

Nama : Oki Kus Mahesa

Nim : L200170064

Kelas : C

Modul 12

REGRESSI LINIER

- Tugas

1. Kasus: Dalam Sebuah Kepala Keluarga telah diperoleh variable pendapatan rata-rata perbulan, jumlah anggota keluarga yang tinggal serumah, dan daya beli rata-rata perbulan.
2. Bagaimana model Regresi Linier yang terbentuk, dan lakukan prediksi terhadap 10 data yang belum diketahui nilai daya belinya.
3. Tabel Hasil Survei 15 Kepala Keluarga

	A	B	C	D
1	NO.RESPONDEN	PENDAPATAN (RUPIAH)	JUMLAH ANGGOTA KELUARGA	DAYA BELI (RUPIAH)
2	1	1,000,000	6	834,000
3	2	1,400,000	7	1,200,000
4	3	200,000	3	134,000
5	4	1,400,000	6	1,167,000
6	5	500,000	3	334,000
7	6	1,700,000	5	1,360,000
8	7	400,000	3	267,000
9	8	1,900,000	5	1,520,000
10	9	300,000	3	200,000
11	10	500,000	4	375,000
12	11	700,000	7	600,000
13	12	1,900,000	3	1,267,000
14	13	800,000	4	600,000
15	14	1,500,000	4	125,000
16	15	1,300,000	7	1,115,000

4. Buatlah proses Regresi Linier Sederhana menggunakan RapidMiner dengan ketentuan sebagai berikut.
 - a. Variable bebas (X) = Pendapatan (X1), Jumlah Anggota Keluarga (X2)
 - b. Variable terikat (Y) = Daya Beli
 - c. Toleransi yang digunakan = 5%

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

no problems.

5. Tentukan apakah variable x1 dan x2 mempengaruhi secara signifikan terhadap nilai Y berdasarkan besarnya nilai t-stat?

a. Table View (Mencari besarnya nilai t-hitung)

Attribute	Coefficient	Std. Error	Std. Coefficient	Tolerance	t-Stat	p-Value	Code
PENDAPATAN (R...	0.605	0.131	0.723	0.857	4.625	0.001	****
JUMLAH ANGGO...	85541.767	48568.994	0.275	0.857	1.761	0.104	
(Intercept)	-284291.899	226294.413	?	?	-1.256	0.233	

Dari tabel tersebut dapat dilihat nilai t-statistic (t-hitung) sebesar:

X1 = Pendapatan = 4.625

X2 = Jumlah Anggota Keluarga = 1.761

Berdasarkan aturan statistik, variable X1 dan X2 dikatakan mempengaruhi secara signifikan terhadap variable Y jika nilai t-hitung > t-table,

Yang artinya bahwa "Pendapatan(x1) dan Jumlah Anggota Keluarga(x2) dapat Mempengaruhi Daya Beli(y).

b. Text View (mencari model regresi)

LinearRegression

```
0.605 * PENDAPATAN (RUPIAH)
+ 85541.767 * JUMLAH ANGGOTA KELUARGA
- 284291.899
```

Berdasarkan hasil pada Text View terlihat sebuah persamaan linier berikut:

$$= 0.605 * \text{PENDAPATAN (RUPIAH)} + 85541.767 * \text{JUMLAH ANGGOTA KELUARGA} - 284291.899$$

Maka dapat dibuat sebuah model persamaan regresi linier sederhana untuk mencari nilai variable y (Daya Beli) berdasarkan variable x_1 (Pendapatan) dan variable x_2 (Jumlah Anggota Keluarga).

