SKS



Grafik Asisten



Grafik Lama_Studi



4. Binomial = Ada 4 (Gender, Asal_Sekolah, Asisten, Lama_Studi)

Polynomial = Ada 1 (Jurusan_SMA)

5. Real = Ada 1 (Rerata_SKS)

6. Rerata_SKS

Maximum : 23

Minimum : 16

Mean : 18.95

StdDev : 1.669

Modul 8 (Klasifikasi : Naïve Bayes)

Klasifikasi Naive Bayes dengan WEKA dan Rapid Miner

CuacaTesting - Notepad

```
File Edit Format View Help
@relation CuacaTesting

@attribute Cuaca {Cerah, Mendung, Hujan}
@attribute Suhu real
@attribute Kelembaban_udara real
@attribute Berangin {YA, TIDAK}
@attribute Bermain_Tenis {YA, TIDAK}

@data
Cerah,75,65,TIDAK,?
Cerah,80,68,YA,?
Cerah,83,87,YA,?
Mendung,70,96,TIDAK,?
Mendung,68,81,TIDAK,?
Hujan,65,75,YA,?
Hujan,64,85,YA,?
```

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Open file... Open URL... Open DB... Generate... Undo Edit... Save...

Filter Choose None Apply

Current relation
Relation: Cuaca Attributes: 5 Instances: 14 Sum of weights: 14

Selected attribute
Name: Cuaca Type: Nominal
Missing: 0 (0%) Distinct: 3 Unique: 0 (0%)

No.	Label	Count	Weight
1	Cerah	5	5.0
2	Mendung	4	4.0
3	Hujan	5	5.0

Attributes
All None Invert Pattern

No.	Name
1	Cuaca
2	Suhu
3	Kelembaban_Udara
4	Berangin
5	Bermain_Tenis

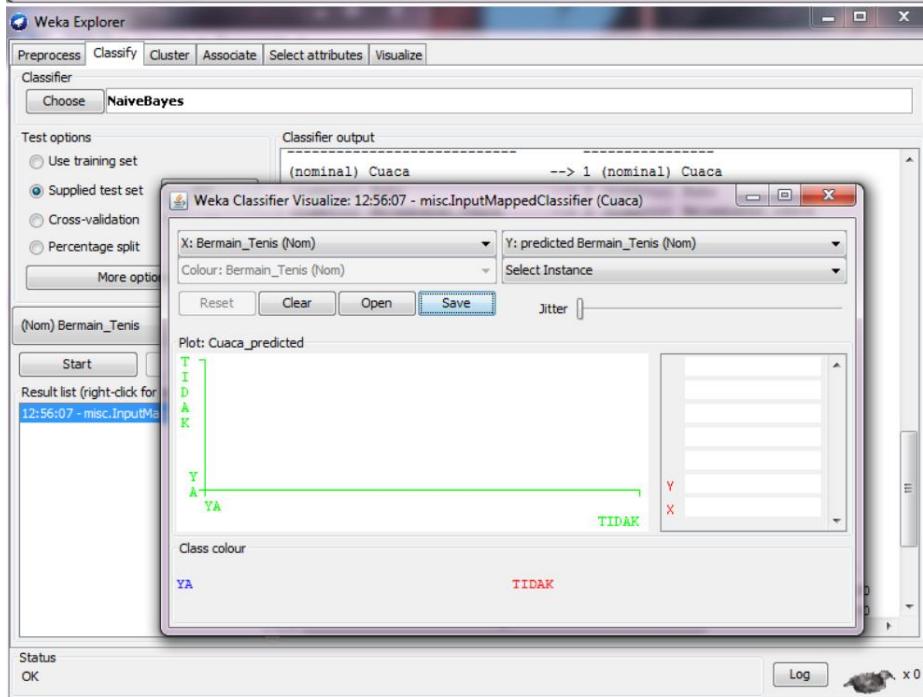
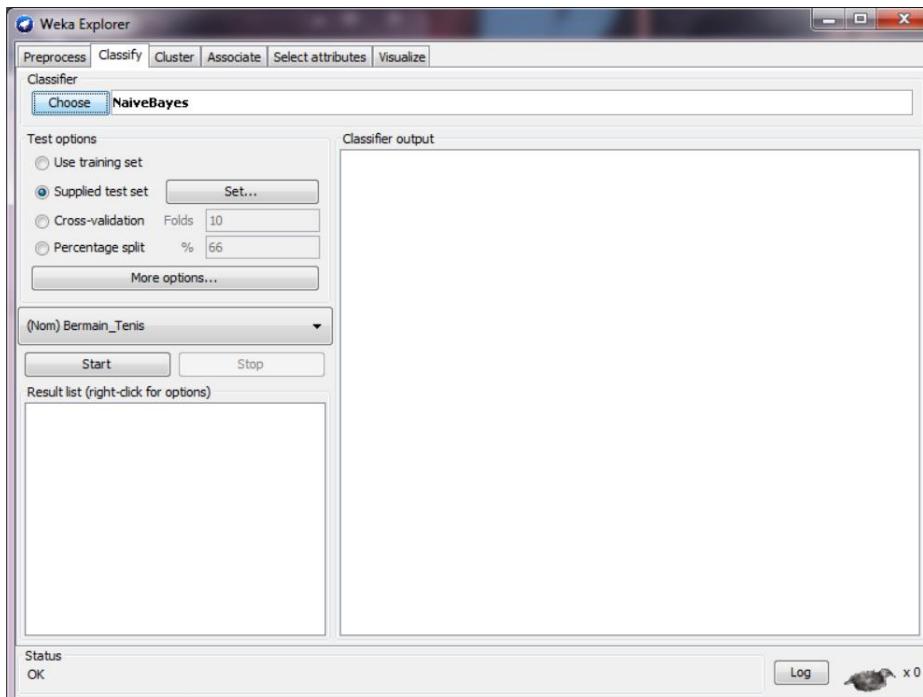
Remove

Class: Bermain_Tenis (Nom) Visualize All

5 4 5

5 4 5

Status OK Log x 0



ARFF-Viewer - C:\Users\LABSI-23\Documents\HasilPrediksi.arff

File Edit View
HasilPrediksi.arff

Relation: Cuaca_predicted

No.	1: Cuaca Nominal	2: Suhu Numeric	3: Kelembaban_Udara Numeric	4: Berangin Nominal	5: prediction margin Numeric	6: predicted Bermain_Tenis Nominal
1	Cerah	75.0	65.0	TIDAK	0.762765	YA
2	Cerah	80.0	68.0	YA	0.087878	YA
3	Cerah	83.0	87.0	YA	-0.676866	TIDAK
4	Mendung	70.0	96.0	TIDAK	0.628523	YA
5	Mendung	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

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	Hujan	74.000	94.000	YA	TIDAK

← Previous Next X Cancel

Import Data - Format your columns.

Format your columns.

Replace errors with missing values ⓘ

	Cuaca	Suhu	Kelembaban_udara	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

no problems. ← Previous Next X Cancel

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

Import Data - Select the cells to import. X

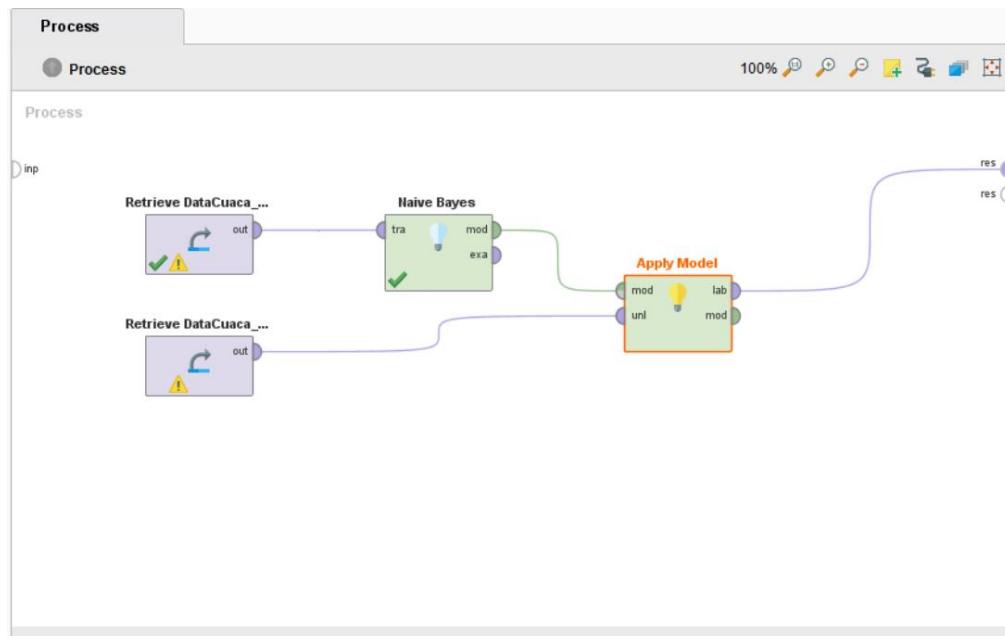
Select the cells to import.

Sheet: Testing ▼ Cell range: A1:D8 Select All Define header row: 1 ▾

A	B	C	D	E
1 Cuaca	Suhu	Kelembaban_udara	Berangin	
2 Cerah	85.000	85.000	TIDAK	
3 Cerah	80.000	90.000	YA	
4 Mendung	83.000	86.000	TIDAK	
5 Hujan	70.000	96.000	TIDAK	
6 Hujan	68.000	80.000	TIDAK	
7 Hujan	65.000	70.000	YA	
8 Mendung	64.000	65.000	YA	

← Previous → Next X Cancel

Row No.	Cuaca	Suhu	Kelembaban...	Berangin
1	Cerah	85	85	TIDAK
2	Cerah	80	90	YA
3	Mendung	83	86	TIDAK
4	Hujan	70	96	TIDAK
5	Hujan	68	80	TIDAK
6	Hujan	65	70	YA
7	Mendung	64	65	YA



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

Filter (7 / 7 examples): all

Row No.	prediction(Bermain_Tenis)	confidence(TIDAK)	confidence(YA)	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

Name

Type

Missing

Statistics

Filter (7 / 7 attributes): [Search for Attributes](#)

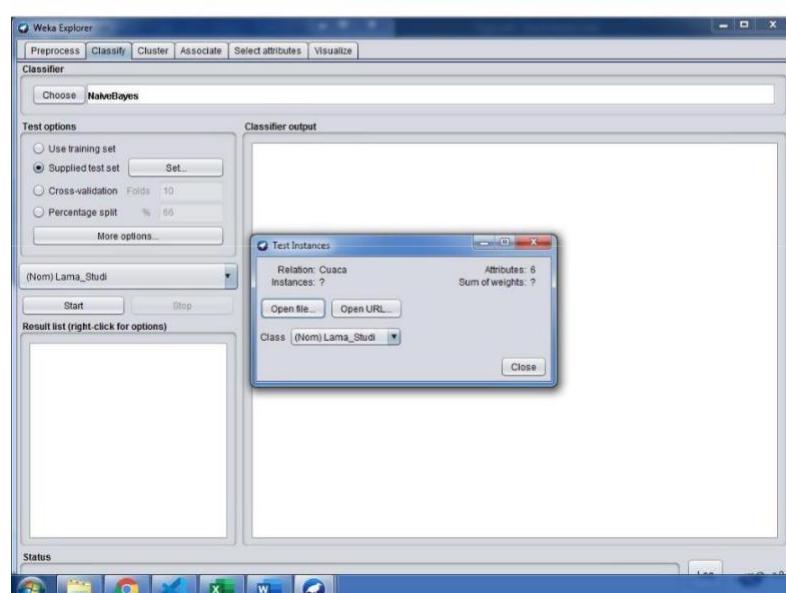
Name	Type	Missing	Statistics	Filter (7 / 7 attributes): Search for Attributes	▼
Prediction prediction(Bermain_Tenis)	Binominal	0	Least TIDAK (2) Most YA (5)	Values YA (5), TIDAK (2)	
Confidence_TIDAK confidence(TIDAK)	Real	0	Min 0.007 Max 0.856	Average 0.353	
Confidence_YA confidence(YA)	Real	0	Min 0.144 Max 0.993	Average 0.647	
Cuaca	Polynomial	0	Least Mendung (2) Most Cerah (3)	Values Cerah (3), Mendung (2)	
Suhu	Integer	0	Min 64 Max 83	Average 72.143	

Tugas

```

1  @relation Tugas
2
3  @attribute Jurusan_SMA{IPS, IPA, LAIN}
4  @attribute Gender{WANITA, PRIA}
5  @attribute Asal_Sekolah{LUAR, SURAKARTA}
6  @attribute Rerata_SKS real
7  @attribute Asisten{TIDAK, YA}
8  @attribute Lama_Studi(TEPAT, TERLAMBAT)
9
10 @data
11 IPS,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

```



a

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose **NaiveBayes**

Test options

- Use training set
- Supplied test set Set...
- Cross-validation Folds: 10
- Percentage split %: 66
- More options...

