

Nama : Yussynta Dewi Aprilya Putri

NIM : L200170084

Kelas : D

## MODUL12

### Mencari Nilai t-hitung dan Model Regresi Linier

1. Buka Ms.Excel, dan buatlah tabel data siswa. Simpan dengan nama **Tabel\_LamaBelajardanNilaiUjian.xls**.

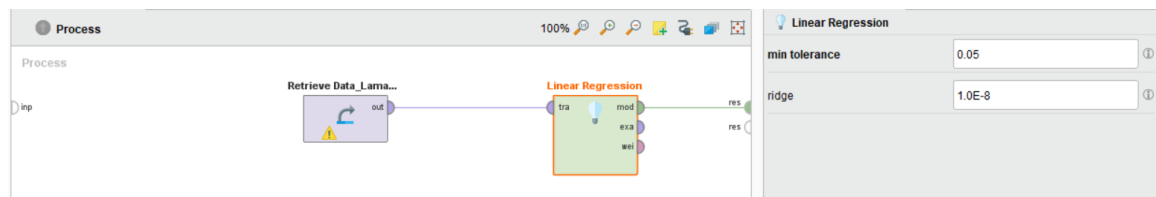
NO_SISWA	NAMA	LAMA BELAJAR (JAM)	NILAI
S-101	JOKO	15	783
S-102	AGUS	18	877
S-103	SUSI	7	505
S-104	DYAH	9	860
S-105	WATI	15	968
S-106	IKA	17	793
S-107	EKO	10	752
S-108	YANTO	5	571
S-109	WAWAN	8	667
S-110	MAHMUD	15	723

2. Jalankan RapidMiner.
3. Gunakan file **Tabel\_LamaBelajardanNilaiUjian.xls** sebagai data yang digunakan dalam proses Regresi Linier. Import file ini ke dalam repositori.
4. Ubahlah tipe data dan jenis masing-masing atribut sebagai berikut :

	NO_SISWA <i>polynominal id</i>	NAMA <i>polynominal</i>	LAMA BELAJAR (JAM) <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

5. Pada langkah ini, beri nama **Data\_LamaBelajardanNilaiUjian** dan masukkan pada repository, klik **Finish**.
6. Gunakan Data\_LamaBelajardanNilaiUjian ini dan masukkan ke dalam area process.
7. Tambahkan operator **Modelling→Predictive→Functions→Linier Regression**.  
Hubungkan port sesuai pada gambar.

8. Klik pada operator **Linier Regression**, tentukan paramete **min tolerance = 0.05**.



9. Jalankan proses dengan menekan tombol Run (atau F11).

10. Hasil proses regresi linier :

- a. Tabel View (mencari besarnya nilai t-hitung)

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

- b. Text View (mencari model regresi)

## LinearRegression

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

Model regresi linier yang terbentuk :

$$Y = 21,608 X1 + 492,769$$

Dengan model tersebut, dapat dicari Nilai Ujian (Y) dengan memasukkan nilai Lama Belajar pada variabel X1.

NO_SISWA	NAMA	LAMA BELAJAR (JAM)	NILAI	Nilai Belajar
S-101	JOKO	15	783	816,889
S-102	AGUS	18	877	881,713
S-103	SUSI	7	505	644,025
S-104	DYAH	9	860	687,241
S-105	WATI	15	968	816,889
S-106	IKA	17	793	860,105
S-107	EKO	10	752	708,849
S-108	YANTO	5	571	600,809
S-109	WAWAN	8	667	665,633
S-110	MAHMUD	15	723	816,889

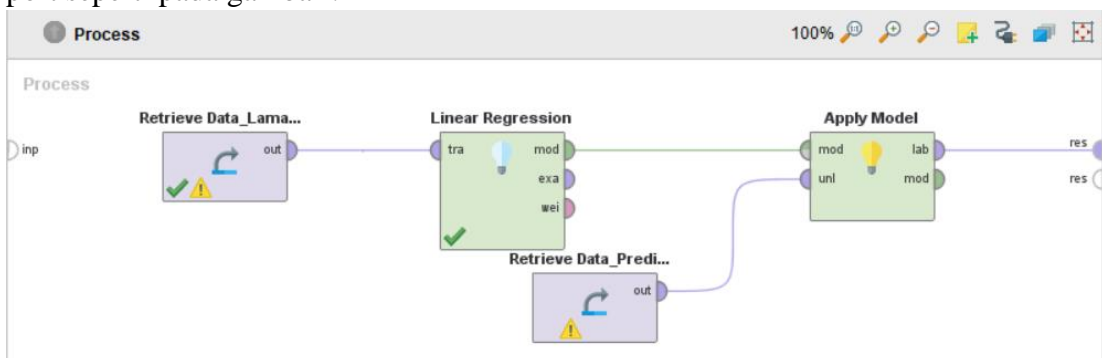
### Mencari Nilai t- dan Model Regresi Linier Menggunakan RapidMiner

1. Buka Ms.Excel, dan buatlah tabel data siswa tersebut. Simpan dengan nama **Tabel\_PrediksiNilaiUjian.xls**
2. Jalankan **RapidMiner**.