Berikut model regresi linier yang terbentuk:

$$Y = 0.605 x_1 + 85541.767 x_2 - 284291.899$$

Dengan model tersebut, dapat dilihat dicari Daya Beli(Y) dengan memasukan nilai Pendapatan pada variable x_1 dan pada variable x_2

6. Gunakan data testing untuk menjawab perintah berikut:

- a. Lakukan prediksi Daya Beli(Y) dengan menggunakan Model Persamaan Regresi linier dari hasil pertanyaan nomor 5!
 - b. Lakukan prediksi Daya Beli (Y) menggunakan RapidMiner!
- Data Testing yang digunakan untuk prediksi :

	A	B	C
1	NO.RESPONDEN	PENDAPATAN (RUPIAH)	JUMLAH ANGGOTA KELUARGA
2	1	900,000	5
3	2	800,000	3
4	3	500,000	2
5	4	1,900,000	6
6	5	600,000	2
7	6	800,000	5
8	7	1,000,000	6
9	8	1,100,000	4
10	9	1,000,000	4
11	10	500,000	3

- Mengubah Tipe Data dan jenis masing-masing atribut sebagai berikut:
 - a. NO.RESPONDEN: integer,id
 - b. PENDAPATAN: integer
 - c. JUMLAH ANGGOTA KELUARGA: integer

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

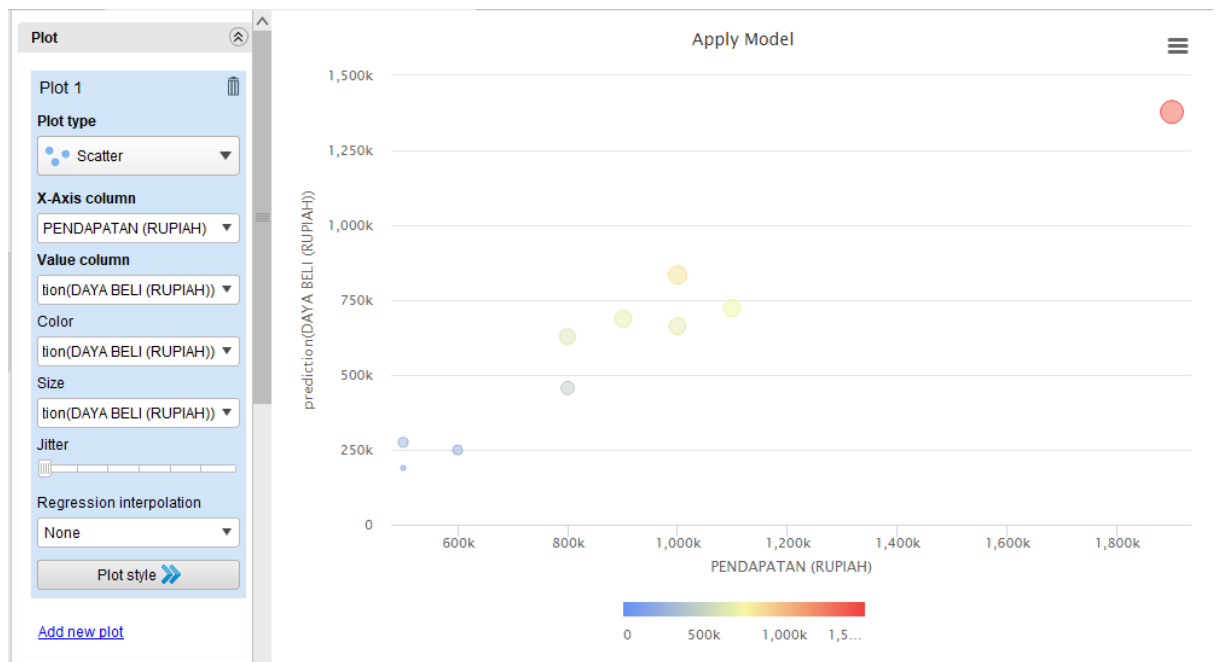
- Data View (Hasil Prediksi Daya Beli)