(Nom) **Lama_Studi**

Start Stop

Result list (right-click for options)

11.30.20 - misc.InputMappedClassifier

Classifier output

Naive Bayes Classifier

Attribute	Class	Value
	TEPAT TERLAMBAT	(0.64) (0.36)
Jurusan_SMA		
IPS	4.0	4.0
IPA	9.0	3.0
LAIN	3.0	3.0
[total]	16.0	10.0
Gender		
WANITA	5.0	2.0
PRIA	10.0	7.0
[total]	15.0	9.0
Asal_Sekolah		
LUAR	5.0	3.0
SURAKARTA	10.0	6.0
[total]	15.0	9.0
Rerata_SKS		
mean	19.5385	17.8571
std. dev.	1.5988	0.9897
weight sum	13	7
precision	1	1
Asisten		

Status

ARFF-Viewer - C:\Users\LABSI-20\Documents\TugasTesting.arff

File Edit View

TugasTesting.arff

Relation: Tugas

No. 1: Jurusan_SMA 2: Gender 3: Asal_Sekolah 4: Rerata_SKS 5: Asisten 6: Lama_Studi

	Nominal	Nominal	Nominal	Numerical	Nominal
1	LAIN	WANITA	SURAKARTA	18.0	TIDAK
2	IPA	PRIA	SURAKARTA	19.0	YA
3	LAIN	PRIA	SURAKARTA	19.0	TIDAK
4	IPS	PRIA	LUAR	17.0	TIDAK
5	LAIN	WANITA	SURAKARTA	17.0	TIDAK
6	IPA	WANITA	LUAR	18.0	YA
7	IPA	PRIA	SURAKARTA	18.0	TIDAK
8	IPA	PRIA	SURAKARTA	19.0	TIDAK
9	IPS	PRIA	LUAR	18.0	TIDAK
10	LAIN	WANITA	SURAKARTA	18.0	TIDAK

<new process> - RapidMiner Studio Trial 9.3.001 @ LABSI-20-PC

File Edit Process View Connections Settings Extensions Help

Import Data - Select the cells to import.

Select the cells to import.

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

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

Activate Wisdom of Crowds Previous Next Cancel

<new process> - RapidMiner Studio Trial 9.3.001 @ LABSI-20-PC

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model

Result History ExampleSet (/Local Repository/Tugas_Testing)

Open in Turbo Prep Auto Model

Filter (10 / 10 examples): all

Data Statistics Visualizations Annotations

Row No.	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_Sek...	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

L@ LABSI-20-PC

File Extensions Help

Import Data - Select the cells to import.

Select the cells to import.

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

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

← Previous → Next Cancel

Activate Wisdom of Crowds

L@ LABSI-20-PC

File Extensions Help

Import Data - Format your columns.

Format your columns.

Replace errors with missing values ⓘ

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

no problems.

← Previous → Next Cancel

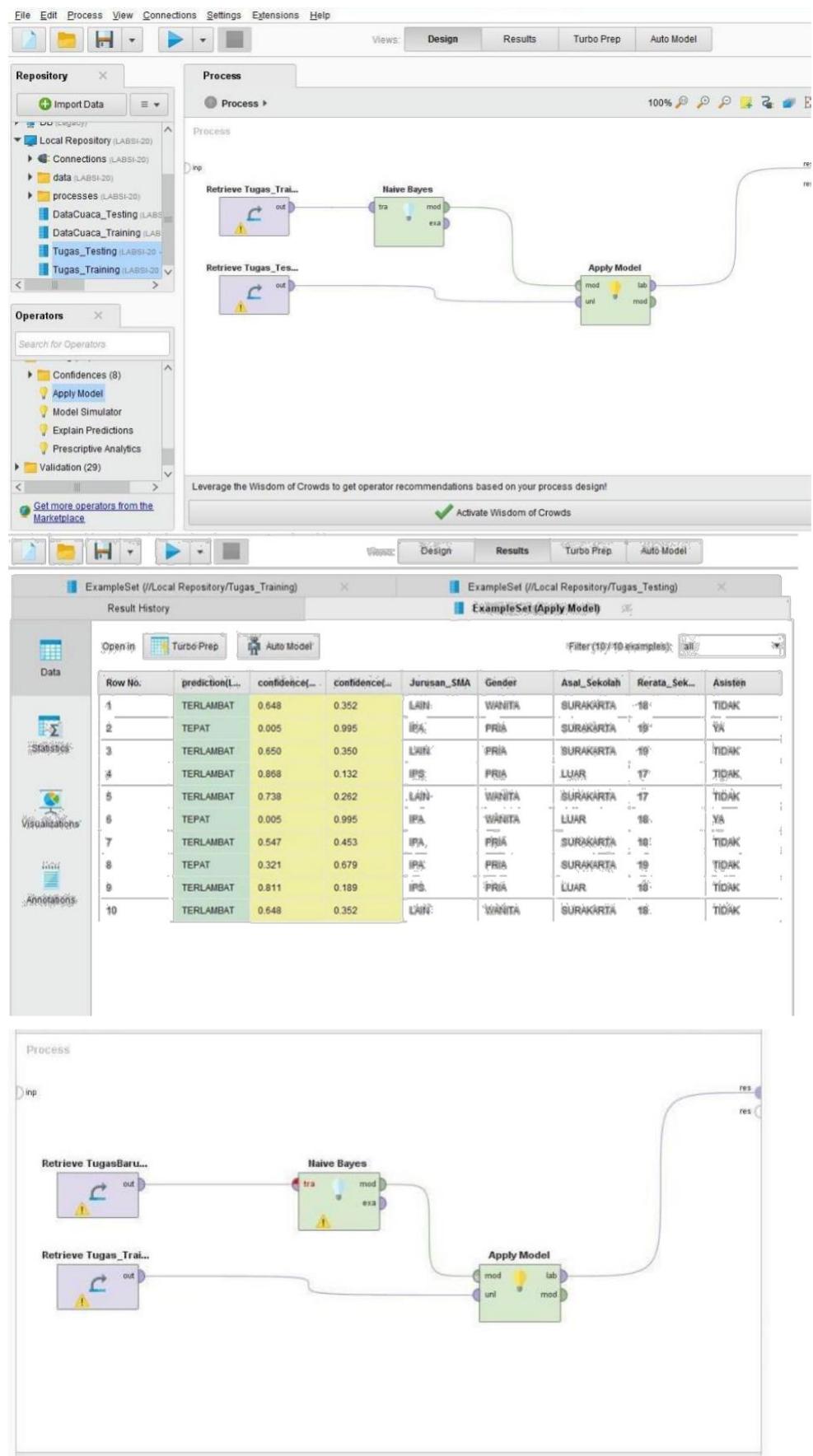
Result History ExampleSet (/Local Repository/Tugas_Training) ExampleSet (/Local Repository/Tugas_Testing)

Data Statistics Visualizations Annotations

Open in Turbo Prep Auto Model Filter (20 / 20 examples): all

Row No.	Lama_Studi	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_Sek...	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
15	TEPAT	IPA	PRIA	LUAR	22	YA

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



Open in

Turbo Prep

Auto Model

Filter (12 / 12 examples):

all



Row No.	prediction(L...	confidence(...	confidence(...	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_Sek...	Asisten
1	TERLAMBAT	0.648	0.352	LAII	WANITA	SURAKARTA	18	TIDAK
2	TEPAT	0.005	0.995	IPA	PRIA	SURAKARTA	19	YA
3	TERLAMBAT	0.650	0.350	LAII	PRIA	SURAKARTA	19	TIDAK
4	TERLAMBAT	0.868	0.132	IPS	PRIA	LUAR	17	TIDAK
5	TERLAMBAT	0.738	0.262	LAII	WANITA	SURAKARTA	17	TIDAK
6	TEPAT	0.005	0.995	IPA	WANITA	LUAR	18	YA
7	TERLAMBAT	0.547	0.453	IPA	PRIA	SURAKARTA	18	TIDAK
8	TEPAT	0.321	0.679	IPA	PRIA	SURAKARTA	19	TIDAK
9	TERLAMBAT	0.811	0.189	IPS	PRIA	LUAR	18	TIDAK
10	TERLAMBAT	0.648	0.352	LAII	WANITA	SURAKARTA	18	TIDAK
11	TEPAT	0.298	0.702	IPA	WANITA	LUAR	18	TIDAK
12	TEPAT	0.076	0.924	LAII	PRIA	SURAKARTA	17	YA

Dewi dan Jono sama sama Lulus dengan Tepat.

Modul 9 (Klasifikasi Decision Tree)

Pohon Keputusan Menggunakan WEKA dan RAPID MINER

Percobaan

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Open file... Open URL... Open DB... Generate... Undo Edit... Save...

Filter Choose None Apply

Current relation
Relation: Cuaca
Instances: 14 Attributes: 5 Sum of weights: 14

Attributes
All None Invert Pattern

No.	Name
1	Cuaca
2	Suhu
3	Kelembaban_Udara
4	Berangin
5	Bermain_Teknis

Selected attribute
Name: Cuaca
Missing: 0 (0%) Distinct: 3 Unique: 0 (0%) Type: Nominal

No.	Label	Count	Weight
1	Cerah	5	5.0
2	Mendung	4	4.0
3	Hujan	5	5.0

Class: Bermain_Teknis (Nom) Visualize All

Status OK Log x 0

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier Choose J48 -C 0.25 -M 2

Test options
 Use training set
 Supplied test set Set...
 Cross-validation Folds 10
 Percentage split % 66
More options...