- Gunakan file **Tabel\_PrediksiNilaiUjian.xls** sebagai data testing. Import file ke dalam repository.
- Ubah tipe data dan jenis masing-masing atribut sebagai berikut :

	NO_SISWA <i>polynominal id</i>	NAMA <i>polynominal</i>	LAMA BELAJAR (JAM) <i>Integer</i>
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

- Pada langkah ini, beri nama **Data\_PrediksiNilaiUjian** dan masukkan pada repository, klik **Finish**.
- Tetap menggunakan proses pada kegiatan sebelumnya, masukkan **Data\_PrediksiNilaiUjian** ke dalam area process.
- Tambahkan **Scoring**→**Apply Model**, letakkan setelah Linier Regression. Hubungkan port seperti pada gambar :

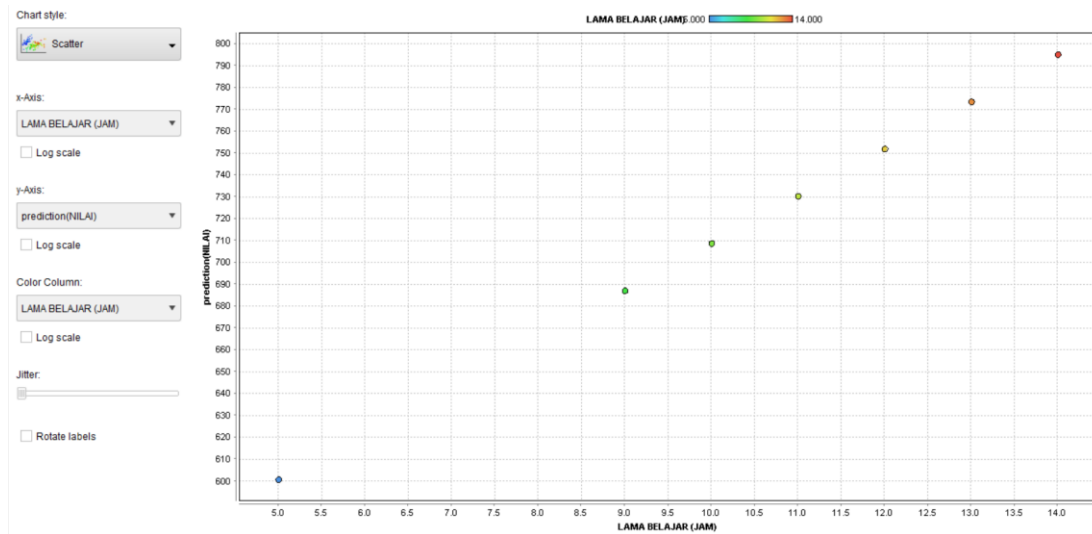


- Dengan menggunakan parameter yang sama pada operator **Regrission Linear**, jalankan proses dengan tombol **Run**.
- Hasil proses prediksi terhadap data testing menggunakan regresi linear :
  - Data View (hasil prediksi nilai ujian)

ExampleSet (10 examples, 2 special attributes, 1 regular attribute)

Row No.	NO_SISWA	prediction(N...	LAMA BELA...
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

b. Chart View (Scatter Plot)



Pembuktian Model Regresi

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

NO_SISWA	NAMA	LAMA BELAJAR (JAM)	Prediction (NILAI)	Prediction (NILAI)
			Tabel	Model Regresi
S-111	BUDI	12	752,061	752,065
S-112	SANTI	13	773,668	773,673
S-113	DIAN	14	795,276	795,281
S-114	DANI	11	730,453	730,457
S-115	AHMAD	5	600,807	600,809
S-116	BAYU	13	773,668	773,673
S-117	RISA	9	687,238	687,241
S-118	RANI	10	708,845	708,849
S-119	YANI	10	708,845	708,849
S-120	RATIH	9	687,238	687,241

Dapat dilihat bahwa hasil prediksi menggunakan RapidMiner menghasilkan nilai yang sama dengan menggunakan Model Persamaan Regresi Linier.

## Tugas

1. Buatlah tabel dengan menggunakan Microsoft Excel seperti berikut :

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

2. Buatlah proses Regresi Linier Sederhana menggunakan RapidMiner dengan ketentuan :
  - a. Variable Bebas(X) = Pendapatan(X1), Jumlah Anggota Keluarga(X2)
  - b. Variable Terikat(Y) = Daya Beli
  - c. Toleransi yang digunakan = 5%

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

## LinearRegression

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

3. Tentukan apakah variable X1 dan X2 mempengaruhi secara signifikan terhadap nilai variable Y berdasarkan nilai t-stat?  
 Dari tabel tersebut dapat dilihat nilai t-stat sebesar 2,142. Berdasarkan aturan statistik, variabel X dikatakan mempengaruhi secara signifikan terhadap Y jika nilai t-stat > t-table.
4. Tuliskan model persamaan regresi linier sederhana yang terbentuk!  

$$Y = 151688,679 X2 + 98652,830$$

5. Gunakan data testing untuk menjawab perintah :
  - a. Lakukan prediksi Daya Beli (Y) dengan menggunakan Model Persamaan Regresi Linier dari hasil pertanyaan nomor 4!
  - b. Lakukan prediksi Daya Beli (Y) menggunakan RapidMiner!

NO. RESPONDEN	PENDAPATAN (RUPIAH)	JUMLAH ANGGOTA KELUARGA	Prediction (NILAI)	
			Model Regresi	Tabel
1	900000	5	857096,225	857096,226
2	800000	3	553718,867	553718,868
3	500000	2	402030,188	402030,189
4	1900000	6	1008784,904	1008784,906
5	600000	2	402030,188	402030,189
6	800000	5	857096,225	857096,226
7	1000000	6	1008784,904	1008784,906
8	1100000	4	705407,546	705407,547
9	1000000	4	705407,546	705407,547
10	500000	3	553718,867	553718,868

6. Gambarkan pola sebaran data menggunakan Plot View (Scatter)