NO.RESPONDEN	PENDAPATAN (RUPIAH)	JUMLAH ANGGOTA KELUARGA	prediction(DAYA BELI (RUPIAH))
1	900000	5	687740.118
2	800000	3	456176.230
3	500000	2	189193.402
4	1900000	6	1378085.419
5	600000	2	249673.756
6	800000	5	627259.765
7	1000000	6	833762.239
8	1100000	4	723159.057
9	1000000	4	662678.704
10	500000	3	274735.170

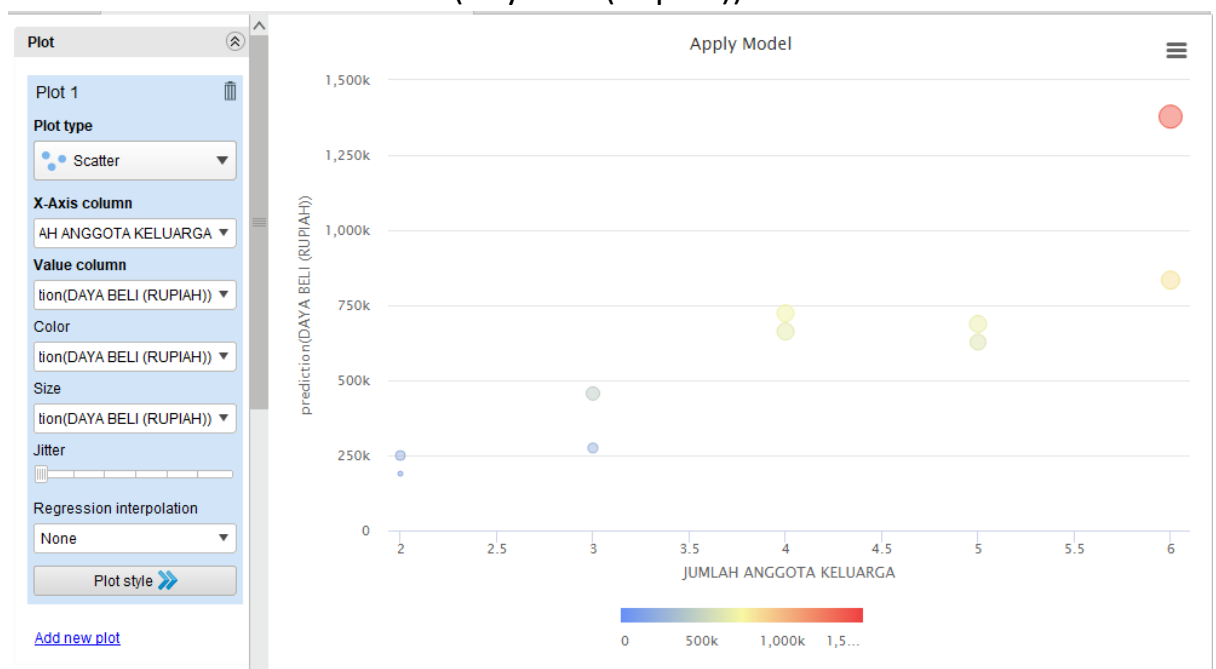
Hasil prediksi Daya Beli terhadap 10 Kepala Keluarga lainnya.

7. Pola sebaran data menggunakan Plot View (Scatter) dengan ketentuan:

- a. X-axis = Pendapatan (Rupiah),
y-axis = Prediction (Daya Beli(Rupiah)),
Color Column = Prediction (Daya Beli(Rupiah))



- b. X-axis = Jumlah Anggota Keluarga,
y-axis = Prediction(Daya Beli(Rupiah)),
Color Column = Prediction (Daya Beli(Rupiah))



Berdasarkan Plot View dapat dilihat bahwa:

- Semakin Besar Jumlah Pendapatan(X1) Maka Jumlah Daya Beli(Y) akan semakin tinggi.
- Semakin Banyak Jumlah anggota Keluarga(X2) Maka Jumlah Daya Beli(Y) akan semakin tinggi