(Nom) Bermain_Teknis Start Stop

Result list (right-click for options)
05:53:07 - trees.J48

Classifier output

```

| Berangin = YA: TIDAK (2.0)
| Berangin = TIDAK: YA (3.0)

Number of Leaves : 5

Size of the tree : 8

Time taken to build model: 0.01 seconds

*** Evaluation on training set ***

Time taken to test model on training data: 0 seconds

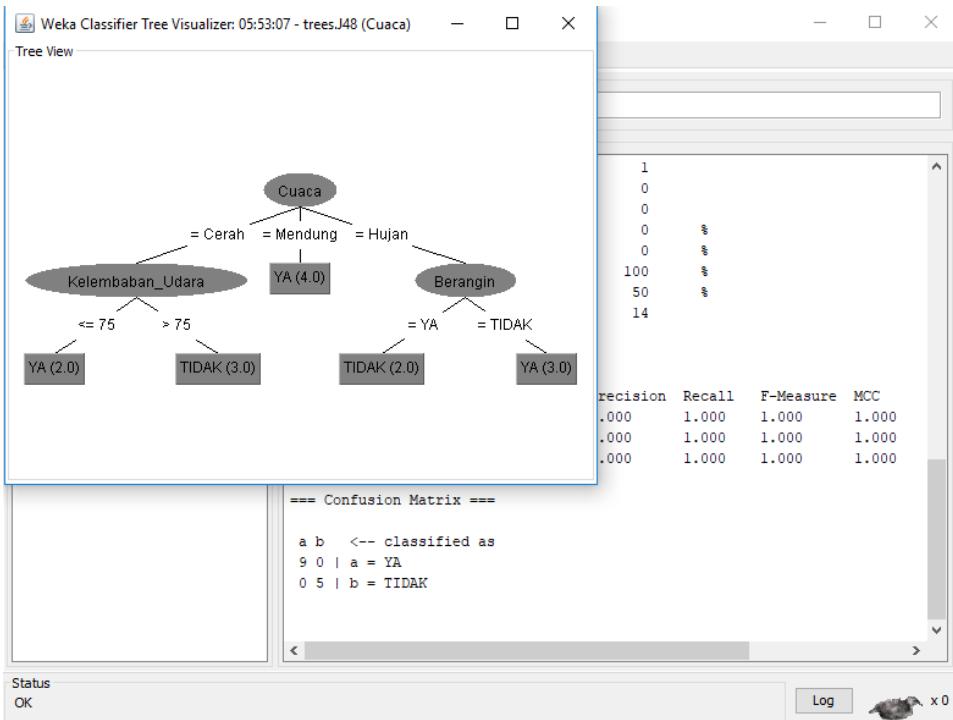
*** Summary ***

Correctly Classified Instances      14          100    %
Incorrectly Classified Instances   0           0      %
Kappa statistic                   1
Mean absolute error               0
Root mean squared error          0
Relative absolute error          0
Root relative squared error     0

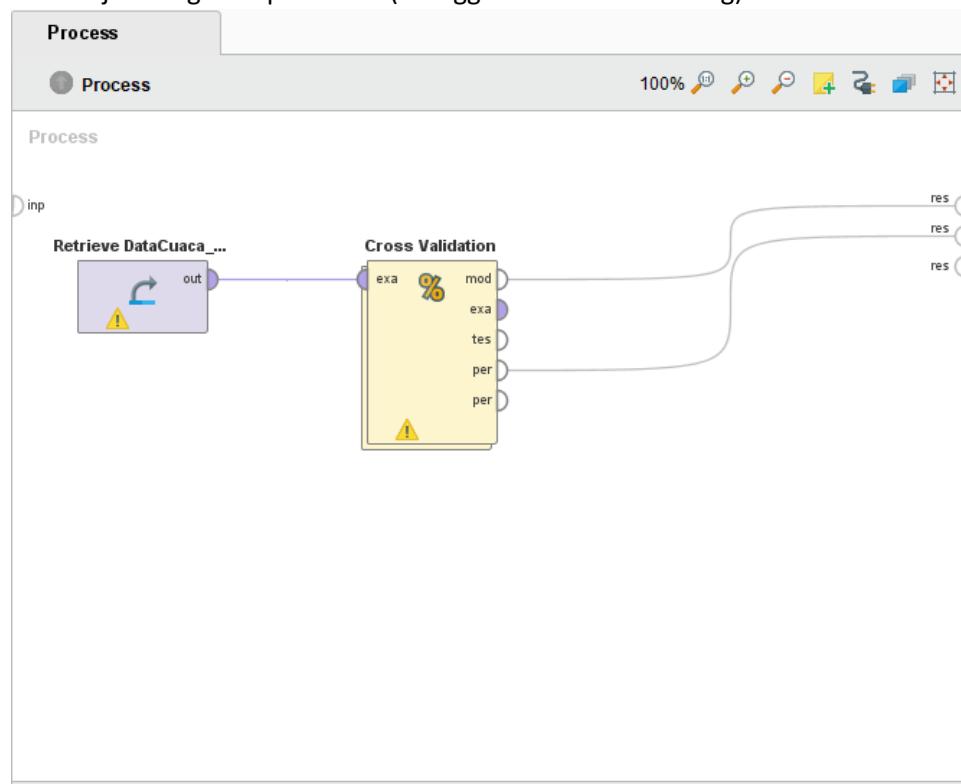
```

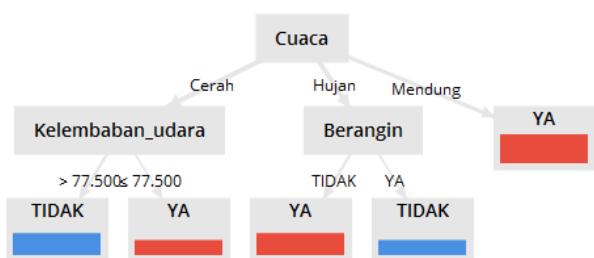
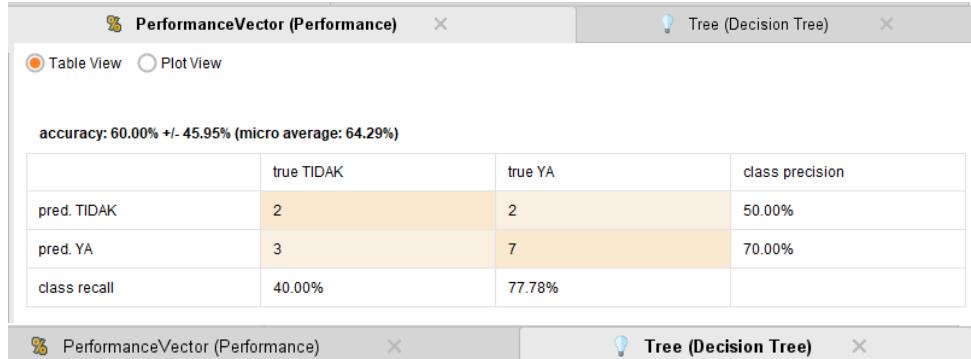
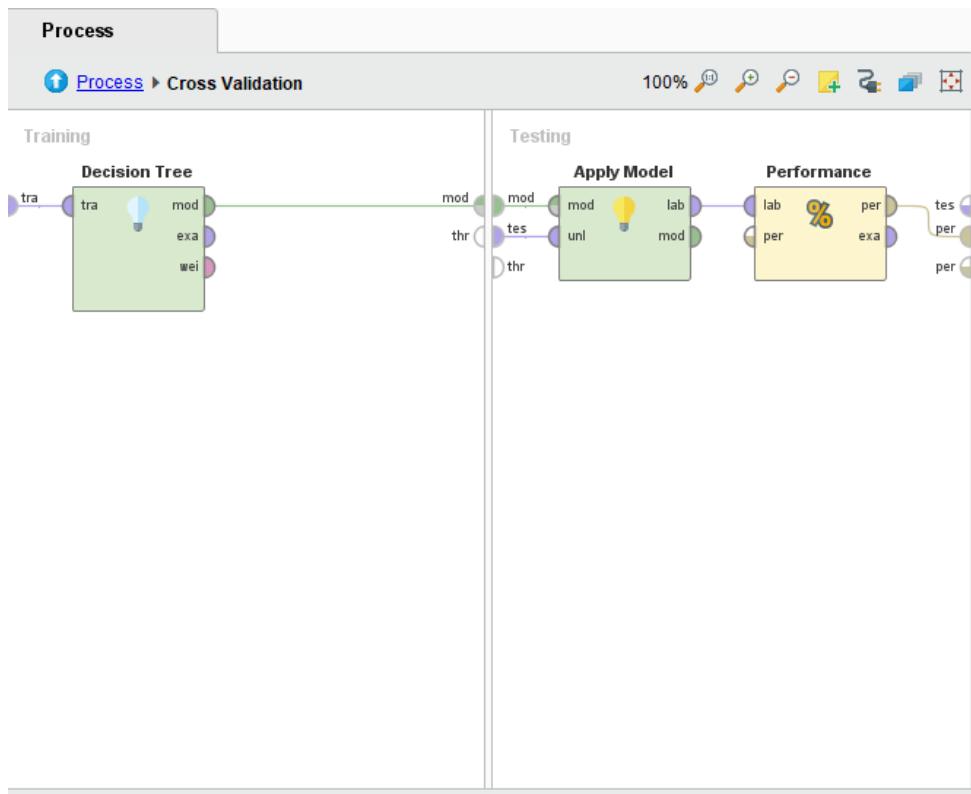
Status OK Log x 0

Choose(Lalu pilih Trees(J48) Pilih Use Training Set. Setelah itu Start)
(Lalu pada Result List Klik kanan dan pilih visualize tree)

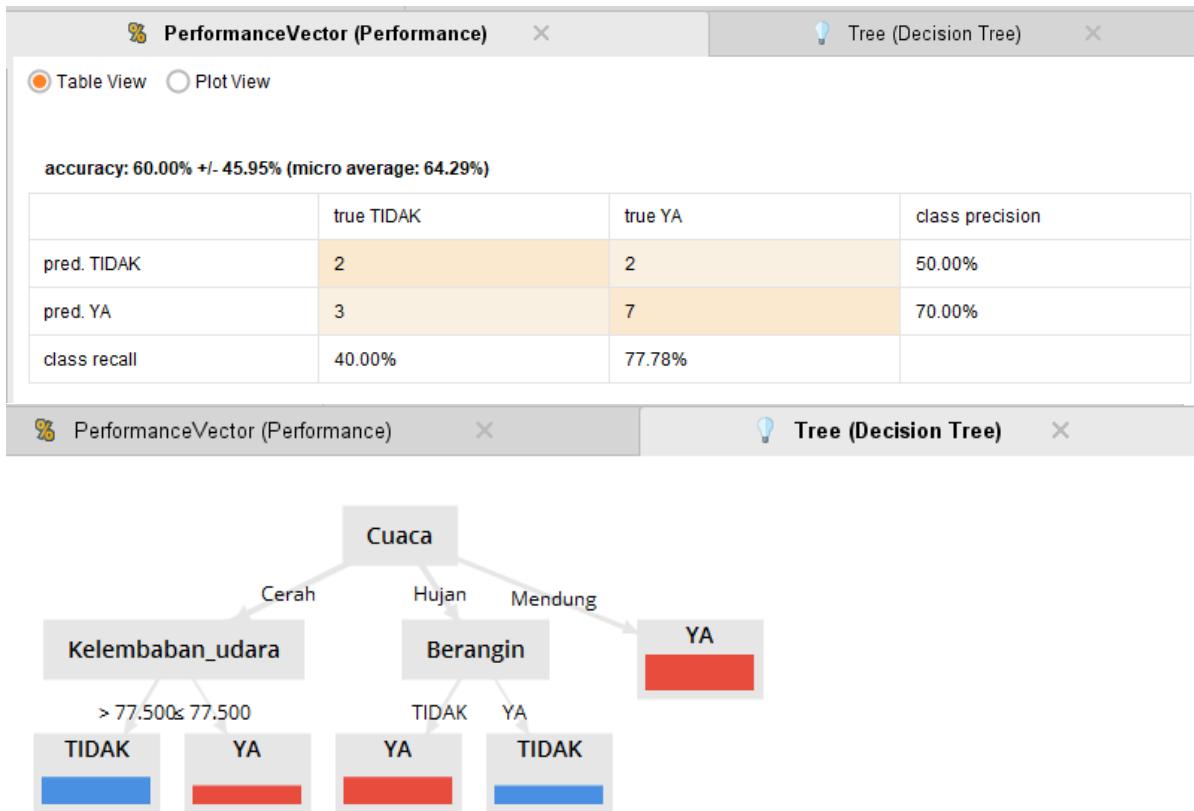


Lalu lanjut dengan Rapid Miner (Menggunakan Data Training)





Tugas



1.

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

2.

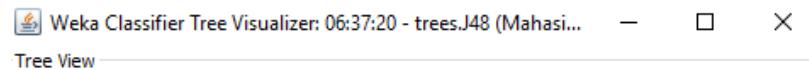
Number of Leaves : 3

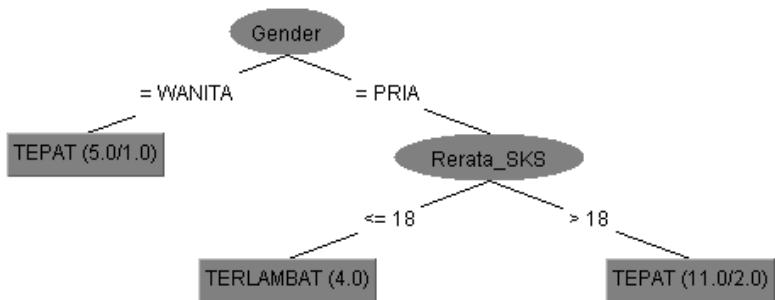
Size of the tree : 5

Time taken to test model on training data: 0 seconds

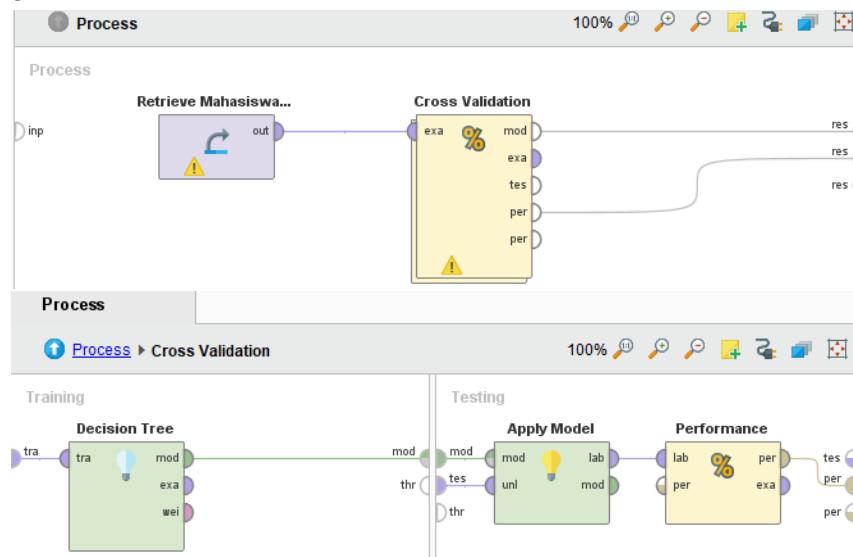
Correctly Classified Instances 17 85 %

Incorrectly Classified Instances 3 15 %

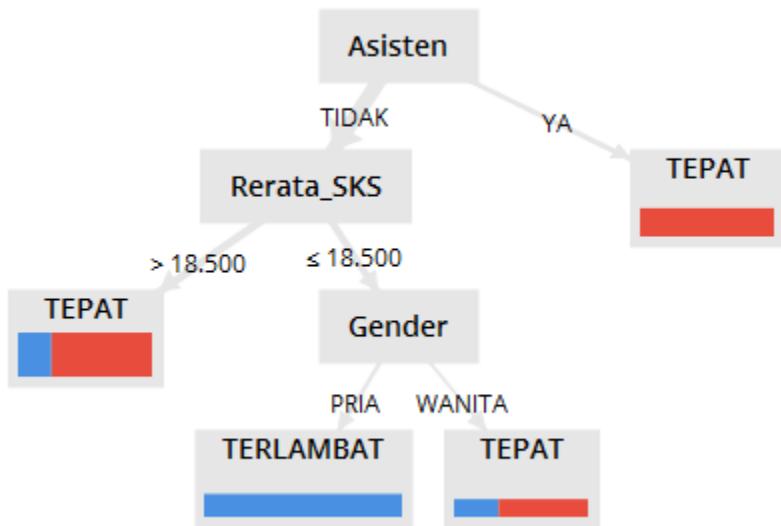
 Weka Classifier Tree Visualizer: 06:37:20 - trees.J48 (Mahasi... — X
Tree View



3.

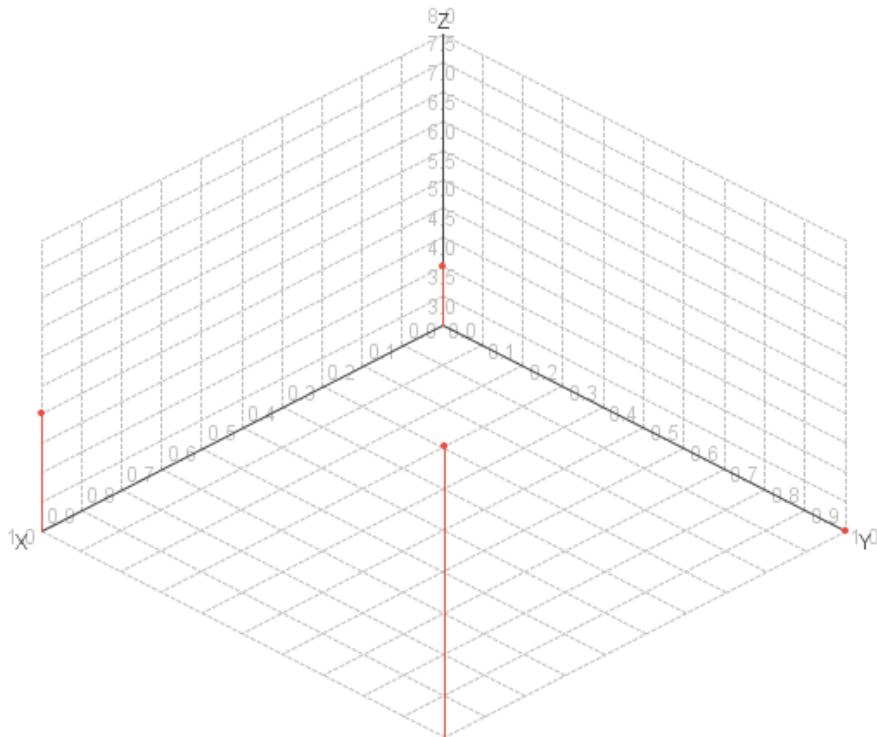


a.



b.

Confusion Matrix (x: true class, y: pred. class, z: counters)



4.

Klasifikasi yang terbentuk yaitu :

- a. Seseorang Lama Studi nya akan **Tepat** Jika Kondisi sebagai berikut :
 - i. Asisten = (YA) TEPAT
 - ii. Asisten = (TIDAK), Rerata_SKS (> 18.500) = **Tepat** (Nilai Atribut lain diabaikan).
 - iii. Asisten = (TIDAK), Rerata_SKS (≤ 18.500), Gender Wanita = **Tepat**
- b. Seseorang Lama Studi nya akan **Terlambat** Jika Kondisi sebagai berikut :
 - i. Asisten = (TIDAK), Rerata_SKS (≤ 18.500), Gender Pria = **Terlambat**

Modul 10 (Clustering K-Means)

Algoritma K-Means menggunakan Rapid Miner

Percobaan

Import Data - Select the cells to import.

Select the cells to import.

Sheet: Sheet1 Cell range: A:D Select All Define header row: 1

A	B	C	D
1 NO_SISWA	NAMA	B.IND	BJNG
2 S-101	JOKO	8.54	8.4
3 S-102	AGUS	9.98	6.81
4 S-103	SUSI	6.2	9.15
5 S-104	DYAH	5.24	7.26
6 S-105	WATI	5.7	5.71
7 S-106	IKA	8.57	5.87
8 S-107	EKO	7.7	7.71
9 S-108	YANTO	6.6	5.7
10 S-109	WAWAN	9.000	8.12
11 S-110	MAHMUD	9.81	9.58

← Previous → Next ✖ Cancel

Import Data - Format your columns.

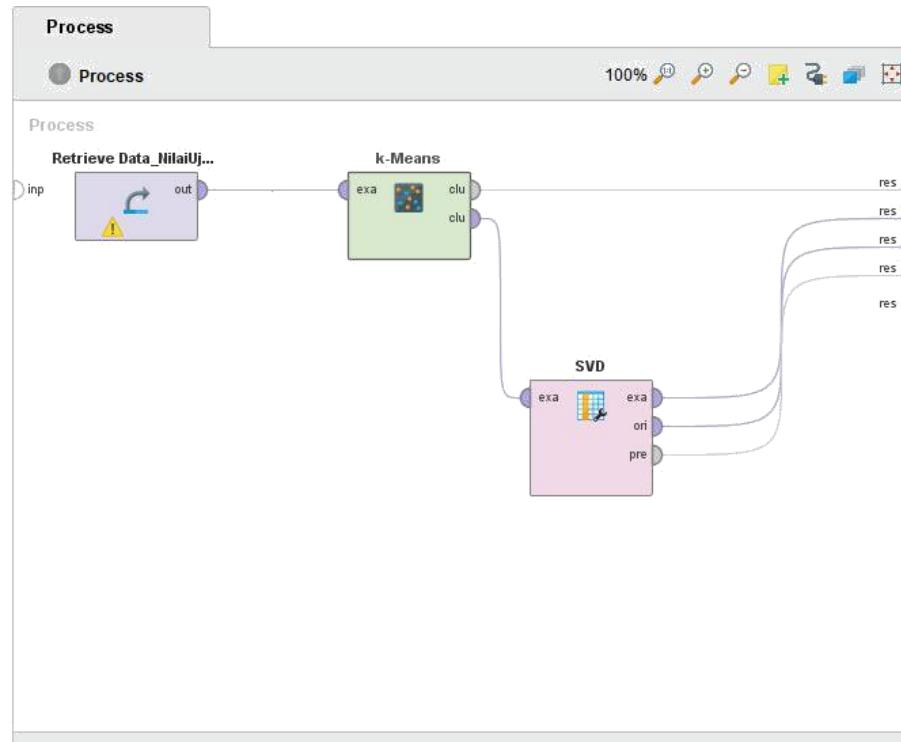
Format your columns.

Replace errors with missing values ⓘ

NO_SISWA polynomial	NAMA polynomial id	B.IND real	BJNG real
1 S-101	JOKO	8.540	8.400
2 S-102	AGUS	9.980	6.810
3 S-103	SUSI	6.200	9.150
4 S-104	DYAH	5.240	7.260
5 S-105	WATI	5.700	5.710
6 S-106	IKA	8.570	5.870
7 S-107	EKO	7.700	7.710
8 S-108	YANTO	6.600	5.700
9 S-109	WAWAN	9.000	8.120
10 S-110	MAHMUD	9.810	9.580

no problems. ← Previous → Next ✖ Cancel

Row No.	NAMA	NO_SISWA	B.IND	BJNG
1	JOKO	S-101	8.540	8.400
2	AGUS	S-102	9.980	6.810
3	SUSI	S-103	6.200	9.150
4	DYAH	S-104	5.240	7.260
5	WATI	S-105	5.700	5.710
6	IKA	S-106	8.570	5.870
7	EKO	S-107	7.700	7.710
8	YANTO	S-108	6.600	5.700
9	WAWAN	S-109	9	8.120
10	MAHMUD	S-110	9.810	9.580



eigenvalue

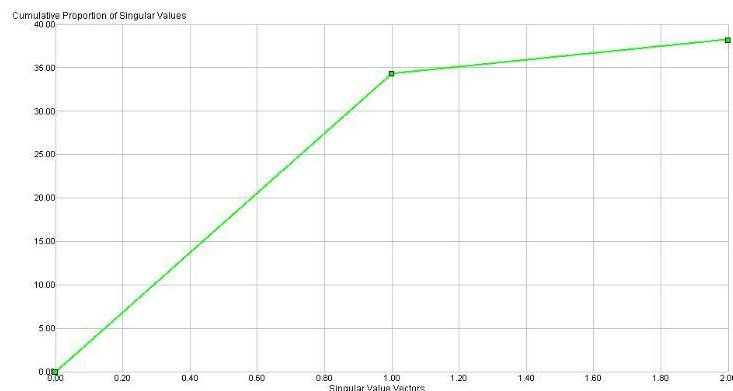
try SVD (SVD) ExampleSet (k-Means) ExampleSet (SVD)

Component	Singular Value	Proportion of Singular Values	Cumulative Singular Values	Cumulative Proportion of Sin...
SVD 1	34.340	0.898	34.340	0.898
SVD 2	3.906	0.102	38.246	1.000

Nilai svd vectors

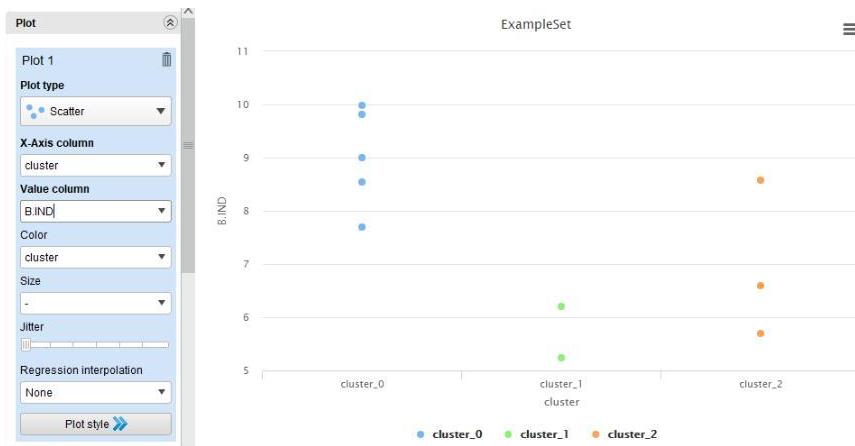
Attribute	SVD Vector 1
B.IND	0.723
B.ING	0.690

Cumulative variance

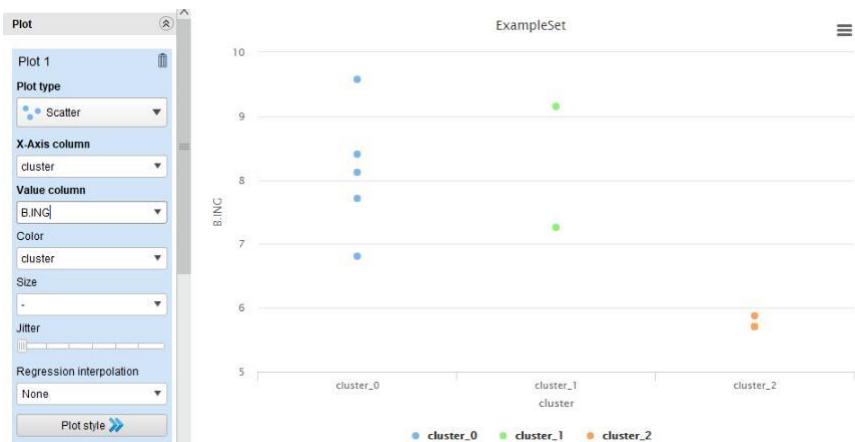


7. exampleset (k-means)

i. b.indonesia



ii. b.ing

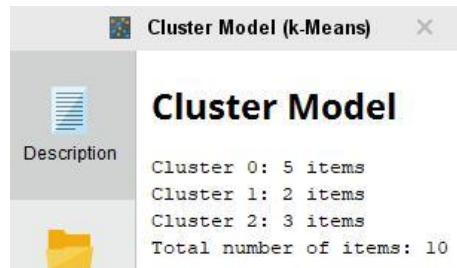


c) exampleset(svd)

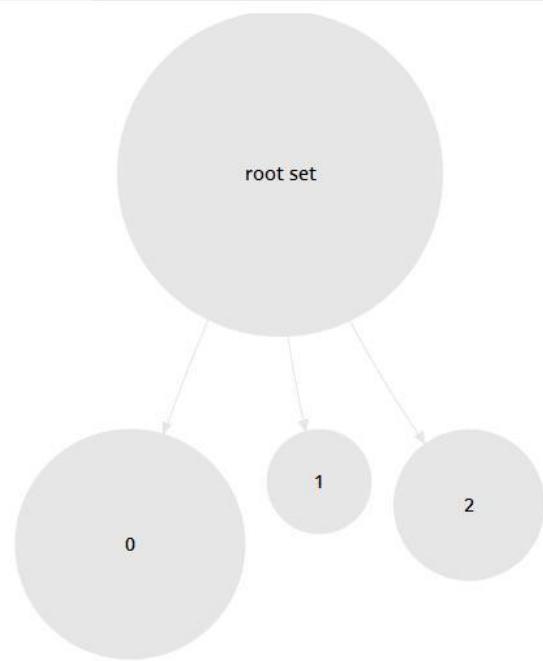
Row No.	NAMA	cluster ↑	B.IND	B.ING
1	JOKO	cluster_0	8.540	8.400
3	SUSI	cluster_0	6.200	9.150
7	EKO	cluster_0	7.700	7.710
9	WAWAN	cluster_0	9	8.120
10	MAHMUD	cluster_0	9.810	9.580
4	DYAH	cluster_1	5.240	7.260
5	WATI	cluster_1	5.700	5.710
8	YANTO	cluster_1	6.600	5.700
2	AGUS	cluster_2	9.980	6.810
6	IKA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

d) cluster model (clustering)

description



Graph

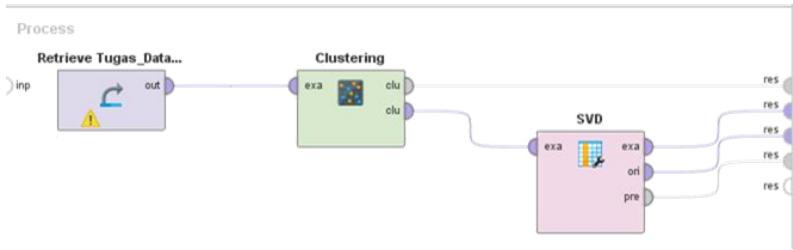


Row No.	NAMA	cluster ↑	svd_1
1	JOKO	cluster_0	0.349
3	SUSI	cluster_0	0.315
7	EKO	cluster_0	0.317
9	WAWAN	cluster_0	0.353
10	MAHMUD	cluster_0	0.399
4	DYAH	cluster_1	0.256
5	WATI	cluster_1	0.235
8	YANTO	cluster_1	0.254
2	AGUS	cluster_2	0.347
6	IKA	cluster_2	0.299

Tugas

Row No.	NAMA	B.IND	B.ING	MTK	IPA
1	JOKO	7.962	9.624	5.978	5.978
2	AGUS	6.093	7.263	7.268	8.443
3	SUSI	8.105	9.480	9.403	6.037
4	DYAH	8.307	5.827	5.619	7.139
5	WATI	5.242	9.690	6.507	7.029
6	IKA	8.645	6.136	8.349	5.233
7	EKO	7.862	6.209	5.829	8.754
8	YANTO	6.806	9.836	9.475	7.900
9	WAWAN	9.516	5.171	6.895	8.175
10	MAHMUD	5.185	6.818	6.958	6.116
11	BUDI	7.923	5.625	5.710	7.049
12	SANTI	7.931	9.095	7.761	6.790
13	DIAN	6.921	8.975	8.740	5.278
14	DANI	6.784	6.596	7.547	9.640

Row No.	NAMA	B.IND	B.ING	MTK	IPA
15	AHMAD	9.889	9.561	7.453	7.389
16	BAYU	7.298	5.063	5.202	5.976
17	RISA	6.698	7.986	7.454	9.470
18	RANI	5.835	5.559	9.387	7.980
19	YANI	7.188	5.362	8.349	6.873
20	RATIH	9.800	5.680	8.395	5.609
21	INDAH	8.043	9.351	7.014	5.467
22	JONO	7.170	5.573	7.422	9.424
23	SARAH	7.599	7.108	7.748	6.024
24	RAMA	6.257	6.897	8.332	5.917
25	BAMBANG	9.919	9.390	9.327	9.341
26	HADI	7.411	8.839	5.292	5.099
27	NANA	7.584	7.260	7.731	9.953
28	FEBRI	6.780	5.638	5.410	7.134
29	DENI	7.298	9.283	9.515	8.475
30	TONI	8.914	8.214	5.483	5.623



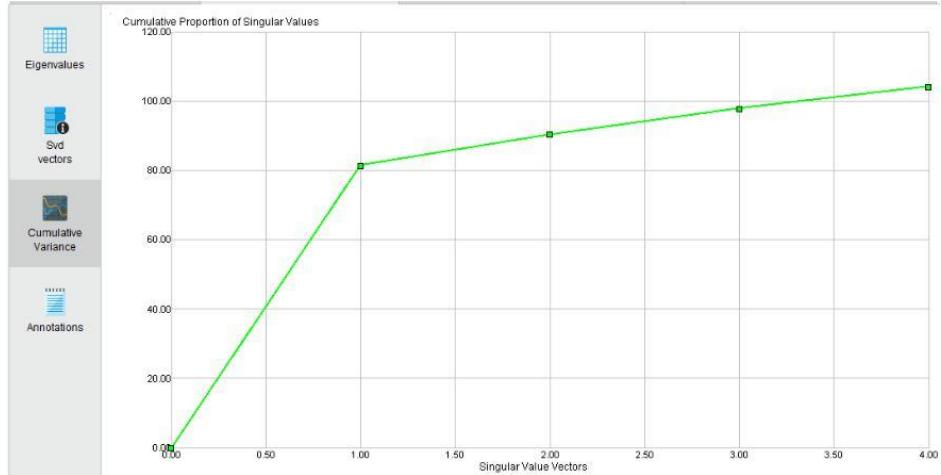
EigenValues

	Component	Singular Value	Proportion of Singular Values	Cumulative Singular Values	Cumulative Proportion of Sin...
Eigenvalues	SVD 1	81.422	0.782	81.422	0.782
Svd vectors	SVD 2	9.013	0.087	90.436	0.868
Eigenvalues	SVD 3	7.466	0.072	97.902	0.940
Svd vectors	SVD 4	6.246	0.060	104.148	1.000

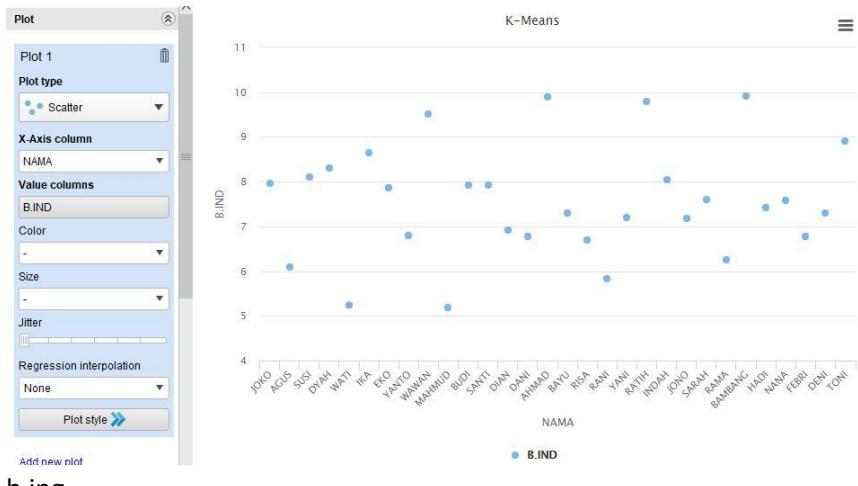
Svd Vectors

	Attribute	SVD Vector 1	SVD Vector 2	SVD Vector 3
Eigenvalues	B.IND	0.510	-0.024	-0.860
Svd vectors	B.ING	0.505	0.719	0.270
Eigenvalues	MTK	0.500	-0.029	0.316
Svd vectors	IPA	0.485	-0.694	0.296

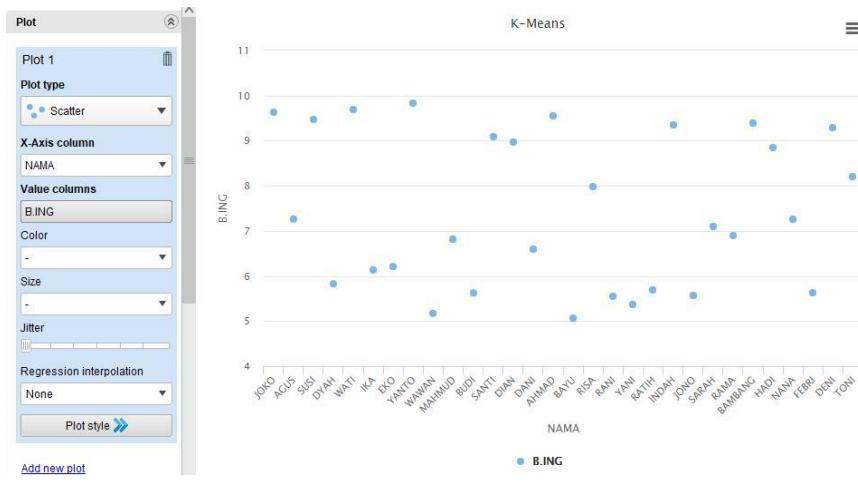
Cumulative Variance



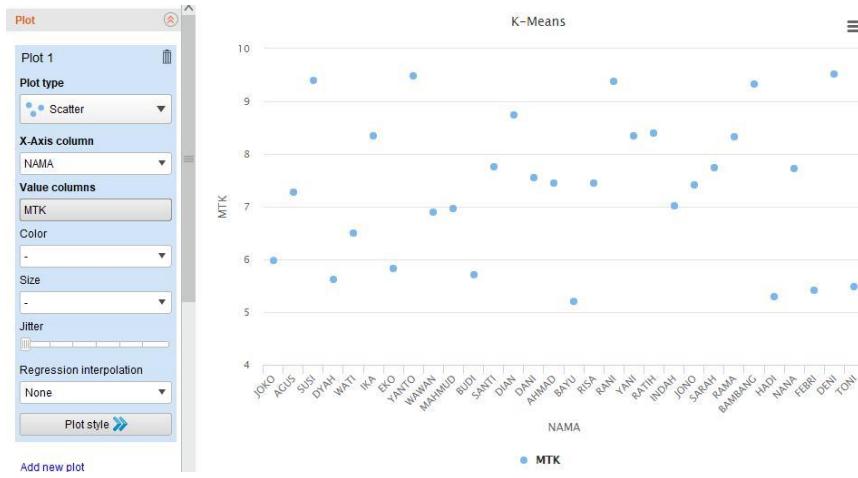
b.ind



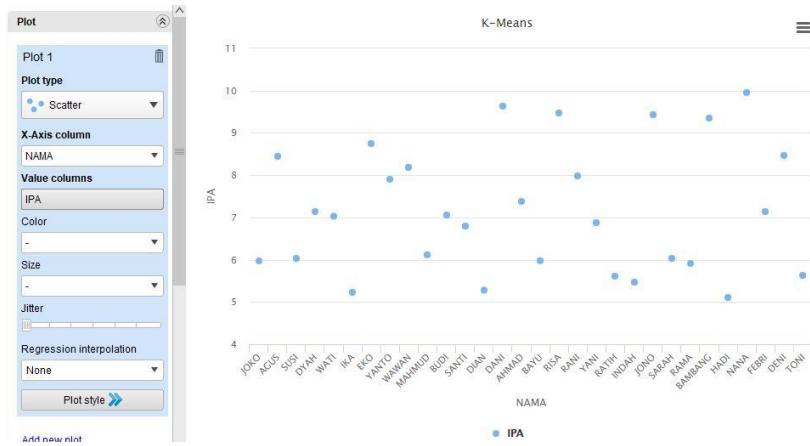
b.ing



Mtk



IPA



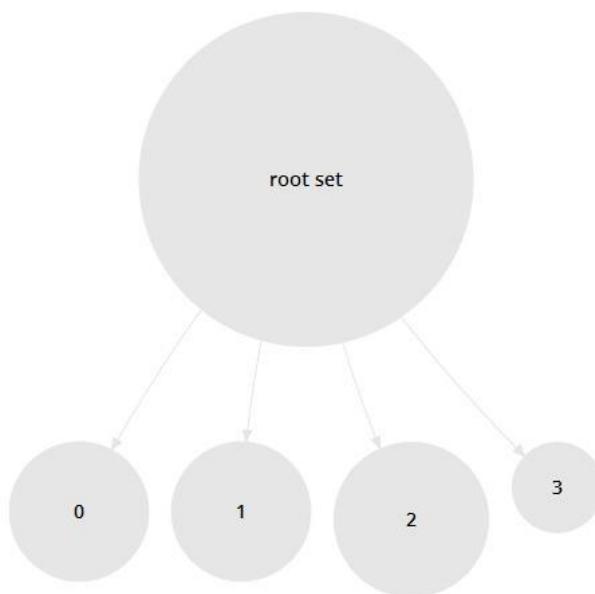
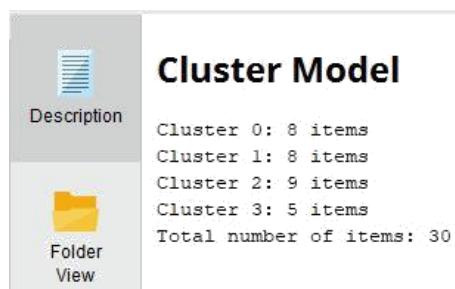
Exampleset svd

Row No.	NAMA	cluster ↑	svd_1
1	JOKO	cluster_0	0.182
5	WATI	cluster_0	0.175
12	SANTI	cluster_0	0.194
13	DIAN	cluster_0	0.184
21	INDAH	cluster_0	0.184
23	SARAH	cluster_0	0.175
26	HADI	cluster_0	0.164
30	TONI	cluster_0	0.174
4	DYAH	cluster_1	0.165
6	IKYA	cluster_1	0.175
7	EKO	cluster_1	0.176
9	WAWAN	cluster_1	0.183
11	BUDI	cluster_1	0.162
16	BAYU	cluster_1	0.145
20	RATIH	cluster_1	0.182
28	FEBRI	cluster_1	0.153
2	AGUS	cluster_2	0.178
10	MAHMUD	cluster_2	0.154
14	DANI	cluster_2	0.187
17	RISA	cluster_2	0.194
18	RANI	cluster_2	0.176
19	YANI	cluster_2	0.170
22	JONO	cluster_2	0.181
24	RAMA	cluster_2	0.168
27	NANA	cluster_2	0.199
3	SUSI	cluster_3	0.203
8	YANTO	cluster_3	0.209
15	AHMAD	cluster_3	0.211
25	BAMBANG	cluster_3	0.233
29	DENI	cluster_3	0.212

Cluster model

Description

Graph



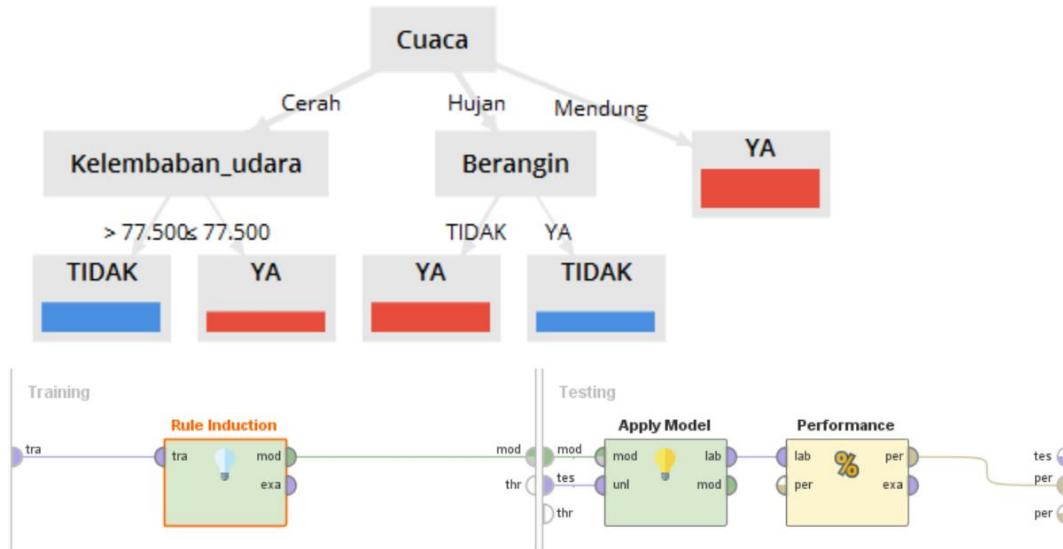
TABLE

Row No.	NAMA	cluster ↑	B.IND	B.ING	MTK	IPA
1	JOKO	cluster_0	7.962	9.624	5.978	5.978
5	WATI	cluster_0	5.242	9.690	6.507	7.029
12	SANTI	cluster_0	7.931	9.095	7.761	6.790
13	DIAN	cluster_0	6.921	8.975	8.740	5.278
21	INDAH	cluster_0	8.043	9.351	7.014	5.467
23	SARAH	cluster_0	7.599	7.108	7.748	6.024
26	HADI	cluster_0	7.411	8.839	5.292	5.099
30	TONI	cluster_0	8.914	8.214	5.483	5.623
4	DYAH	cluster_1	8.307	5.827	5.619	7.139
6	IKA	cluster_1	8.645	6.136	8.349	5.233
7	EKO	cluster_1	7.862	6.209	5.829	8.754
9	WAWAN	cluster_1	9.516	5.171	6.895	8.175
11	BUDI	cluster_1	7.923	5.625	5.710	7.049
16	BAYU	cluster_1	7.298	5.063	5.202	5.976
Row No.	NAMA	cluster ↑	B.IND	B.ING	MTK	IPA
20	RATIH	cluster_1	9.800	5.680	8.395	5.609
28	FEBRI	cluster_1	6.780	5.638	5.410	7.134
2	AGUS	cluster_2	6.093	7.263	7.268	8.443
10	MAHMUD	cluster_2	5.185	6.818	6.958	6.116
14	DANI	cluster_2	6.784	6.596	7.547	9.640
17	RISA	cluster_2	6.698	7.986	7.454	9.470
18	RANI	cluster_2	5.835	5.559	9.387	7.980
19	YANI	cluster_2	7.188	5.362	8.349	6.873
22	JONO	cluster_2	7.170	5.573	7.422	9.424
24	RAMA	cluster_2	6.257	6.897	8.332	5.917
27	NANA	cluster_2	7.584	7.260	7.731	9.953
3	SUSI	cluster_3	8.105	9.480	9.403	6.037
8	YANTO	cluster_3	6.806	9.836	9.475	7.900
15	AHMAD	cluster_3	9.889	9.561	7.453	7.389
25	BAMBANG	cluster_3	9.919	9.390	9.327	9.341
29	DENI	cluster_3	7.298	9.283	9.515	8.475

Modul 11 (Induksi dan Aturan Asosiasi)

Induksi dan Aturan Asosiasi menggunakan data cuaca dan aplikasi rapid miner.

	true TIDAK	true YA	class precision
pred. TIDAK	2	2	50.00%
pred. YA	3	7	70.00%
class recall	40.00%	77.78%	



RuleModel

```

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

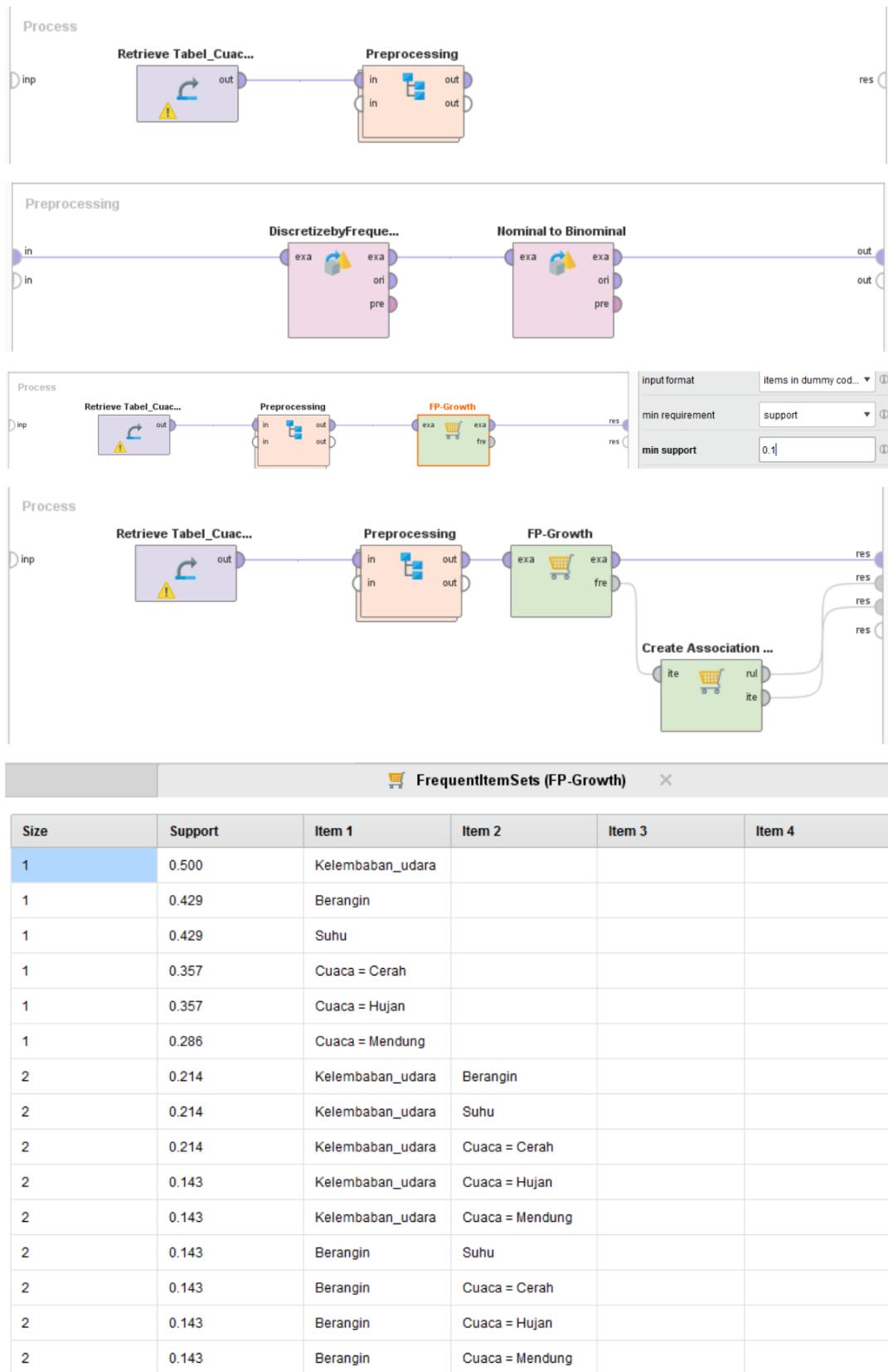
```

correct: 12 out of 13 training examples.

accuracy: 65.00% +/- 45.00% (micro average: 71.43%)

	true TIDAK	true YA	class precision
pred. TIDAK	2	1	66.67%
pred. YA	3	8	72.73%
class recall	40.00%	88.89%	

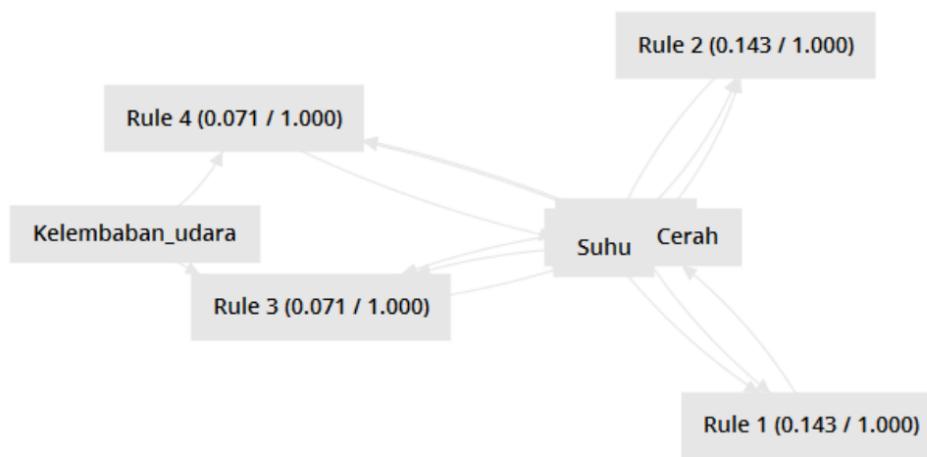
Aturan Asosiasi Data Cuaca

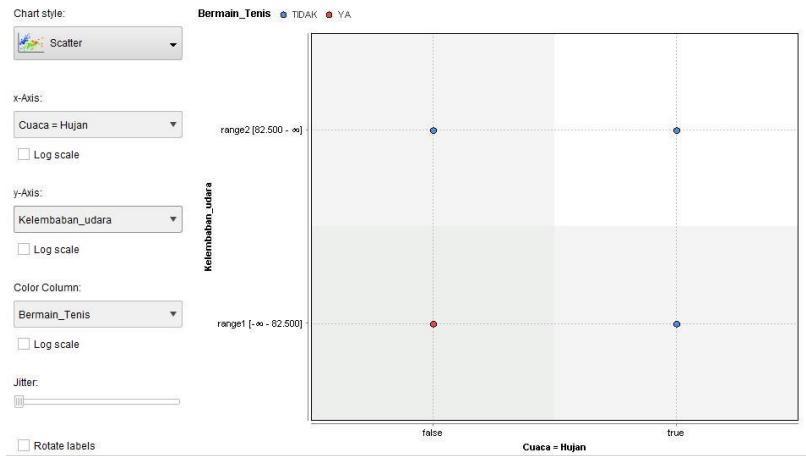


2	0.214	Suhu	Cuaca = Cerah			
2	0.071	Suhu	Cuaca = Hujan			
2	0.143	Suhu	Cuaca = Mendung			
3	0.071	Kelembaban_udara	Berangin	Suhu		
3	0.071	Kelembaban_udara	Berangin	Cuaca = Cerah		
3	0.071	Kelembaban_udara	Berangin	Cuaca = Hujan		
3	0.071	Kelembaban_udara	Berangin	Cuaca = Mendung		
3	0.143	Kelembaban_udara	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	

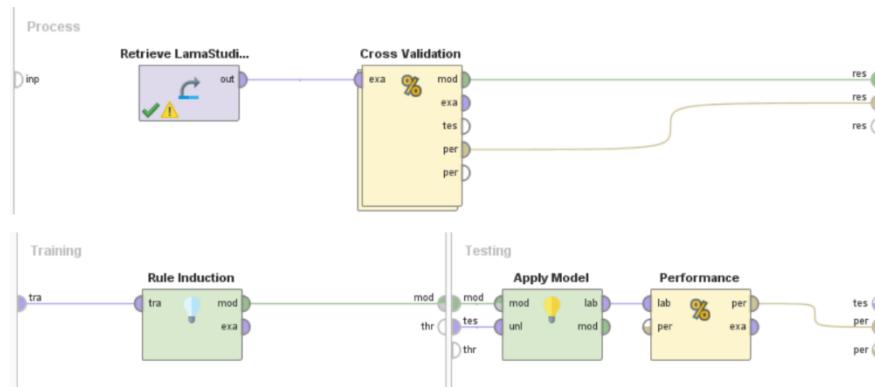
No.	Premises	Conclusion	Support
1	Berangin, Suhu	Cuaca = Cerah	0.143
2	Berangin, Cuaca = Cerah	Suhu	0.143
3	Kelembaban_udara, Berangin, Suhu	Cuaca = Cerah	0.071
4	Kelembaban_udara, Berangin, Cuaca = Cerah	Suhu	0.071

Confidence	LaPlace	Gain	p-s	Lift	Convicti...
1	1	-0.143	0.092	2.800	∞
1	1	-0.143	0.082	2.333	∞
1	1	-0.071	0.046	2.800	∞
1	1	-0.071	0.041	2.333	∞



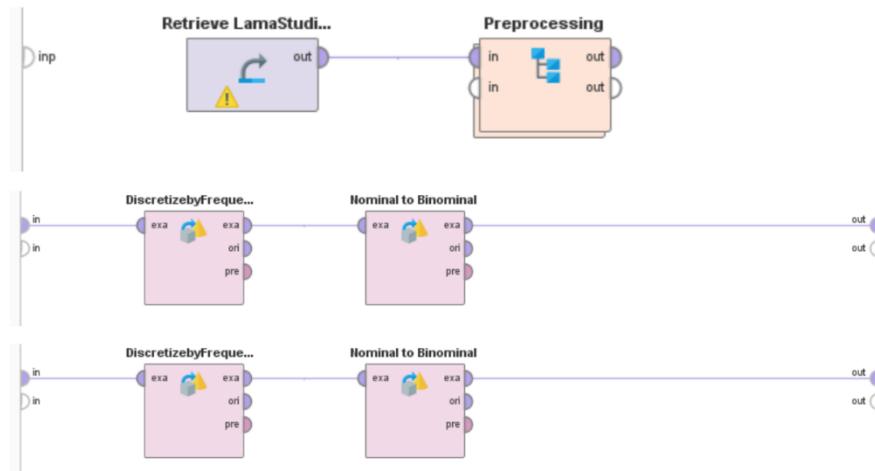


Tugas



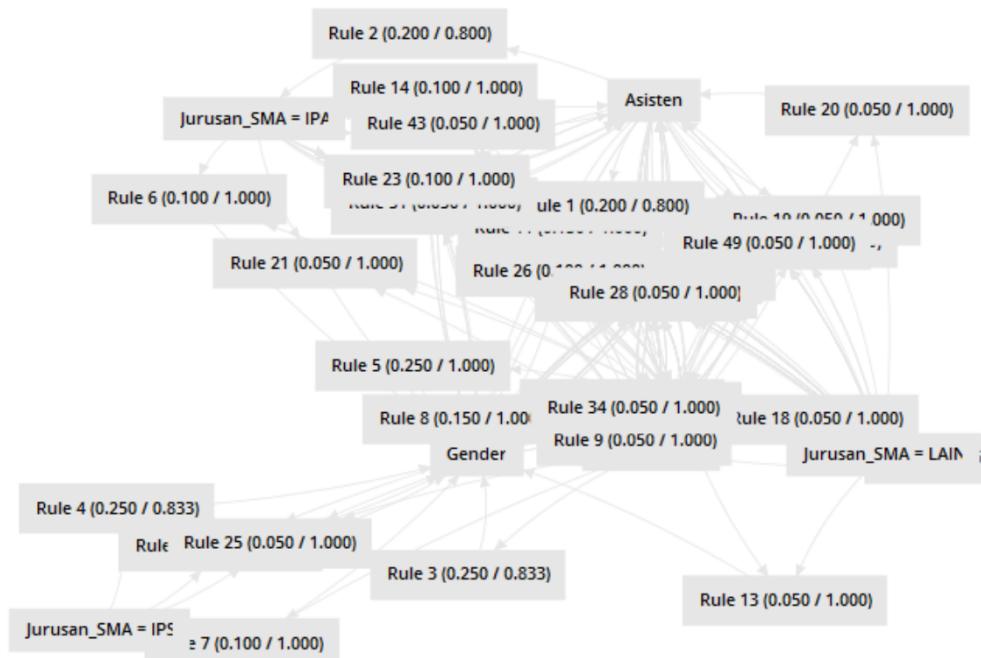
accuracy: 65.00% +/- 32.02% (micro average: 65.00%)

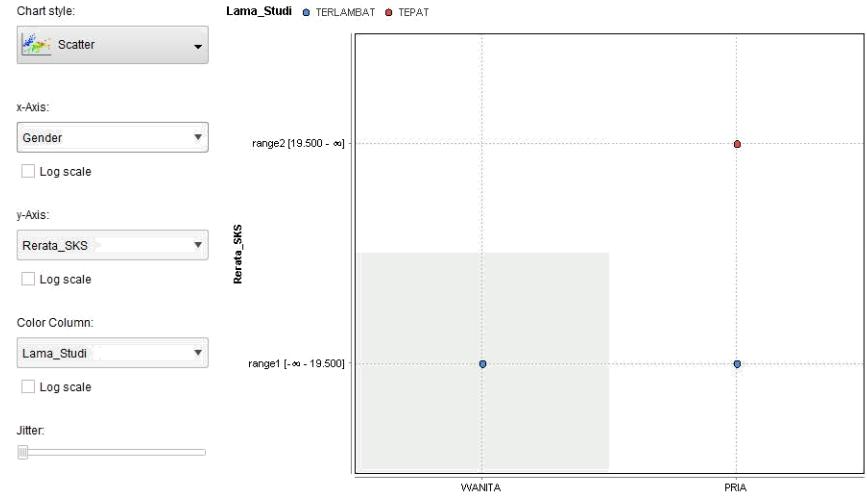
	true TERLAMBAT	true TEPAT	class precision
pred. TERLAMBAT	4	4	50.00%
pred. TEPAT	3	9	75.00%
class recall	57.14%	69.23%	



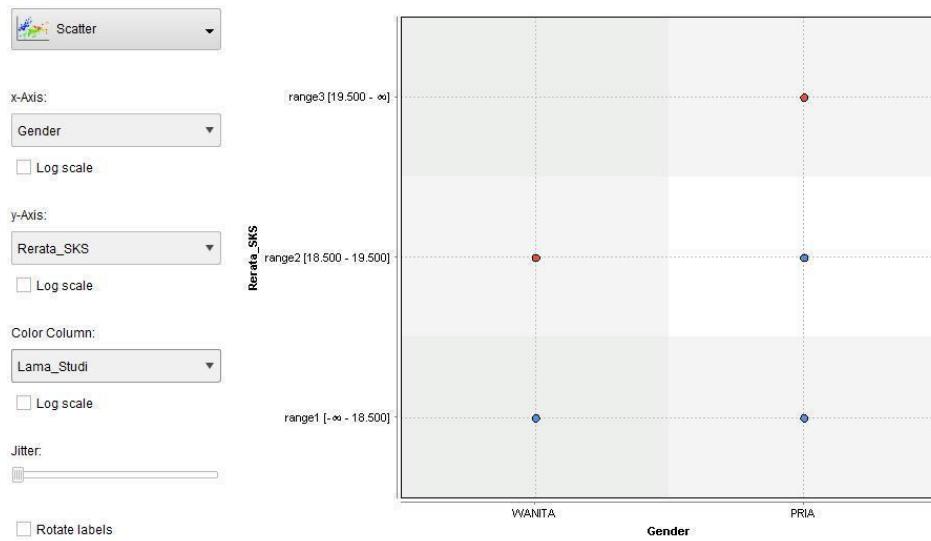
3	0.050	Gender	Asal_Sekolah	Jurusan_SMA = ...		
3	0.100	Gender	Jurusan_SMA = ...	Rerata_SKS		
3	0.150	Gender	Asisten	Rerata_SKS		
3	0.050	Gender	Asisten	Jurusan_SMA = ...		
3	0.050	Gender	Rerata_SKS	Jurusan_SMA = ...		
3	0.100	Jurusan_SMA = ...	Asal_Sekolah	Asisten		
3	0.050	Jurusan_SMA = ...	Asal_Sekolah	Rerata_SKS		
3	0.100	Jurusan_SMA = ...	Asisten	Rerata_SKS		
3	0.050	Asal_Sekolah	Jurusan_SMA = ...	Rerata_SKS		
3	0.100	Asal_Sekolah	Asisten	Rerata_SKS		
3	0.050	Asal_Sekolah	Asisten	Jurusan_SMA = ...		
3	0.050	Asal_Sekolah	Rerata_SKS	Jurusan_SMA = ...		
3	0.050	Asisten	Rerata_SKS	Jurusan_SMA = ...		
4	0.050	Gender	Jurusan_SMA = ...	Asal_Sekolah	Asisten	
4	0.050	Gender	Jurusan_SMA = ...	Asal_Sekolah	Rerata_SKS	
3	0.050	Gender	Asal_Sekolah	Jurusan_SMA = ...		
3	0.100	Gender	Jurusan_SMA = ...	Rerata_SKS		
3	0.150	Gender	Asisten	Rerata_SKS		
3	0.050	Gender	Asisten	Jurusan_SMA = ...		
3	0.050	Gender	Rerata_SKS	Jurusan_SMA = ...		
3	0.100	Jurusan_SMA = ...	Asal_Sekolah	Asisten		
3	0.050	Jurusan_SMA = ...	Asal_Sekolah	Rerata_SKS		
3	0.100	Jurusan_SMA = ...	Asisten	Rerata_SKS		
3	0.050	Asal_Sekolah	Jurusan_SMA = ...	Rerata_SKS		
3	0.100	Asal_Sekolah	Asisten	Rerata_SKS		
3	0.050	Asal_Sekolah	Asisten	Jurusan_SMA = ...		
3	0.050	Asal_Sekolah	Rerata_SKS	Jurusan_SMA = ...		
3	0.050	Asisten	Rerata_SKS	Jurusan_SMA = ...		
4	0.050	Gender	Jurusan_SMA = ...	Asal_Sekolah	Asisten	
4	0.050	Gender	Jurusan_SMA = ...	Asal_Sekolah	Rerata_SKS	

Support	Confidence	LaPlace	Gain	p-s	Lift	Convict...
0.250	0.833	0.962	-0.350	0.025	1.111	1.500
0.250	0.833	0.962	-0.350	0.025	1.111	1.500
0.250	1	1	-0.250	0.062	1.333	∞
0.100	1	1	-0.100	0.025	1.333	∞
0.100	1	1	-0.100	0.025	1.333	∞
0.150	1	1	-0.150	0.038	1.333	∞
0.050	1	1	-0.050	0.012	1.333	∞
0.100	1	1	-0.100	0.025	1.333	∞
0.150	1	1	-0.150	0.038	1.333	∞
0.050	1	1	-0.050	0.012	1.333	∞
0.050	1	1	-0.050	0.012	1.333	∞
0.100	1	1	-0.100	0.075	4	∞
0.050	1	1	-0.050	0.038	4	∞
0.050	1	1	-0.050	0.035	3.333	∞
0.050	1	1	-0.050	0.038	4	∞





Number of bins = 3



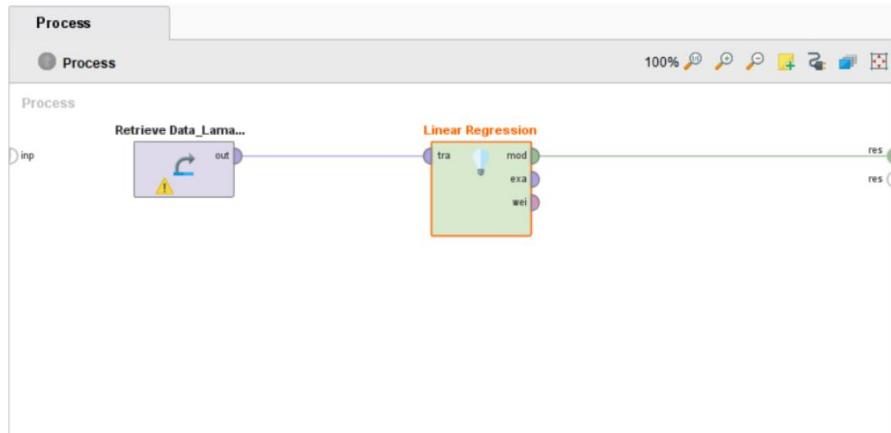
Modul 12 (Regresi Linier Sederhana)

Contoh Kasus :

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

1	No_Siswa	Nama	Lama Belajar (Jam)	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

1	No_Siswa polynomial id	Nama polynomial	Lama Belajar (Jam) integer	Nilai integer label
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



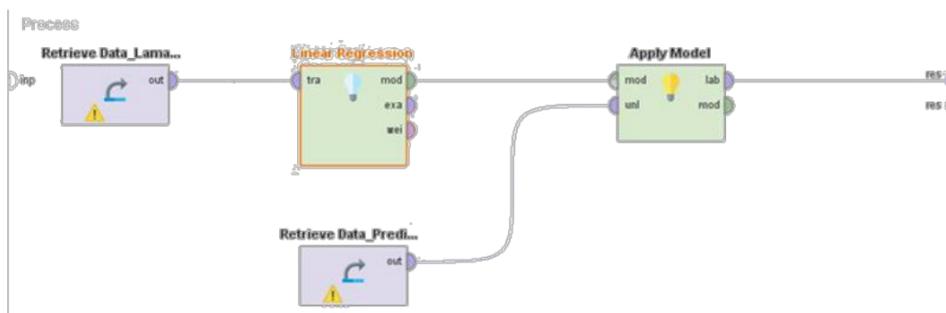
Attribute	Coefficient	Std. Error	Std. Coefficient	Tolerance	t-Stat	p-Value	Code
Lama Belajar (J...	21.608	7.645	0.707	1	2.827	0.022	**
(Intercept)	492.769	96.909	?	?	5.085	0.001	***

Linear Regression

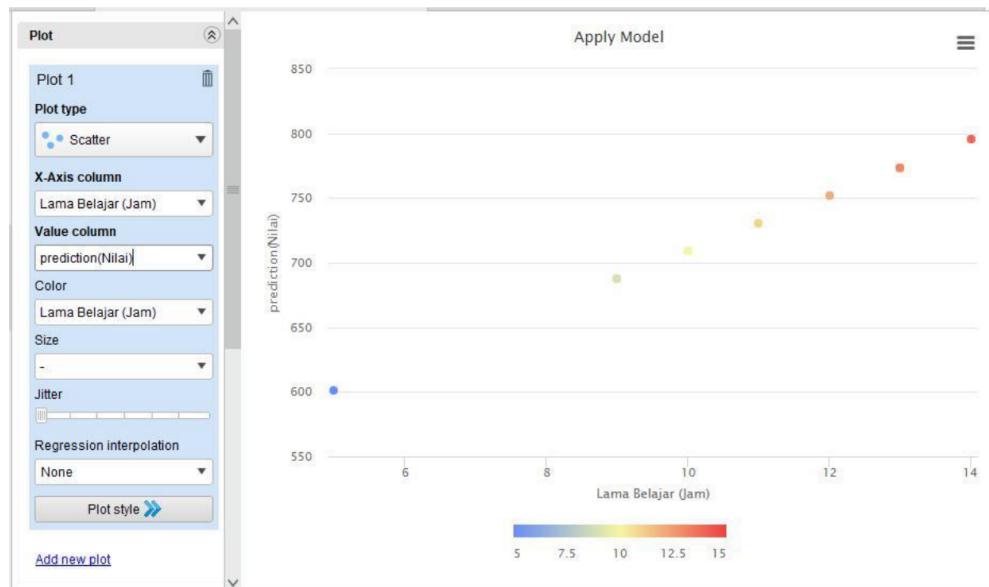
21.608 * Lama Belajar (Jam)
+ 492.769

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

1	No_Siswa polynomial id	Nama polynomial	Lama Belajar (Jam) integer
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



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



1	No_Siswa	Nama	Lama Belajar	Prediction (Nilai)	Prediction(Nilai)
				Tabel	Model Regresi
3	S-111	BUDI	12	752,061	752,065
4	S-112	SANTI	13	773,668	773,673
5	S-113	DIAN	14	795,276	795,281
6	S-114	DANI	11	730,453	730,457
7	S-115	AHMAD	5	600,807	600,809
8	S-116	BAYU	13	773,668	773,673
9	S-117	RISA	9	687,238	687,241
10	S-118	RANI	10	708,845	708,849
11	S-119	YANI	10	708,845	708,849
12	S-120	RATIH	9	687,238	687,241

TUGAS

Data Testing

1	NO RESPONDEN	PENDAPATAN (RUPIAH)	JUMLAH ANGGOTA KELUARGA	DAYA BELI (RUPIAH)
2	1	1000000	6	834000
3	2	1400000	7	1200000
4	3	200000	3	134000
5	4	1400000	6	1167000
6	5	500000	3	334000
7	6	1700000	5	1360000
8	7	400000	3	267000
9	8	1900000	5	1520000
10	9	300000	3	200000
11	10	500000	4	375000
12	11	700000	7	600000
13	12	1900000	3	1267000
14	13	800000	4	600000
15	14	1500000	4	1125000
16	15	1300000	7	1115000

Import Data - Format your columns.

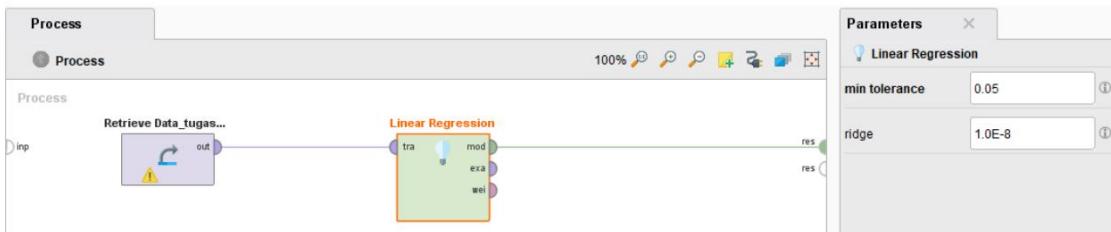
Format your columns.

Replace errors with missing values ⓘ

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.

Previous Next Cancel



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

LinearRegression

0.788 * PENDAPATAN (RUPIAH)
-7467.579

Import Data - Format your columns.

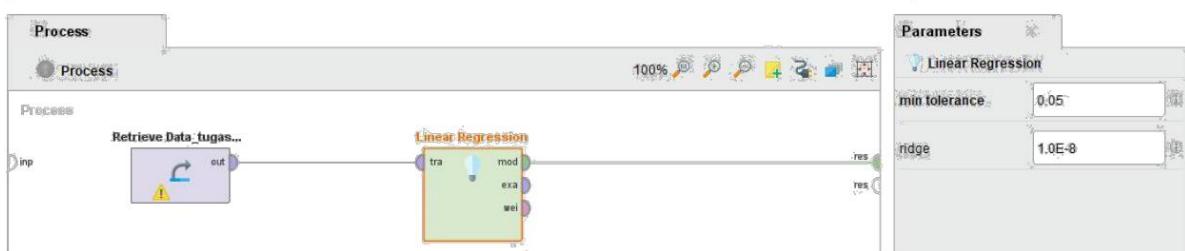
Format your columns.

Replace errors with missing values ⓘ

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

no problems.

◀ Previous ▶ Next ✖ Cancel



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

LinearRegression

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

Data Testing

1	NO RESPONDEN	PENDAPATAN (RUPIAH)	JUMLAH ANGGOTA KELUARGA	DAYA BELI (RUPIAH)
2	1	900000	5	834000
3	2	800000	3	1200000
4	3	500000	2	134000
5	4	1900000	6	1167000
6	5	600000	2	334000
7	6	800000	5	1360000
8	7	1000000	6	267000
9	8	1100000	4	1520000
10	9	1000000	4	200000
11	10	500000	3	375000

1	NO RESPONDEN	PENDAPATAN (RUPIAH)	JUMLAH ANGGOTA KELUARGA	DAYA BELI (RUPIAH)	X1	X2
2	1	900000	5	834000	701731,421	857096,225
3	2	800000	3	1200000	622931,421	553718,867
4	3	500000	2	134000	386531,421	402030,188
5	4	1900000	6	1167000	1489731,421	1008784,904
6	5	600000	2	334000	465331,421	402030,188
7	6	800000	5	1360000	622931,421	857096,225
8	7	1000000	6	267000	780531,421	1008784,904
9	8	1100000	4	1520000	859331,421	705407,546
10	9	1000000	4	200000	780531,421	705407,546
11	10	500000	3	375000	386531,421	55371,867

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

